



Cisgender Women Underestimate Support for the Inclusion of Transgender Women in the Group ‘Women’: An Investigation of Norm Misperceptions

Christina Maxwell¹ · Hema Preya Selvanathan¹ · Jolanda Jetten¹ · Charlie R. Crimston^{1,2}

Accepted: 1 August 2025 / Published online: 26 August 2025
© The Author(s) 2025

Abstract

Individuals sometimes misjudge the views of others in their group, believing that their own attitudes differ from those held by the majority. Across five studies ($N=1159$) conducted in Australia and the United Kingdom, we assessed the similarity between cisgender women’s own beliefs and their perceptions of the group norm (i.e., other cisgender women) regarding the inclusion of transgender women within the boundaries of womanhood. We found robust evidence that cisgender women in our sample underestimated the level of support for inclusion of trans women among their group relative to their own beliefs. Brief interventions that communicated norm information were shown to be effective in increasing participants’ perceptions of the norm. However, overall, the intervention did not affect willingness to express inclusion opinions or engage in actions to support trans women’s rights and their social and legal inclusion into the group ‘women.’ These findings highlight the extent to which group norms surrounding trans inclusion can be misperceived and suggest that correcting these misperceptions may help counter the influence of exclusionary minority voices.

Keywords Perceived norms · Pluralistic ignorance · Trans women · Ingroup membership · Trans exclusion · Social identity · Social norms · Gender identity · Inclusion

There has been considerable debate among the cisgender (cis) public about what it means to be a woman, centred on whether transgender (trans) women should be socially and legally included in the group ‘women.’ Two cultural contexts in which this debate about inclusion has been gaining attention includes Australia and the United Kingdom. Notably, British author JK Rowling has advocated for exclusionary group boundaries determined by “biological sex” (see Gardner, 2023, for an overview). Based on this definition, trans women are denied entry to the group ‘women.’ Australian and British political candidates have also organised campaigns highlighting conflict between trans women’s rights and cis women’s rights (John, 2022; Martin & McGowan, 2022). Furthermore, news media regularly platform exclusionary

voices that depict cis women as being concerned about the inclusion of trans women within traditionally single-sex cis women’s spaces, such as prisons, bathrooms, and sports (García & Badge, 2021; LGBTIQ+ Health Australia, 2022; Pearce et al., 2020).

Despite media that promote a conflictual relationship between cis and trans women, most cis women may not actually be opposed to including trans women within the category of women. Emerging literature in the United Kingdom has shown that cis women are supportive of sharing single-sex spaces with trans women (Pain et al., 2021; Pringle & Denison, 2022). Research in Australia is less forthcoming but shows similar trends within sport (Pringle & Denison, 2022). There are also online campaigns by feminist groups in both the United Kingdom and Australia (Engender, 2023; Hislop, 2022) who protest the rise of anti-trans policies that disproportionately impact trans women under the guise of protecting cis women.

Exclusionary narratives may thus be driven by a small but vocal minority of cis women who receive a large amount of media attention (Hines, 2019; Serano, 2018). Despite being

✉ Christina Maxwell
christina.maxwell@student.uq.edu.au

¹ School of Psychology, The University of Queensland, Brisbane, QLD 4072, Australia

² The Australian National University, Canberra, Australia

a numerical minority, the dominance of these marginal voices in the public domain indicates that their involvement in this issue is not benign. As individuals' own beliefs and behaviours are influenced by the social norms around them (i.e., the perceived standards of behaviours within a group; Turner et al., 1987), mistaking vocality for representativeness may have consequences for those who hold majority opinions.

In the current research, we propose that the concept of pluralistic ignorance can offer valuable insights into current debates concerning whether cis women should include trans women within the group 'women.' Pluralistic ignorance refers to the mistaken belief that one's opinions and behaviours differ from the group norm (Sargent & Newman, 2021). Across five studies conducted in the United Kingdom and Australia, we hypothesise that cis women underestimate the level of support for inclusion and that this misperception of group norms (i.e., pluralistic ignorance) can suppress motivations to advocate for the expansion of group boundaries. We then test whether pluralistic ignorance can be corrected with a brief, targeted intervention that can subsequently affect collective action intentions and behaviours.

The Social Norms of Inclusion

Social norms refer to the range of behaviours that are commonly accepted (i.e., injunctive norms) and/or performed (i.e., descriptive norms) within a group (Cialdini et al., 1991). They provide clarity around what it means to be a group member and how they should behave. It is through categorisation of oneself as belonging to a particular group that the norms of that group can exert influence on behaviour, as posited by self-categorisation theory (SCT; Turner et al., 1987). These norms can be further used to emphasise similarities within the group ("us") and differences between other groups ("them"). SCT complements social identity theory (SIT; Tajfel & Turner, 1979) and both theoretical traditions are referred to as the social identity approach (Turner & Reynolds, 2001). According to the social identity approach, groups are motivated to engage in favourable intergroup comparisons that contribute to a positive ingroup identity. To aid in this process, groups construct, maintain, and protect their boundaries that dictate ingroup eligibility.

However, group boundaries can vary in permeability, that is, the ease with which intergroup movement can occur and that outgroup members are accommodated into the ingroup. Past research in this area has addressed the attitudes and actions of the incoming outgroup (who are often lower in status than the receiving ingroup) as they engage in mobility attempts (e.g., Lalonde & Silverman, 1994; Tajfel & Turner, 1979). However, this framing ignores the influence

that the receiving group has in gatekeeping boundaries by limiting or preventing opportunities for intergroup movement. Research that has turned the investigative lens onto the receiving group has focused on this group's perceptions of or attitudes toward experimenter-manipulated or pre-existing boundaries (e.g., Johnson et al., 2005; Richard & Wright, 2010; Verkuyten & Reijerse, 2008; Zhang et al., 2014) rather than acknowledging their agentic role in determining boundary permeability. For instance, enforcement of boundaries at the intergroup level might involve establishing inclusion criteria or social norms (e.g., pertaining to behaviours, values, or physical attributes), instigating outgroup hostility, or engaging in campaigns and collective action efforts (Bennett, 2014; Ryan & Reicher, 2019; Travers & Deri, 2011). Understanding the involvement of receiving groups in intergroup movement attempts can shed light on intergroup relations as well as potential opportunities for receiving group members to act as allies to encourage inclusivity. The current research therefore extends SIT to understand this active role that the receiving group plays in the social mobility efforts of the incoming group members through boundary maintenance within the context of 'womanhood.'

In social identity terms, the current debates surrounding the boundaries of womanhood indicate instability and permeability of the group. That is, there are definitions of 'women' based on self-identification and shared experiences of oppression (Pain et al., 2021), while other definitions stipulate access solely via biology (Armitage, 2020). Regardless, as the category of 'women' is still often understood as being synonymous with 'cis women,' to what degree are cis women willing to expand the boundaries of womanhood to encompass trans women within category definitions? Emerging research shows evidence of cis women supporting the inclusion of trans women within single-sex women's spaces. For example, Auerbach and colleagues (2020) reported that 78% of their cis women participants living with HIV in the United States encouraged the integration of trans women within their single-sex healthcare services. Support was also found among cis women regarding abuse support services in the United Kingdom (Pain et al., 2021), women's rugby teams in the United Kingdom and Australia (Pringle & Denison, 2022), and women's colleges in the United States (Outten & Lawrence, 2020). Together, this evidence suggests a descriptive norm of inclusion among cis women in certain contexts (Cialdini et al., 1991). However, the actual norm is not always accurately perceived by ingroup members. In the following section, we describe one of the systematic errors group members make when perceived norms conflict with actual norms, resulting in an effect known as pluralistic ignorance.

Pluralistic Ignorance

Pluralistic ignorance refers to a situation in which a perceiver incorrectly believes “others’ cognitions (attitudes, beliefs, feelings) and/or behaviors differ systematically from their own (i.e., a directional misperception)” (Sargent & Newman, 2021, p. 166, see also Miller & McFarland, 1987). Within group settings, pluralistic ignorance occurs when group members inaccurately come to believe that ingroup norms are incongruent with their own personal beliefs. Pluralistic ignorance has been documented across a range of contexts (see Sargent & Newman, 2021, for a review), including attitudes towards the lesbian, gay, and bisexual (LGB) community (Bowen & Bourgeois, 2001; Dubuque et al., 2002; Eisner et al., 2021). In this research, individuals incorrectly believed that they held more positive attitudes towards the LGB community than others. The identification of pluralistic ignorance is therefore an important first step to understanding individual and group behaviour.

The reasons why the norm is misperceived are varied (e.g., see Tankard & Paluck, 2016, for an overview), although one oft-cited cause is the news media (Noelle-Neumann, 1993; Perryman et al., 2018). Within the context of LGB attitudes, the news media has a history of publishing fear-mongering narratives and negative depictions of the LGB community (Jacobs & Meeusen, 2021; Steiner et al., 1993). However, the growing acceptance of LGB rights has resulted in the reduction of this negative media slant and its subsequent transference to the transgender community. Specifically, media coverage has focused on the accommodation of trans women into single-sex ‘women’s’ spaces (García & Badge, 2021). As such, the pluralistic ignorance effect previously observed for LGB attitudes may be extended to beliefs regarding the inclusion of trans women into the group ‘women.’

Misperceived norms can also be caused by vocal minorities (Halbesleben et al., 2007). In the context of trans women’s inclusion, these vocal minorities include trans exclusionary radical feminists (TERFs), whose mobilisation efforts both online and offline have intensified in recent years (Pearce et al., 2020; Willem et al., 2022). In the United Kingdom, TERF rhetoric is amplified by influential cis women such as JK Rowling who has considerable public followership and regularly contributes to a narrative of exclusion that is broadcast by the media (Gardner, 2023). The frequent platforming of these vocal minorities by the media may lead to confusing availability of opinion with popularity of opinion (Halbesleben et al., 2007). As such, this literature also underscores the heterogeneity within receiving groups (particularly within the group ‘women’) and the subsequent competition over defining

group boundaries, with the dominant narrative ultimately shaping collective perceptions of boundary permeability.

Together with previously presented evidence that cis women may support inclusion, this conflict between the publicised norm and personal beliefs presents the ideal conditions for pluralistic ignorance to form. The first aim of this paper is therefore to examine whether cis women underestimate the degree to which other cis women hold inclusionary beliefs towards trans women (i.e., the perceived norm), resulting in pluralistic ignorance.

Consequences of Norm Misperception

When issues are contested, such as the debate around trans women’s inclusion into the group ‘women,’ people may look to social norms as a guide for their own opinions and behaviour. Group members engage in comparative processes between their own behaviours and the perceived norms (regardless of the accuracy of this perception), amending their behaviour accordingly if inconsistencies are observed (Turner et al., 1987). By conforming one’s behaviour to group norms, social sanctions can be avoided while group-based identities can be maintained and strengthened (Miller & McFarland, 1987; Noelle-Neumann, 1993). However, when perceived norms are inaccurate (i.e., there is pluralistic ignorance), individuals may suppress their true opinions and actions, which can reinforce the status quo.

The suppression of norm-inconsistent behaviour can have important implications. For example, bystander literature provides examples of how prosocial behaviour is reduced when a norm of inaction is perceived (Latané & Darley, 1970; Prentice & Miller, 1993). Relatedly, individuals who believe themselves to hold opinions incongruent with the perceived norm may opt for self-censorship. When minority voices are publicised as representing majority opinion, a ‘spiral of silence’ (Noelle-Neumann, 1993) effect may occur such that the majority suppresses their views to remain in line with the perceived social norm, thereby strengthening the voices of the minority opinion. Research has shown that people are less likely to express their views when they perceive a lack of majority support (Nekmat & Gozenbach, 2013; Portelinha & Elcheroth, 2016).

The second aim of this paper is to test whether underestimating support for inclusion inhibits cis women’s motivations to engage in collective action around inclusion both online and offline (e.g., speaking out about inclusion, attending protests, signing petitions, and supporting policy that protects trans women’s rights). We also examine whether the norm influences perceptions of collective efficacy (van Zomeren et al., 2008), that is, the belief that cis women as a group can successfully push for change around the inclusion of trans women. Specifically, cis women who perceive low

ingroup support for their inclusion opinions may believe that changing the boundaries of womanhood is an unachievable goal.

Correcting Norms for Behaviour Change

Perceptions of the norm are malleable and are constantly revised in response to new information (Tankard & Paluck, 2016). This feature offers promising opportunities for interventions to be developed and implemented that could correct norm perceptions and therefore reduce pluralistic ignorance. One commonly used method to change norm perceptions is group summary intervention (Yamin et al., 2019) in which accurate norm information (i.e., average rates of a behaviour or endorsement of a belief) is communicated to group members.

The group summary approach has successfully changed perceptions about group norms across a variety of contexts. For instance, previous studies have shown corrective rates of behaviour and beliefs to participants via pie charts and posters to change norm perceptions across a range of contexts, including college drinking (Mattern & Neighbors, 2004; Vallentin-Holbech et al., 2018), gender bias in the fields of science, technology, engineering, and mathematics (De Souza & Schmader, 2021), mental health service use (Sargent, 2021), and bullying in schools (Perkins et al., 2011). Norm correction interventions can also affect behaviour; donation behaviour and policy support for climate change initiatives increase when norms about climate change are corrected (Andre et al., 2022), and men's allyship behaviour towards women increases when they receive corrected norms about other men's endorsement of gender equality (Bursztyjn et al., 2020). Corrected norms can effectively counteract the spiral of silence by increasing willingness to speak out about divisive issues both in-person (Geiger & Swim, 2016) and online (Nekmat & Gozenbach, 2013).

In Studies 3–5, we employ a group summary approach to provide cis women with previously collected data regarding group norms on the inclusion of trans women. We expect that pluralistic ignorance surrounding inclusion will be reduced following this intervention. It should be noted that as social norms develop over time through several sources, corrective interventions are deemed successful if they *reduce*, but not necessarily eliminate, the norm misperception (Miller & Prentice, 2016).

We also test whether receiving corrective norm information affects willingness to express opinions about trans women's inclusion. Specifically, we predict that participants who view corrective information about the inclusion norm may have lower or higher intentions to engage in speaking out behaviours as they reassess their position in the group

from holding minority or majority opinions, relative to their own inclusion beliefs.

Overview of Current Research

Social norms, accurate or not, communicate important information about the societal climate of inclusion. We postulate that how public support is perceived regarding the inclusion of trans women (i.e., the norm) may impact the motivation of cis women to engage in behaviours that support this inclusion. Across five studies in the United Kingdom and Australia, we investigated the degree to which pluralistic ignorance occurs, if at all, regarding cis women's beliefs on whether trans women should be included into the group 'women.' By focusing on cis women as the receiving ingroup (relative to the incoming group of trans women), we intend to advance current theorising on boundary maintenance to examine the role of gatekeepers in intergroup mobility. To do so, we developed a new measure to assess inclusion beliefs for oneself and the perceived inclusion norm, that is, the beliefs of other cis women. We also examined the effects of norm misperceptions on behaviours intended to support the inclusion of trans women.

Study 1 was a correlational proof-of-concept investigation to determine whether pluralistic ignorance existed for this issue among the focal population (cis women). The second study sought to replicate Study 1, while beginning to examine the effects of norm misperceptions on cis women's willingness to express their opinions about the inclusion of trans women. The final three studies employed a norm correction intervention using a post-test randomised experimental design (Studies 3 and 4) and a within-participants longitudinal experimental design (Study 5). The first two studies were exploratory and not pre-registered. The remaining studies were pre-registered (Study 3: https://aspredicted.org/FKZ_MZF, Study 4: https://aspredicted.org/FZB_TK1, Study 5: https://aspredicted.org/75F_5PB). The supplemental material, data, analysis code, and materials for all studies are available at https://osf.io/d5wga/?view_only=e2a3b40df7f04e2caa54806309eeae6e.

Study 1

Study 1 investigated pluralistic ignorance within the context of cis women's inclusion of trans women using two exploratory research questions: RQ1) What are the personal inclusion beliefs of a subsample of Australian cis women? and RQ2) Are these average inclusion beliefs aligned with participants' estimations of the inclusion norm, that is, the degree to which participants believed other cis women supported the inclusion of trans women within the group 'women?'

Method

Participants and Procedure

Australian cis women ($M_{age} = 33.05$ years; $SD_{age} = 11.68$ years) were recruited via Prolific. We sought to recruit 300 participants in line with past research (Dubuque et al., 2002; Outten et al., 2019). One participant did not identify as a cis woman, resulting in a total of 299 responses. A sensitivity analysis using G*power (version 3.1.9.7; Faul et al., 2007) with alpha set at .05, power set at 80%, and a sample size of 299 determined that we could detect an effect size d of 0.16 using a paired-samples t -test. Participants were predominantly White ($n=224$, 74.9%) and heterosexual ($n=210$, 70.2%) (see Supplemental Material in the OSF for participant demographics for all studies). For Study 1 and the studies that followed, all participants received a detailed participant information sheet before providing informed consent. Ethics approval for all studies was granted from the first author's University Ethics Committee.

Measures

Inclusion Beliefs

To measure the degree of inclusion of trans women into the group 'women' desired by cis women, we developed a new measure based on Aron et al.'s (1992) Inclusion of Other in the Self Scale and Schubert and Ottens's (2002) Assessment of Ingroup-Outgroup Overlap Scale. Participants were presented with five pairs of circles (of which they were directed to select one pair) in various stages of overlap where each circle represented the boundaries of either the group 'women' or 'trans women' (Fig. 1). This novel measure aimed to understand participants' conceptions of the boundaries of womanhood in terms of the overlap between the superordinate category of 'woman' with the category 'trans women.' Responses were recorded on a 5-point scale where higher scores indicated greater overlap between groups and greater desired inclusivity.

The inclusion beliefs scale was presented twice to create two separate variables: own inclusion beliefs and perceived inclusion norm (i.e., norm estimation). To assess own

inclusion beliefs, participants were asked: "To what extent do you think trans women should be included in the group 'women'?" For the perceived inclusion norm, participants were asked: "To what extent do you think that the average cis woman believes that trans women should be included in the group 'women'?"

To assess the convergent validity of the own inclusion beliefs measure, participants were asked two questions on a scale from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*): "Trans women are women" and "Trans women do not belong in the group women" (reverse-scored). Collectively, the two items demonstrated a strong positive correlation ($\rho=.82$, $p<.001$) with the own inclusion beliefs scale.

Results and Discussion

Data were imported into SPSS (version 28) for analysis. Participants were generally supportive of including trans women into the group 'women,' as observed by responses that were significantly higher than the midpoint on the own inclusion beliefs scale, $M=3.79$, $SD=1.28$, $t(298)=10.62$, $p<.001$, 95% CI [0.64, 0.93], $d=0.61$. Further, the most inclusive option (E in Fig. 1) was the most frequently selected among participants (38%). However, when we compared participants' own inclusion beliefs with their perceptions of the inclusion norm, we found that the former was significantly higher than the latter ($M=2.84$, $SD=0.95$), providing evidence for a pluralistic ignorance effect, $t(298)=12.80$, $p<.001$, 95% CI [0.80, 1.09], $d=0.74$ (see Fig. 2).

Study 1 showed that while the Australian cis women in our sample held high inclusion beliefs regarding the inclusion of trans women into the group 'women' (RQ1), they underestimated the degree to which other cis women were supportive of this inclusion (RQ2). This gap between own inclusion beliefs and the perceived inclusion norm is indicative of pluralistic ignorance. This finding is also consistent with previous research that shows a contingent of Australian, North American, and British cis women who support trans women's inclusion in single-sex spaces such as sports, universities, domestic and sexual abuse services, and HIV treatment services (Auerbach et al., 2020; Outten & Lawrence, 2020; Pain et al., 2021; Pringle & Denison, 2022).

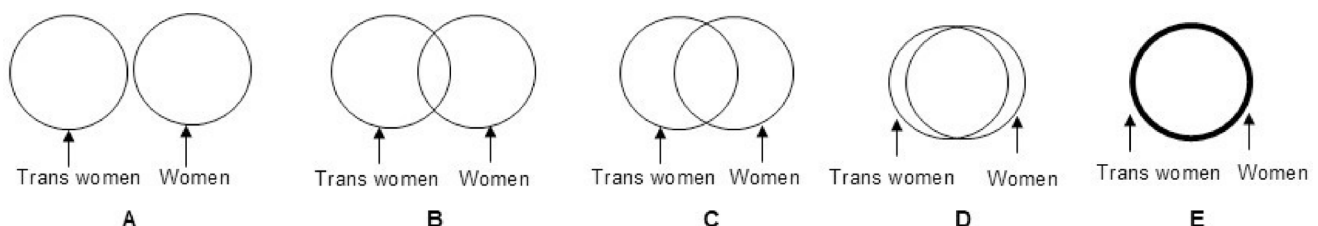


Fig. 1 Inclusion beliefs scale adapted from Aron et al. (1992) and Schubert and Otten (2002) to assess own beliefs and norm beliefs about the inclusion of trans women into the group women.

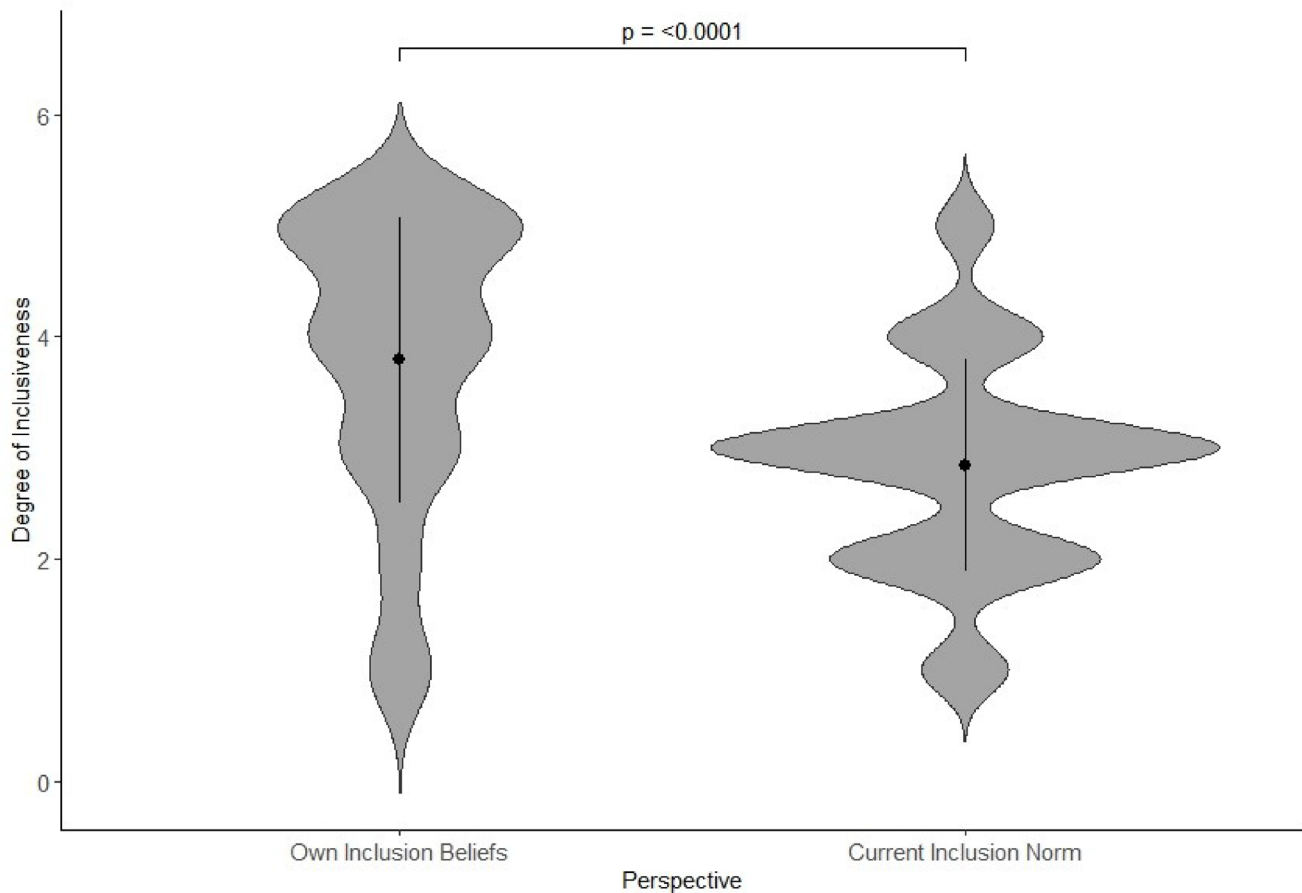


Fig. 2 Dispersion of own and perceived norm inclusion beliefs by participant perspective, with mean and standard deviation. Both variables were measured on a 5-point scale, with higher scores representing higher inclusiveness. $N=299$.

Study 1 added to the sparse literature around Australian cis women's perspectives on trans inclusion. This study also introduced and validated a pictorial scale to capture beliefs about boundary permeability (i.e., inclusion). This scale enabled us to examine the boundary maintenance beliefs of receiving ingroups during the contested process of inter-group mobility, something that has been under-investigated in the SIT literature to-date.

Study 2

The second study aimed to replicate the findings of Study 1 in a different context while exploring the extent to which pluralistic ignorance is associated with willingness to express one's beliefs about trans inclusion. Subsequently, we hypothesised that participants would demonstrate a pluralistic ignorance effect when we compared their own inclusion beliefs to the perceived inclusion norm (H1). Further, according to the spiral of silence theory (Noelle-Neumann, 1993), those who believe themselves to hold minority positions in society (regardless of the accuracy of this belief) are less willing to express their opinions than those who believe

they have majority public support. Therefore, we predicted that willingness to express inclusion beliefs would be highest (vs. lowest) when participants perceived more congruence (vs. incongruence) between their own beliefs and the perceived inclusion norm (H2). Studies 2–5 investigated pluralistic ignorance within a British context.

Method

Participants and Procedure

Of the 300 cis women in the United Kingdom recruited via Prolific, 20 participants were rejected for failing the attention check and one participant did not identify as a cis woman. The resulting 279 participants ($M_{age} = 32.50$ years; $SD_{age} = 9.86$ years) were predominantly White ($n=263$, 94.3%) and heterosexual ($n=241$, 86%). A sensitivity analysis with alpha set at 0.05, power set at 80%, and a sample size of 279 found that we could detect an effect size f^2 of 0.03 for linear multiple regression, R^2 increase. The survey procedure followed Study 1.

Measures

Own inclusion beliefs and the perceived inclusion norm were measured using the same items used in Study 1.

Willingness to Express Opinion

Four items were adapted from Geiger and Swim (2016) and Sokoloski et al. (2018) to assess how likely participants were to express their inclusion beliefs across four modalities (i.e., displaying badges, attending protests, posting on social media, and discussing opinions with those who held divergent opinions; 1 = *Extremely Unlikely* to 7 = *Extremely Likely*) ($\alpha = .85$, $M = 3.95$, $SD = 1.58$).

Analytic Strategy

To examine whether congruence between own inclusion beliefs and the perceived inclusion norm affected willingness to express opinions about trans women's inclusion, we conducted polynomial regression and response surface analysis (RSA; Edwards, 2001, 2002; Shanock et al., 2010) using the *RSA* R package (version 0.10.6; Schönbrodt & Humberg, 2023). While difference scores are often used in the literature to assess the relationship between norm misperceptions and outcome variables, there is increasing awareness of the limitations of this analytic technique (see Edwards, 2001, for a review). Specifically, concerns around the use of difference scores include their reductionist approach to the data by combining the two independent variables into a singular score, the inability to manage non-linear relationships, and the false equivalence applied to difference score values (Edwards, 2001; Miyajima & Yamaguchi, 2017; Munsch et al., 2018). Polynomial regression with RSA has been proposed as an alternative method that overcomes these limitations (Barranti et al., 2017; Edwards, 2002; Shanock et al., 2010). In their scoping review of pluralistic ignorance literature, Sargent and Newman (2021) recommended polynomial regression with response surface analysis for use with social norm misperceptions. Indeed, the social norms literature is moving towards these more holistic methods of dealing with discrepancy data as opposed to difference scores (Lees et al., 2020; Munsch et al., 2018).

The RSA involved creating four coefficients ($a1$, $a2$, $a3$, and $a4$) to plot the effects of congruence and incongruence between two variables on the outcome. There are two key lines of interest on the graph: the line of congruence (LOC) which shows the effect when the perceived norm aligns with personal beliefs, created using the $a1$ (slope) and $a2$ (curvature) coefficients, and the line of incongruence (LOIC), which shows the effect when the perceived norm opposes personal beliefs, created using the $a3$ (slope) and

$a4$ (curvature) coefficients. The hypothesis predicts a flat LOC line (non-significant $a1$ and $a2$ coefficients), indicating higher intention to express opinions when personal beliefs and perceived norms are aligned, and an inverted U-shaped LOIC line (non-significant $a3$ and a significant negative $a4$ coefficient) indicating lower intention to express opinions as the divergence between personal beliefs and perceived norms increases.

Results and Discussion

Replicating Study 1, participants' own inclusion beliefs were significantly higher than the mid-point on the own inclusion beliefs scale ($M = 3.81$, $SD = 1.28$), indicating a degree of support towards the inclusion of trans women, $t(278) = 10.60$, $p < .001$, 95% CI [0.66, 0.96], $d = 0.64$. Additionally, participants most frequently chose the option that was the most inclusive (38%). Similarly, own inclusion beliefs were significantly higher than the perceived inclusion norm ($M = 2.92$, $SD = 0.88$), indicating that this sample of cis women believed themselves to hold more positive attitudes towards the inclusion of trans women than their peers (i.e., an effect of pluralistic ignorance), $t(278) = 11.81$, $p < .001$, 95% CI [0.74, 1.04], $d = 0.71$.

Willingness to Express Opinion

The results of the polynomial regression and RSA are presented in Table 1 and visually in Fig. 3. Although the overall model was significant ($R^2 = .35$, $p < .001$), only own inclusion beliefs significantly contributed to the model (Table 1). Further, the response surface values (represented by the a coefficients) were not consistent with predictions; the significant $a1$ and $a3$ coefficients indicate that participants with high inclusion beliefs had higher intentions to express their inclusion opinions than participants with lower inclusion beliefs, regardless of the perceived congruence or incongruence with the norm.

Study 2 replicated and extended the results of Study 1, demonstrating a robustness of the norm misperception effect (H1). On average, cis women in our sample from the United Kingdom (consistent with our sample of cis women in Australia) incorrectly believed themselves to hold minority opinions about the inclusion of trans women, consistent with a pluralistic ignorance effect. However, considerations of the perceived norm did not appear to factor into cis women's willingness to express their inclusion opinions, which is inconsistent with the predictions of H2. In fact, the main effect of own inclusion beliefs indicated that cis women with higher inclusion beliefs tended to be more willing to

Table 1 Study 2 Results of Polynomial Regression Analysis

Variable	<i>b</i>	<i>SE</i>	LLCI	ULCI	<i>p</i>
Own Beliefs	0.63	0.06	2.87	3.39	<.001
Perceived Norm	0.06	0.11	-0.16	0.28	.567
Own ²	0.16	0.05	0.06	0.25	<.001
Perceived Norm ²	-0.05	0.07	-0.20	-0.04	.471
Own*Norm	0.01	0.08	-0.14	0.16	.885
<i>a</i> ₁	0.69	0.01	0.50	0.88	<.001
<i>a</i> ₂	0.11	0.08	-0.04	0.26	.133
<i>a</i> ₃	0.56	0.15	0.26	0.86	<.001
<i>a</i> ₄	0.09	0.14	-0.19	0.37	.520

Note. *n* = 279. Unstandardised coefficients are presented. *SE* = Standard Error. LLCI = Lower Limit Confidence Interval. ULCI = Upper Limit Confidence Interval. *a*₁ = LOC slope; *a*₂ = LOC curvature; *a*₃ = LOIC slope; *a*₄ = LOIC curvature

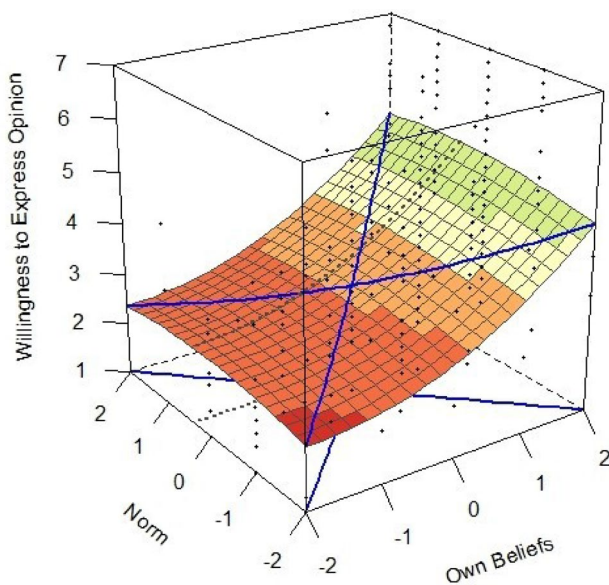


Fig. 3 Study 2 results of response surface analysis. The colour gradient indicates scores across the outcome variable. Own beliefs and perceived norm variables are scale-centred.

express those beliefs than cis women with lower inclusion beliefs.

Study 3

Having shown that pluralistic ignorance exists for this issue, Study 3 aimed to reduce pluralistic ignorance around the inclusion of trans women using a group summary intervention, and then to measure the effect of this intervention on willingness to express inclusion opinions. Despite failing to see an effect of misperceived norms on the measured outcomes in Study 2, corrected norms may nonetheless influence behaviour, particularly among group members who previously felt misaligned with the perceived norms (Noelle-Neumann, 1993; Tankard & Paluck, 2016). Further, while preliminary exploratory evidence in Study 2 indicated that own beliefs had a stronger effect on behavioural

intention than perceived norms, previous research has argued that personal beliefs are more resistant to change than norms (Flezzani & Benshoff, 2003; Tankard & Paluck, 2016). Therefore, the practical significance of investigating the norm misperception and any possible impacts was deemed more important than pursuing the results pertaining to own beliefs. As such, we hypothesised that participants in the correction condition would report a smaller difference between their own beliefs and the perceived norm (i.e., a reduced pluralistic ignorance effect) relative to participants in the control condition (H1).

Further, we hypothesised that participants with high inclusion beliefs in the intervention condition (relative to the control) would increase their willingness to express their inclusion opinions both offline (H2a) and online (H2b) and their actual speaking out behaviour (H3), with opposite effects observed for participants with low inclusion beliefs. We believed it prudent to examine opinion expression online due to the growing network of exclusionary subgroups in the United Kingdom (i.e., TERFs) on social media such as Twitter (e.g., Gardner, 2023; Lu & Jurgens, 2022). The visibility of these groups may hinder willingness to express inclusion opinions among cis women online relative to offline, however online behaviours may also present fewer barriers (i.e., time, effort, perceived anonymity) to collective action involvement than offline behaviours (Schumann & Klein, 2015). As such, we do not make predictions about any differences in opinion expression between offline and online settings.

Method

Participants and Procedure

A total of 250 UK-based cis women who had not been involved in Study 2 were recruited via Prolific. Due to the prevalence of anti-trans rhetoric on Twitter in the United Kingdom (Lu & Jurgens, 2022), we selected this social

media platform to test willingness to express opinion online; we therefore only recruited participants who had a Twitter account. Of the 250 participants who completed the study, two were excluded for not identifying as cis women and five were excluded for failing the attention check. A post-hoc decision was made to focus on the moderation test (pre-registered as exploratory) rather than the pre-registered independent *t*-test to better align with the theoretical framework. Consequently, a sensitivity analysis was conducted. The sensitivity analysis identified that we could detect an effect size f^2 of .033 with alpha set at .05, power set at 80%, and a sample size of 243 participants for linear multiple regression, R^2 increase. After exclusions, 119 participants were randomly assigned to the norm correction and 124 participants were randomly assigned to the control. Participants ($M_{age} = 38.02$ years; $SD_{age} = 12.50$ years) were predominantly White ($n = 198$, 81.5%) and heterosexual ($n = 191$, 78.6%).

Measures

Own inclusion beliefs and the perceived inclusion norm was measured using the same items from the previous studies. The Willingness to Express Opinion measure from Study 2 was adapted to measure more offline-affiliated behavioural intentions by replacing the item “I would post on social media about my opinion” with “I would sign a petition that supported my opinion” ($\alpha = .82$).

The Majority of Cis Women in the UK are Positive Towards Inclusion

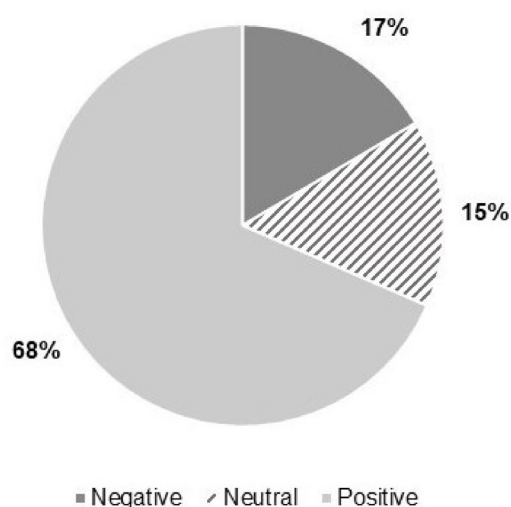


Fig. 4 Information presented to participants in the Study 3 correction condition.

Willingness To Express Opinion Online

Three items assessed participants’ desire to respond to an anti-inclusion tweet that advocated for single-sex spaces that accommodated cis women only. The anti-inclusion tweet was adapted from an existing tweet to increase realism and created using Tweetgen (<https://www.tweetgen.com>; see p. 4 of the Supplemental Material in the OSF). Participants were instructed to imagine that they had come across this tweet while online and then indicated their likelihood of responding to the tweet through liking, retweeting, and commenting (1 = *Extremely Unlikely* to 7 = *Extremely Likely*) ($\alpha = .86$). A sentiment analysis and inductive content analysis of these responses can be found in the Supplemental Material in the OSF (pp. 7–9).

Procedure

Participants were randomly assigned to one of two conditions: a norm correction condition or a control (no information) condition. The norm correction condition was informed by similar interventions (Hornsey et al., 2004; White et al., 2002). To create the correction material, a pie chart was developed based on the results from Study 2 informing participants that most cis women in the United Kingdom held positive inclusion attitudes toward trans women (see Fig. 4. Figure appeared in colour for participants).

To increase the strength of the information, participants were asked to summarise the findings. Both conditions then completed own and perceived others’ inclusion beliefs, followed by the two willingness to express opinion measures. In addition to indicating intention to respond to the tweet, participants were given the opportunity to reply by entering up to 280-characters (the Twitter posting limit at the time of this study) in a text box. Participants were debriefed at the end of the study and informed that they had read real data collected by the research team in Prolific for a previous study.

Results

Descriptive statistics for all variables are presented in Table 2.

Manipulation Check

Results showed that the perceived inclusion norm was higher in the correction condition than in the control condition, providing evidence that the manipulation worked as intended, $t(235.15) = -5.18$, $p < .001$, 95% CI [-0.90, -0.40],

Table 2 Study 3 Means and Standard Deviations for Key Variables by Condition

	Control Condition		Correction Condition	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Own Inclusion Beliefs	3.35 ^a	1.43	3.55 ^a	1.30
Perceived Inclusion Norm	2.64 ^a	1.07	3.29 ^b	0.88
Willingness to Express Opinion Online	2.44 ^a	1.65	2.78 ^a	1.70
Willingness to Express Opinion Offline	3.85 ^a	1.57	4.20 ^a	1.41

Note. Different superscript letters indicate significant differences across conditions at $p < .05$.

Own inclusion beliefs and perceived inclusion norm were assessed on a 5-point scale where higher scores indicated higher inclusion. The Willingness to Express Opinion measures were assessed on a 7-point scale where higher scores indicated higher willingness

$d = -0.66$, equal variances not assumed, Levene's test: $F(1, 241) = 5.25$, $p = .023$.

Pluralistic Ignorance

Participants systematically under-estimated the level of perceived support for the inclusion of trans women into the group 'women' in both the control condition, $t(123) = 5.86$, $p < .001$, 95% CI [.47, .95], $d = 0.53$, and the correction condition, $t(118) = 2.72$, $p = .008$, 95% CI [.07, .47], $d = 0.25$, relative to their own inclusion beliefs. Using the *TOSTER* package in R (version 0.7.1; Caldwell, 2022; Lakens, 2017), it was determined that the two effect sizes were not equivalent, $z = -2.10$, $p = .036$. These results suggest that the intervention was successful in reducing the gap between own inclusion beliefs and the perceived inclusion norm relative to the control, consistent with H1.

Willingness to Express Opinion

To explore whether own inclusion beliefs moderated the relationship between condition and willingness to express opinions online and offline, a series of moderation analyses were conducted using Hayes' (2022) PROCESS macro Model 1 (version 3.5.3). Regarding willingness to express opinions online (H2a), no main effects were observed for own inclusion beliefs ($p = .791$) or condition ($p = .118$) and no interaction was present ($p = .554$). The main effect of condition ($p = .144$) and the interaction ($p = .203$) were also non-significant for willingness to express opinions offline (H2b). However, a main effect emerged for own inclusion beliefs: $b = 0.59$, $SE = 0.13$, $t(239) = 7.13$, $p < .001$, 95% CI [0.43, 0.76].

Speaking Out Behaviour

Most participants ($n = 239$, 98.4%) opted to leave a comment to the anti-inclusion tweet. A moderated logistic regression was conducted to explore whether own inclusion beliefs moderated the relationship between condition and speaking out behaviour (coded as 0 = not tweeted and 1 = tweeted). Due to low observations in the no-tweet category, Firth's

penalized likelihood logistic regression was employed using the *logistf* R package (version 1.26.0; Heinze et al., 2023). Contrary to H3, statistical significance was not observed for either the main effects of condition ($p = .520$) or own beliefs ($p = .398$), nor for their interaction ($p = .655$).

Discussion

Study 3 showed that this sample of British cis women exhibited pluralistic ignorance whereby they underestimated the degree to which they perceived that other ingroup members supported the inclusion of trans women. The norm correction intervention was successful in reducing this pluralistic ignorance effect relative to the control condition. However, this reduction did not correspond to a change in behaviour or intention to express inclusion opinions online or offline. Partially replicating Study 2, a main effect emerged for own inclusion beliefs on willingness to express opinions offline, but not online.

While these results suggest that own beliefs have a relatively stronger influence on behavioural intention than perceived norms, the study's design may have attenuated the effects of the norm. Specifically, the manipulation material used in this study showed that 68% of cis women held high inclusion beliefs, which may have been perceived as too low to constitute a clear majority. To strengthen the norm correction manipulation, these proportions were adjusted in Study 4 to show a more inclusive norm held by a large majority of cis women.

Study 4

Method

Participants and Procedure

Prolific was used to recruit 250 cis women from the United Kingdom who used Twitter and had not participated in the previous studies. After three participants were removed for

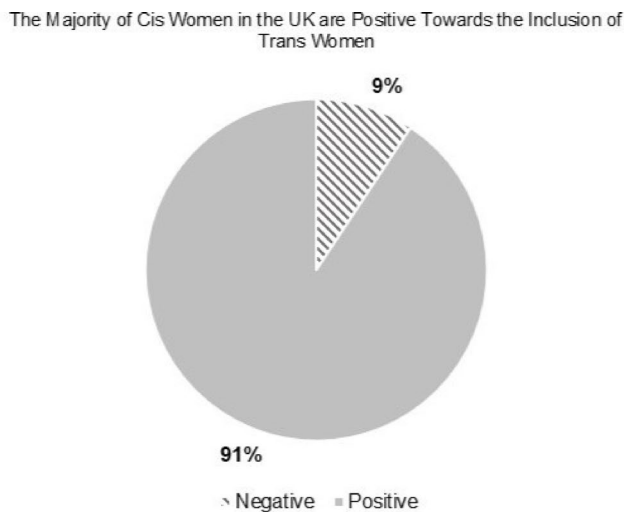


Fig. 5 Information presented to participants in Study 4 correction condition.

failing the attention check (all in the correction condition), 127 participants were randomly assigned to the control condition and 120 participants were randomly assigned to the correction condition. Participants ($M_{age} = 40.36$ years, $SD_{age} = 12.71$ years) were mostly heterosexual ($n = 210$, 85%), and White ($n = 227$, 91.9%).

Independent t -tests and chi-square tests showed that participants in both conditions were equally balanced across demographic variables. The procedure and hypothesis testing were the same as Study 3.

Measures

Own inclusion beliefs and the perceived inclusion norm were assessed in Study 4 using the same measures as Studies 1–3.

Willingness to Express Opinion Offline

Four items measured willingness to share inclusion opinions by: attending a protest, displaying signs or badges, discussing opinions in community meetings, and meeting with local government representatives to discuss inclusion-related

policies (1 = *Extremely Unlikely* to 7 = *Extremely Likely*) ($\alpha = .86$).

Willingness to Express Opinion Online

Four items measured online inclusion opinion expression, including: publicly signing an online petition, posting materials to social media, defending others online who share their inclusion opinion from criticism, and commenting on online news articles (1 = *Extremely Unlikely* to 7 = *Extremely Likely*) ($\alpha = .89$).

Speaking Out Behaviour

To measure actual speaking out behaviour, we reworded the anti-inclusion Tweet response item from Study 3 so it was clear that providing an open-ended response was optional. Analysis of these responses can be found in the Supplemental Material in the OSF (p. 15).

The correction material from Study 3 was also amended by removing the ‘neutral’ segment from the pie chart and collapsing all data that indicated at least some support for inclusion (i.e., response selections “B” through “E” on the Inclusion Beliefs Scale, see Fig. 1) into the “Positive” segment (see Fig. 5).

Results

Manipulation Check

Providing evidence that the manipulation worked as intended, the perceived inclusion norm was higher in the correction condition than the control condition, $t(245) = -8.03$, $p < .001$, 95% CI $[-1.13, -0.69]$, $d = -1.02$. In addition, participants in the correction condition reported higher own inclusion beliefs than those in the control condition, $t(244.775) = -3.14$, $p = .002$, 95% CI $[-0.84, -0.19]$, $d = -0.40$, equal variances not assumed, Levene’s test: $F(1, 245) = 5.62$, $p = .019$.

Table 3 Study 4 Means and Standard Deviations for Inclusion Beliefs by Condition

	Control Condition		Correction Condition	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Own Inclusion Beliefs	3.29 ^a	1.35	3.81 ^b	1.24
Perceived Inclusion Norm	2.61 ^a	0.88	3.52 ^b	0.90
Willingness to Express Opinion Online	3.73 ^a	1.72	3.93 ^a	1.50
Willingness to Express Opinion Offline	2.79 ^a	1.56	2.75 ^a	1.49

Note. Different superscript letters indicate significant differences across conditions at $p < .05$

Own inclusion beliefs and the perceived inclusion norm were assessed on a 5-point scale where higher scores indicated higher inclusion. The Willingness to Express Opinion measures were assessed on a 7-point scale where higher scores indicated higher willingness

Pluralistic Ignorance

Descriptive statistics for all variables are presented in Table 3. Participants in both the correction condition $t(119)=3.23$, $p=.002$, and the control condition $t(126)=6.54$, $p<.001$, misperceived the norm as being more exclusionary towards trans women than their own inclusion beliefs, indicating pluralistic ignorance. The magnitude of this difference was reduced in the correction condition, $d=0.29$, 95% CI [0.11, 0.48], as compared to the control condition, $d=0.58$, 95% CI [0.39, 0.77]. Using the *TOSTER* package, it was determined that these effect sizes were not equivalent, $z=-2.17$, $p=.030$. These results suggest that the intervention was successful in reducing the gap between own inclusion beliefs and the perceived inclusion norm.

Willingness to Express Opinion

Similarly to Study 3, there was no main effect for condition on willingness to express opinions online ($p=.707$) or offline ($p=.131$). Additionally, the interactions were not significant for online ($p=.958$) or offline expression ($p=.303$). However, in contrast to the previous study, main effects were observed for own inclusion beliefs in both expression models, online: $b=.52$, $SE=.10$, $t(243)=5.32$, $p<.001$, 95% CI [.33, .71]; offline: $b=.39$, $SE=.09$, $t(243)=4.16$, $p<.001$, 95% CI [0.20, 0.57].

Speaking Out Behaviour

Thirty-three participants (13.4%) opted to leave a comment to the tweet. A moderated logistic regression replicated Study 3, finding no main effects (condition: $p=.290$; own beliefs: $p=.118$) nor interaction ($p=.864$) between condition on own beliefs on speaking out behaviour.

Discussion

Consistent with Study 3, Study 4 showed that our sample of British cis women experienced pluralistic ignorance regarding trans women's inclusion norms. As expected, the difference between own beliefs and the perceived inclusion norm was lower in the correction condition relative to the control condition, suggesting the intervention was successful in reducing this pluralistic ignorance. Participants in the intervention condition also appeared to express more inclusive own beliefs than those in the control condition based on the independent t -test results. The random allocation of participants to condition suggests that these findings may be attributable to social influence and conformity effects (Turner et al., 1987). Namely, participants may have reported more

inclusionary beliefs to remain aligned with the communicated group norm.

However, our hypotheses that corrected norm perceptions would affect speaking out intentions and behaviour relative to one's own beliefs were not supported. The replicated main effect of own inclusion beliefs for both intention expression types (as compared to Study 3 when only offline expression demonstrated a significant main effect of own beliefs) may reflect the changes made to the online measure to make the two expression modalities comparable. In Study 5, we broaden our investigation to examine the impacts of correcting pluralistic ignorance beyond opinion expression. Specifically, we examine the impacts of pluralistic ignorance and its correction for collective action intentions such as attending protests and signing petitions, as well as supporting policy change that protects trans women's rights. We also expect that misperceiving social support for one's opinion may hamper beliefs that attempts at collective change towards inclusion will be successful (i.e., collective efficacy; van Zomeren et al., 2008).

Study 5

We used a pre-post within-participants experimental design for Study 5 to assess changes due to the intervention over time, with each participant serving as their own baseline. Previous norm correction research has utilised similar repeated measures designs without a control group (LaBrie et al., 2010; Mattern & Neighbors, 2004; Perkins et al., 2011). We hypothesised that pluralistic ignorance would reduce from Time 1 to Time 2 as a function of the norm correction intervention (H1). Additionally, we predicted that policy support, collective action intentions, and collective efficacy would vary over time as a function of own inclusion beliefs, consistent with spiral of silence theorising (H2). Additional outcome measures were pre-registered, however due to space we mention only those variables related to collective action outcomes. See Supplemental Material in the OSF for complete measures and analyses (pp. 19–26).

Method

Participants

One hundred cisgender women in the United Kingdom were recruited of which three failed embedded attention checks, two elected to withdraw their data upon debriefing, two exceeded Prolific time restrictions, and two returned their submissions part-way through the Time 2 survey. Reminder emails were sent to participants when the second survey was made available and all participants from Time 1 returned to

Table 4 Study 5 Means and Standard Deviations for Key Variables by Time

	Pre-Intervention (Time 1)		Post-Intervention (Time 2)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Own Inclusion Beliefs	5.44 ^a	3.76	6.29 ^b	3.60
Perceived Inclusion Norm	3.89 ^a	2.28	7.36 ^b	2.41
Policy Support	2.53 ^a	1.72	2.60 ^a	1.86
Collective Efficacy	4.73 ^a	0.99	4.92 ^a	1.17
Collective Action	3.71 ^a	1.32	3.67 ^a	1.37

Note. Different superscript letters indicate significant differences across time at $p < .05$

Own inclusion beliefs and the perceived inclusion norm were assessed on an 11-point scale where higher scores indicated higher desired inclusion. Policy support, collective efficacy, and collective action intentions were measured on 7-point scales where higher scores indicate higher support, perceived efficacy, and action intentions

complete Time 2. A post-hoc decision was made to focus on the moderation test (pre-registered as exploratory) rather than the pre-registered paired t -test to better align with the theoretical framework. Consequently, a simulation power analysis was conducted using the *simr* package (Green & MacLeod, 2016) in R to detect the interaction between time (pre- and post-intervention) and own inclusion beliefs in the three models. Based on 1000 simulations, the estimated power was between 81.80% (policy support, 95% CI [79.27%, 84.15%]), 46.20% (collective efficacy, 95% CI [43.08, 49.35%]), and 70% (collective action, 95% CI [67.05%, 72.83%]), indicating varying power to detect the interaction effect at an alpha of .05. Eligible data were obtained from 91 participants ($M_{age} = 41.98$ years, $SD_{age} = 14.69$ years) who were primarily heterosexual ($n = 79$, 86.8%) and White ($n = 74$, 81.3%).

Measures

The corrective information was the same as Study 4. The 5-point pictorial scale to measure inclusion beliefs was replaced with a 11-point pictorial slider scale to capture more variability.

Policy Support

One item measured support for the removal of legal protections of trans peoples in the UK Equality Act 2010, a topical issue at the time of this study (Ali, 2022; 1 = *Strongly Oppose* to 7 = *Strongly Support*). Higher endorsement of this policy is reflective of higher exclusionary behaviour intentions.

Collective Action Intentions

Five items adapted from the willingness to express opinion measures examined participants' likelihood of engaging in collective action behaviours to change the boundaries of womanhood as consistent with their own inclusion beliefs (e.g., "Attend a protest or a march that was consistent with

my opinion"; 1 = *Extremely Unlikely* to 7 = *Extremely Likely*) (Time 1 $\alpha = .80$, Time 2 $\alpha = .83$).

Collective Efficacy

Four items adapted from van Zomeren et al. (2010) measured how much participants believed that their group (defined as cis women who held the same inclusion beliefs as them) would be successful in achieving their goals of changing the boundaries of womanhood (e.g., "If we work together, we can change public opinion on inclusion"; 1 = *Extremely Unlikely* to 7 = *Extremely Likely*) (Time 1 $\alpha = .78$, Time 2 $\alpha = .90$).

Procedure

This study adopted a one-day delay between pre-post assessments to reduce attrition and minimise effects of contextual changes outside experimental control. Due to this short delay, we did not anticipate changes to our variables that were not attributable to our intervention. At Time 1, participants completed pre-test assessments of their own inclusion beliefs, the perceived inclusion norm, and outcome measures. After a one-day delay, participants were invited to complete the Time 2 survey. At Time 2, participants read the corrective information and completed the same measures from Time 1.

Results

Manipulation Check

Paired t -tests indicated that the perceived inclusion norm was significantly higher at Time 2 than Time 1 providing evidence that the correction worked as intended, $t(90) = -13.93$, $p < .001$, $d = -1.46$. Contrary to expectations, own inclusion beliefs also increased across time, $t(90) = -3.67$, $p = .001$, $d = -0.39$.

Pluralistic Ignorance

Descriptive statistics for all variables are shown in Table 4. Evidence for the pluralistic ignorance effect emerged at Time 1, whereby the perceived inclusion norm was lower than one's own inclusion beliefs, $t(90)=4.69$, $p<.001$, $d=0.49$. This effect reversed at Time 2 with the norm being perceived as more inclusive than oneself, $t(90) = -3.22$, $p=.002$, $d = -0.34$.

Effect of Norm Correction

Due to changes in our intended moderator (own inclusion beliefs) from Time 1 to Time 2, we were unable to conduct the planned moderation analyses. Instead, we employed a mixed effects model using the *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages in R. Participants were entered as a random effect to account for individual differences in the within-participant design, and time (pre- and post-intervention) and own inclusion beliefs were included as fixed effects.

For collective action, a significant main effect of own inclusion beliefs was observed, indicating that those with higher own inclusion beliefs were less likely to support the anti-inclusionary policy: $b=.07$, $SE=.03$, $t(157.87)=2.72$, $p=.007$. For collective efficacy, an interaction between time and own inclusion beliefs emerged, $b=.06$, $SE=.03$, $t(90.70)=2.04$, $p=.044$. Simple slopes testing revealed that the effect of time on collective efficacy was only significant for cis women with high level of own inclusion beliefs, $b=.34$, $SE=.15$, $t(153)=2.35$, $p=.021$. No other main effects or interactions were statistically significant across policy support, collective action, and collective efficacy (see pp. 19–26 of the Supplemental Material in the OSF for results).

Discussion

Consistent with Studies 1–4, pluralistic ignorance was observed at Time 1 regarding the inclusion of trans women into the group ‘women.’ We also found that the norm correction intervention was successful in increasing perceptions of the inclusion norm as well as one's own inclusion beliefs. However, unlike Studies 3 and 4, the norm misperception at baseline reversed after the intervention, such that the norm was perceived as being *more* inclusionary than one's own beliefs. It is possible that participants re-evaluated their own beliefs in relation to the new information about the norm through processes of social comparison, although it is unclear why this finding was not observed in the previous two norm correction studies.

Despite these changes in pluralistic ignorance and own beliefs following the norm intervention, there were no differences observed across time for policy support or collective actions intentions. However, the interaction effect observed for the relationship between time and collective efficacy, moderated by participants' own inclusion beliefs, was consistent with predictions. The results imply that correcting pluralistic ignorance could lead to shifts in collective efficacy, particularly among those with highly inclusive personal beliefs.

General Discussion

In the current research, we investigated norm perceptions around cis women's inclusion of trans women into the group ‘women.’ We also tested the effect of inclusion norm perceptions on cis women's motivations to express their opinions about trans women's inclusion. Across five studies, we consistently observed evidence of a norm misperception around inclusion, such that most cis women from our samples in Australia (Study 1) and the United Kingdom (Studies 2–5) supported at least some degree of trans women's inclusion within the boundaries of ‘womanhood’ but underestimated the degree of this support among their group members. That is, our participants believed that they were more inclusionary than other cis women, resulting in a robust pluralistic ignorance effect. However, a brief norm correction intervention (Studies 3–5) was found to increase participants' perceptions of the inclusionary norm and, in turn, reducing this pluralistic ignorance. Further, cis women with high inclusion beliefs showed greater perceived collective efficacy to redefine the boundaries of womanhood boundaries post-intervention compared to pre-intervention. In contrast, cis women's willingness to express their views on trans women's inclusion remained largely unaffected by the intervention. In the remainder of the discussion, we expand on these findings and consider their implications for both theory and practice.

Beliefs and Perceptions of Transgender Inclusion

As aforementioned, we showed consistent support for at least some inclusion of trans women among our samples. This finding emerged across two different contexts which have both recorded increasing transphobic rhetoric, particularly towards trans women, in recent years that suggested an exclusionary norm among cis women (García & Badge, 2021; Pearce et al., 2020). The current research suggests that this exclusionary norm is driven by an unrepresentative and vocal minority of cis women. Our research is also consistent with literature that finds support among cis women

towards the inclusion of trans women in single-sex spaces (e.g., Auerbach et al., 2020; Outten & Lawrence, 2020; Pain et al., 2021; Pringle & Denison, 2022).

Despite the overall trend of cis women supporting the inclusion of trans women among our samples, our participants consistently underestimated the inclusion norm thus demonstrating a norm misperception in the form of pluralistic ignorance. Pluralistic ignorance has been well documented across a range of situations (Sargent & Newman, 2021), and our research adds a new issue to this literature. By identifying a pluralistic ignorance effect regarding cis women's inclusion of trans women within the boundaries of womanhood, we contribute to understandings of group processes and social perception, offering a foundation for future work on meaningful behaviour, belief, and attitude change. This finding also raises further questions around how norm perceptions are formed and the (potentially inaccurate) cues that group members are using to determine the norms.

An alternative interpretation of our findings is that the pluralistic ignorance effect we observed could also be interpreted through the lens of the 'better-than-average' effect. Here, group members overestimate their own rate of a desirable behaviour relative to their group members to bolster their own feelings of superiority (Alicke & Govorun, 2005; Zell et al., 2020). To the degree that including trans women into the group 'women' is seen as the desirable or morally superior course of action, cis women may wilfully underestimate the degree to which other group members engage in this behaviour. However, we observed in Study 5 that participants appeared to re-evaluate their own inclusion beliefs relative to the norm after receiving the intervention, resulting in higher norm perceptions than own inclusion beliefs and suggesting that the better-than-average effect may not have been implicated.

Results from exploratory qualitative analyses in Studies 3–5 also lend support for the pluralistic ignorance effect as opposed to the better-than-average effect. Specifically, after viewing the norm information, participants were asked to reflect on their thoughts and feelings about this information. Coding of emotional responses (see pp. 4–5 for Study 3, p. 11 for Study 4, and p. 17 for Study 5 in the Supplemental Material in the OSF for details) indicated that participants were pleasantly surprised to find themselves sharing opinions with most of their group. In the case of self-enhancement, receiving feedback that one's behaviours, beliefs, or attitudes align with the broader group (i.e., are average) is often perceived as undesirable and threatening to one's self-image, thereby triggering defensive responding (Gaertner et al., 2012; Howell & Ratliff, 2017). Consequently, the relief and validation expressed by participants when learning that their beliefs were reflective of the group is more consistent

with a pluralistic ignorance explanation than the better-than-average effect.

The Effectiveness of the Norm Correction Intervention

At the same time, sharing the inclusive group norm via group summary interventions also increased participants' perceptions of the inclusionary norm. Our research joins other group summary intervention studies to demonstrate the effectiveness of communicating norm information to change norm perceptions (e.g., De Souza & Schmader, 2021; Perkins et al., 2011; Yamin et al., 2019). This simple approach to changing perceptions of the norm may have potential for challenging perceptions of the exclusionary norm on a larger scale, outside of lab settings, as discussed further below.

The intervention was also successful for increasing perceptions of collective efficacy among cis women with high inclusion beliefs in Study 5. This finding suggests opportunities for enhancing solidarity and fostering action toward inclusionary goals. The belief that one's group can achieve their goals through their joint efforts (i.e., collective efficacy) is considered an integral predictor of collective action (van Zomeren et al., 2008). The current research suggests that highly inclusive cis women who learn that they have ingroup support for their positions on inclusion are more likely to consider collective action that expands the boundaries of womanhood as an achievable goal. This collective efficacy, in turn, may have implications for future inclusive action.

We also observed increases in participants' own inclusion beliefs after viewing the correction material in Studies 4 and 5, relative to the control and pre-correction conditions, respectively. This finding is inconsistent with the rationale underpinning the study that perceived norms are more susceptible to influence than personal beliefs (Flezzani & Benshoff, 2003; Tankard & Paluck, 2016). However, previous literature has similarly found changes in personal beliefs upon receiving information about the norm (De Souza & Schmader, 2021; Duncan et al., 2020; Stangor et al., 2001; Tankard & Paluck, 2017). This re-evaluation of one's own beliefs when compared against the group may reflect the influential power of group norms (Turner et al., 1987). Alternatively, participants may have evaluated their exclusionary views as socially undesirable when presented with a norm of inclusion, resulting in a hesitancy to report their true attitudes post-intervention. Future research could explore the longevity of these observed changes in own inclusion beliefs through longitudinal studies that incorporate longer delays and additional time points while presenting normative information. The lack of significant change

in own beliefs in Study 3 could reflect the stronger manipulation material used in Studies 4 and 5 which indicated a greater majority position of inclusion.

Finally, neither our norm correction intervention nor pluralistic ignorance itself affected collective action intentions, willingness to express opinion, or policy support for the inclusion of trans women within the group ‘women,’ contrary to predictions. One possible reason for this outcome is that participants’ own inclusion beliefs were associated with their willingness to express their inclusion beliefs more than their perceptions of the norm. Past research has also found that own beliefs, and not perceptions of the norm, are predictive of allyship behavioural intentions such as men’s intentions for gender equality (De Souza & Schmader, 2021) and allies’ collective action intentions for same-sex rights (Hornsey et al., 2004). These findings, along with our own, run counter to the predictions of social influence theories and the spiral of silence, that is, that conformity to the norm will supersede one’s own personal beliefs and desires (Noelle-Neumann, 1993; Turner et al., 1987). However, and as aforementioned, longitudinal assessments of norm interventions that incorporate repeated exposures to normative information may help to slowly change behaviours.

Practice Implications

The robust pluralistic ignorance effect may have important implications for policymakers who make decisions based on perceptions about public opinion on an issue. Specifically, recognising that sections of the public privately endorse a cause that is not reflected by the perceived norm may be useful for politicians or leaders to engage with these ‘silent majority’ norms more effectively. Educating policymakers about norm misperceptions while also supporting more accurate data collections methods on public opinion may help to develop policy decisions that better represent their constituents.

The group summary intervention presented in the current study may also be scaled up to affect wider change to norms perceptions. For example, the norm information may be diffused to a broader audience by displaying posters in public spaces (e.g., Perkins et al., 2011; Thomas et al., 2017). Examples of impactful spaces where inclusion norms are likely to be salient within the context of trans women’s inclusion include single-sex spaces (i.e., bathrooms, changing rooms). Feminist and LGBTQIA+ organisations could also be engaged to disperse corrective norms on inclusion to amplify support for trans rights, encourage visible allyship from the ‘silent majority’ of cis women, and challenge the narrative that exclusionary perspectives are dominant in Australia and the United Kingdom.

At the same time, it is important to note that the non-significant effects in the current research can be informative for designing effective norm correction interventions. For instance, the incremental refinement of our study design revealed the conditions under which norm correction interventions are and are not successful. Specifically, we showed that pluralistic ignorance was reduced regardless of the majority proportion shown to participants (i.e., 68% in Study 3 vs. 91% in Studies 4 and 5). However, own beliefs only became significantly more inclusive following the interventions in Studies 4 and 5 where the pro-inclusion proportion constituted a strong majority (91%). These findings suggest that advocacy organisations seeking to employ information interventions like the one in the current study may find greater success with the presentation of a clear and unambiguous majority norm.

Limitations and Future Directions

Future research in this area may be guided by addressing the limitations of the current work. First, due to the different socio-political contexts around the world, the current findings may not be applicable to countries outside of Australia and the United Kingdom. Namely, the recognition of trans identities and trans rights in general remains contested globally (Ipsos, 2023) which may result in different norms and perceptions of those norms. Additionally, the use of convenience samples from Australia and the United Kingdom may limit the representativeness of the current research’s findings within these countries as well. Finally, we acknowledge that there is also evidence of cis women’s exclusionary attitudes towards trans women (e.g., Morgenroth et al., 2024). Further research may help to elucidate whether the pro-inclusion norm identified across the five studies here reflects a majority opinion across different contexts, however, given the context-dependent nature of inclusion beliefs, determining a generalised set of beliefs among a representative sample may not be meaningful. Regardless, our data show evidence that, within the contexts of the current work, on average cis women held high inclusion beliefs and were willing to express their opinions on inclusion.

Future research may also seek to identify and disentangle the unique effects of perceiving the descriptive norm (how group members behave) and the injunctive norm (the behaviours that group members approve of; Cialdini et al., 1991). While we examined the descriptive norm regarding trans women’s inclusion into the group ‘women,’ participants may have been responding to an injunctive norm that they perceived as pro-inclusionary. Correspondingly, it follows that cis women with high inclusion beliefs would be emboldened to discuss their inclusion opinions, even when perceiving a low inclusion descriptive norm.

A further limitation pertains to the intervention approach implemented. Researchers have theorised that for a corrected norm to affect behavioural intentions, one must see evidence of this new norm. For example, De Souza and Schmader (2021) showed that men were more likely to confront a sexist comment when they observed that other men were concerned about gender inequality as opposed to when they were only informed (i.e., through a group summary intervention) that there was a norm of concern. In the present research, we only provided an informational intervention about the inclusion norms which was insufficient to elicit behavioural change. Future research may encourage participants to reflect on how they have seen and experienced inclusion in their own lives, demonstrate inclusion from confederates under experimental conditions (e.g., as in Study 3 of De Souza & Schmader, 2021), or read about real-world examples of inclusion (such as solidarity from athletes who are cis women; e.g., Rapinoe, 2021).

Our inability to detect behavioural changes may also be due to the chosen reference group. Social norms literature and SCT state that for norms to influence behaviour, the reference group should be relevant to participants' own self-concept (Tankard & Paluck, 2016; Turner et al., 1987). That is, the more distant a group is from one's sense of self, the less influence that group will have on one's behaviour. Further, collection action theorising argues that a strong social identification is a required precursor to mobilisation (van Zomeren et al., 2008). As such, the use of general groups such as 'students,' 'the public,' or, in our study, 'the average cis woman in the United Kingdom' may not be specific enough to influence behaviour (see also Halbesleben et al., 2007).

Similarly, the group identity made salient in these studies ('cis women in the United Kingdom/Australia') may not have been associated with mobilisation objectives. That is, participants may not have adopted what van Zomeren and colleagues (2008) refer to as a politicised identity, which is one of the pre-conditions for collective action. However, the group 'cis women' is based on gender identity and not necessarily a desire to improve collective group standing. Instead, a group based on feminist identification would be a politicised identity that might be more associated with collective action engagement and strengthened when normative information is presented and should be the subject of future research.

Although these interventions did not appear to affect cis women's intentions to push for social change as measured in this research, future research may explore whether corrected norms can influence other types of beliefs and/or behaviours. For instance, norm misperceptions may have implications beyond the population studied such as for trans women. Aversa and colleagues (2022) use the term

"anticipatory prejudice" to describe the stressors felt by the trans community as they predict whether a space will be safe for them to enter. If cis women are underestimating the inclusion norms among their peers, we may expect that trans women are as well. Future research should seek the perspectives of trans women regarding inclusion-based norms and the subsequent implications of norm perceptions on their health and wellbeing.

Future research may also extend the usage of our inclusion beliefs scale to other situations where intergroup mobility is contested. This scale (in either the static 5-point or dynamic 11-point form) was specifically designed to examine the openness of group boundaries from the perspectives of receiving groups, which has not been sufficiently examined in research to-date (Armenta et al., 2017; Charlton & Bettencourt, 2001). In doing so, it conceptualises participants as gatekeepers who have an active role in determining not only whether their group boundaries are open, but also the degree to which they are open to incoming outgroup members. These tokenistic boundaries (Wright, 2001) allow receiving groups to maintain control over their ingroup identities by setting the conditions for entry. Our scale and the current research more broadly extend SIT to understand the attitudes and behaviours of receiving groups and provides a tangible way in which to empirically capture the nuances of their relationship with incoming groups.

Conclusion

Our research presents evidence for a norm of including trans women into the group 'women' among samples of cis women in Australia and the United Kingdom. This norm contrasts with media narratives and vocal minorities who suggest an intractable conflict between cis women and trans women. These findings demonstrate a need to critically reflect upon the representativeness of exclusionary discourses that may be dominant in the media, as they may not reflect the norms of the majority. A failure to critique these narratives may result in the norm misperception effect we found whereby cis women—the majority of whom in our samples held inclusionary attitudes—do not think that others share their beliefs.

Acknowledgements The authors would like to thank Flavien Raynaud for his help in creating the materials for Study 5.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability The data, codebook, analysis code, and materials for all studies are available at https://osf.io/d5wga/?view_only=5b7b33f5d3094eeda2c77c16d5ab41fc The first two studies were exploratory

and not pre-registered. The remaining studies were pre-registered (Study 3: https://aspredicted.org/FKZ_MZF, Study 4: https://aspredicted.org/FZB_TK1, Study 5: https://aspredicted.org/75F_5PB)

Declarations

Ethics Approval and Consent to Participate All studies received ethics approval from The University of Queensland Ethics Committee. Study 1 (Approval: 2021/HE001517), Study 2 (Approval: 2021/HE001517), Study 3 (Approval: 2022/HE000979), Study 4 (Approval: 2022/HE000979–Amendment), Study 5 (2022/HE000979–Amendment).

Competing Interests The Authors declare that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ali, J. (2022, November 2). *Rishi Sunak reportedly planning Equality Act review to underline that 'sex means biological sex rather than gender'*. Wales Online. <https://www.walesonline.co.uk/news/rishi-sunak-reportedly-planning-review-25403893>
- Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 85–106). Psychology.
- Andre, P., Boneva, T., Chopra, F., & Falk, A. (2022). Misperceived social norms and willingness to act against climate change. *The Review of Economics and Statistics*, 1–46. https://doi.org/10.1162/rest_a_01468
- Armenta, B., Stroebe, K., Scheibe, S., Van Yperen, N., Stegeman, A., & Postmes, T. (2017). Permeability of group boundaries: Development of the concept and a scale. *Personality and Social Psychology Bulletin*, 43(3), 418–433. <https://doi.org/10.1177/0146167216688202>
- Armitage, L. (2020). Explaining backlash to trans and non-binary genders in the context of UK gender recognition act reform. *Journal of the International Network for Sexual Ethics & Politics*, 11–35. <https://doi.org/10.3224/insep.si2020.02>
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596–612. <https://doi.org/10.1037/0022-3514.63.4.596>
- Auerbach, J. D., Moran, L., Watson, C., Weber, S., Keatley, J. A., & Sevelius, J. (2020). We are all women: Barriers and facilitators to inclusion of transgender women in HIV treatment and support services designed for cisgender women. *AIDS Patient Care and STDs*, 34(9), 392–398. <https://doi.org/10.1089/apc.2020.0056>
- Aversa, I., Coleman, T., Travers, R., Coulombe, S., Wilson, C., Woodford, M. R., Davis, C., Burchell, D., & Schmid, E. (2022). I'm always worried: Exploring perceptions of safety and community inclusion among transgender people. *International Journal of Community Well-Being*, 5, 37–54. <https://doi.org/10.1007/s42413-021-00123-9>
- Barranti, M., Carlson, E., & Côté, S. (2017). How to test questions about similarity in personality and social psychology research: Description and empirical demonstration of response surface analysis. *Social Psychological and Personality Science*, 8(4), 465–475. <https://doi.org/10.1177/1948550617698204>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bennett, M. (2014). Intergroup social exclusion in childhood: Forms, norms, context, and social identity. *Journal of Social Issues*, 70(1), 183–195. <https://doi.org/10.1111/josi.12054>
- Bowen, A. M., & Bourgeois, M. J. (2001). Attitudes toward lesbian, gay, and bisexual college students: The contribution of pluralistic ignorance, dynamic social impact, and contact theories. *Journal of the American College Health Association*, 50(2), 91–96. <https://doi.org/10.1080/07448480109596012>
- Bursztn, L., González, A. L., & Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10), 2997–3029. <https://doi.org/10.1257/AER.20180975>
- Caldwell, A. R. (2022). *Exploring equivalence testing with the updated TOSTER R package*. PsyArXiv. <https://doi.org/10.31234/osf.io/ty8de>
- Charlton, K., & Bettencourt, B. A. (2001). In-group members' perceptions of a social mobility attempt: Effects of boundary permeability and outcome. *Group Dynamics: Theory Research and Practice*, 5(2), 81–91. <https://doi.org/10.1037/1089-2699.5.2.81>
- Cialdini, R., Kallgren, C., & Reno, R. (1991). A focus theory of normative conduct: A theoretical refinement and Reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology*, 24, 201–234. [https://doi.org/10.1016/S0065-2601\(08\)60330-5](https://doi.org/10.1016/S0065-2601(08)60330-5)
- De Souza, L., & Schmader, T. (2021). The misjudgment of men: Does pluralistic ignorance inhibit allyship? *Journal of Personality and Social Psychology*, 122(2), 265–285. <https://doi.org/10.1037/psp0000362>
- Dubueque, E., Ciano-Boyce, C., & Shelley-Sireci, L. (2002). Measuring misperceptions of homophobia on campus. In A. D. Berkowitz (Ed.), *The report on social norms: Working paper #4* (pp. 2–3). PaperClip Communications.
- Duncan, M., Pelled, A., Wise, D., Ghosh, S., Shan, Y., Zheng, M., & McLeod, D. (2020). Staying silent and speaking out in online comment sections: The influence of spiral of silence and corrective action in reaction to news. *Computers in Human Behavior*, 102, 192–205. <https://doi.org/10.1016/j.chb.2019.08.026>
- Edwards, J. (2001). Ten difference score myths. *Organizational Research Methods*, 4(3), 265–287. <https://doi.org/10.1177/109442810143005>
- Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow, & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 350–400). Jossey-Bass/Wiley.
- Eisner, L., Turner-Zwinkels, F., & Spini, D. (2021). The impact of laws on norms perceptions. *Personality and Social Psychology Bulletin*, 47(7), 1071–1083. <https://doi.org/10.1177/0146167220959176>
- Engender (2023, January 17). *A statement from Scottish civil society organisations on the UK Government's intervention on the Gender Recognition Reform (Scotland) Bill*. <https://www.engender.org.uk/news/blog/a-statement-from-scottish-civil-society-organisations-on-the-uk-governments-intervention-o/>

- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Flezzani, J., & Benshoff, J. (2003). Understanding sexual aggression in male college students: The role of self-monitoring and pluralistic ignorance. *Journal of College Counselling*, 6(1), 69–79. <https://doi.org/10.1002/j.2161-1882.2003.tb00228.x>
- Gaertner, L., Sedikides, C., & Cai, H. (2012). Wanting to be great and better but not average: On the pancultural desire for self-enhancing and self-improving feedback. *Journal of Cross-Cultural Psychology*, 43, 521–526. <https://doi.org/10.1177/0022022112438399>
- García, A., & Badge, J. (2021, June 9). *Transgender people in the Australian press: Bombarded by outright harassment*. Sydney Corpus Lab. <https://sydneycorpuslab.com/transgender-people-in-the-australian-press-bombarded-by-outright-harassment/>
- Gardner, A. (2023, April 25). *A complete breakdown of the J.K. Rowling transgender-comments controversy*. Glamour. <https://www.glamour.com/story/a-complete-breakdown-of-the-jk-rowling-transgender-comments-controversy>
- Geiger, N., & Swim, J. K. (2016). Climate of silence: Pluralistic ignorance as a barrier to climate change discussion. *Journal of Environmental Psychology*, 47, 79–90. <https://doi.org/10.1016/j.jenvp.2016.05.002>
- Green, P., & MacLeod, C. J. (2016). Simr: An R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Halbesleben, J. R. B., Wheeler, A. R., & Buckley, R. M. (2007). Understanding pluralistic ignorance in organizations: Application and theory. *Journal of Managerial Psychology*, 22(1), 65–83. <https://doi.org/10.1108/02683940710721947>
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis, second edition: A regression-based approach* (3rd ed.). Guilford Press.
- Heinze, G., Ploner, M., Jiricka, L., & Steiner, G. (2023). *logistf: Firth's bias-reduced logistic regression*. R package (version 1.26.0). <http://CRAN.R-project.org/package=logistf>
- Hines, S. (2019). The feminist frontier: On trans and feminism. *Journal of Gender Studies*, 28(2), 145–157. <https://doi.org/10.1080/09589236.2017.1411791>
- Hislop, M. (2022, May 10). *Women's rights and safety organisations condemn targeting of trans women by political candidates*. Women's Agenda. <https://womensagenda.com.au/latest/womens-rights-and-safety-organisations-condemn-targeting-of-trans-women-by-political-candidates/>
- Hornsey, M. J., Terry, D. J., & McKimmie, B. M. (2004). Willingness to speak out about gay law reform: Some cause for optimism. *Journal of Homosexuality*, 47(2), 47–61. https://doi.org/10.1300/J082V47N02_03
- Howell, J., & Ratliff, K. (2017). Not your average bigot: The better-than-average effect and defensive responding to implicit association test feedback. *British Journal of Social Psychology*, 56(1), 125–145. <https://doi.org/10.1111/bjso.12168>
- Ipsos (2023). *LGBT+ pride 2023*. <https://www.ipsos.com/en-uk/ipso-s-pride-2023-global-survey-two-in-three-britons-think-transgender-people-face-discrimination>
- Jacobs, L., & Meeusen, C. (2021). Coming out of the closet, also on the news? A longitudinal content analysis of patterns in visibility, tone and framing of LGBTs on television news (1986–2017). *Journal of Homosexuality*, 68(13), 2144–2168. <https://doi.org/10.1080/00918369.2020.1733352>
- John, T. (2022, July 18). *Britain's Conservative party leadership race is turning into a transphobic spectacle*. CNN. <https://edition.cnn.com/2022/07/17/uk/uk-conservative-leadership-trans-intl-gbr/index.html>
- Johnson, D., Terry, D. J., & Louis, W. R. (2005). Perceptions of the intergroup structure and anti-Asian prejudice among white Australians. *Group Processes and Intergroup Relations*, 8(1), 53–71. <https://doi.org/10.1177/1368430205048616>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1–26. <https://doi.org/10.18637/jss.v082.i13>
- LaBrie, J. W., Hummer, J. F., Grant, S., & Lac, A. (2010). Immediate reductions in misperceived social norms among high-risk college student groups. *Addictive Behaviors*, 35(12), 1094–1101. <https://doi.org/10.1016/j.addbeh.2010.08.003>
- Lakens, D. (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- Lalonde, R. N., & Silverman, R. A. (1994). Behavioral preferences in response to social injustice: The effects of group permeability and social identity salience. *Journal of Personality and Social Psychology*, 66(1), 78–85. <https://doi.org/10.1037/0022-3514.66.1.78>
- Latané, B., & Darley, J. (1970). *The unresponsive bystander: Why doesn't he help?* Appleton-Century-Crofts.
- Lees, J., Cetron, J. S., Vollberg, M. C., Reggev, N., & Cikara, M. (2020). *Intentions to comply with COVID-19 preventive behaviors are associated with personal beliefs, independent of perceived social norms*. PsyArXiv. <https://psyarxiv.com/97jry/>
- LGBTIQ+ Health Australia (2022, June 10). *Media outlets must prioritise trans and gender diverse voices in FINA decision coverage*. https://www.lgbtiqhealth.org.au/media_outlets_must_prioritise_trans_and_gender_diverse_voices_in_fina_decision_coverage
- Lu, C., & Jurgens, D. (2022). The subtle language of exclusion: Identifying the toxic speech of trans-exclusionary radical feminists. *Proceedings of the Sixth Workshop on Online Abuse and Harms*, 79–91. <https://doi.org/10.18653/v1/2022.woaah-1.8>
- Martin, S., & McGowan, M. (2022, April 19). Katherine Deves claims key role in controversial bill to ban trans women from women's sport. *The Guardian*. <https://www.theguardian.com/australia-news/2022/apr/19/katherine-deves-claims-key-role-in-controversial-bill-to-ban-trans-women-from-womens-sport>
- Mattern, J. L., & Neighbors, C. (2004). Social norms campaigns: Examining the relationship between changes in perceived norms and changes in drinking levels. *Journal of Studies on Alcohol*, 65(4), 489–493. <https://doi.org/10.15288/JSA.2004.65.489>
- Miller, D., & McFarland, C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology*, 53(2), 298–305. <https://doi.org/10.1037/0022-3514.53.2.298>
- Miller, D. T., & Prentice, D. A. (2016). Changing norms to change behavior. *Annual Review of Psychology*, 67, 339–361. <https://doi.org/10.1146/ANNUREV-PSYCH-010814-015013>
- Miyajima, T., & Yamaguchi, H. (2017). I want to but I won't: Pluralistic ignorance inhibits intentions to take paternity leave in Japan. *Frontiers in Psychology*, 8, Article 1508. <https://doi.org/10.3389/FPSYG.2017.01508>
- Morgenroth, T., Kirby, T. A., Zehnter, M. K., & Ryan, M. K. (2024). Sex wars and TERF wars: The divisiveness of who is included in feminism. *Sex Roles*, 90(6), 676–705. <https://doi.org/10.1007/S11199-024-01473-2>
- Munsch, C., Weaver, J., Bosson, J., & O'Connor, L. (2018). Everybody but me: Pluralistic ignorance and the masculinity contest. *Journal of Social Issues*, 74(3), 551–578. <https://doi.org/10.1111/josi.12282>

- Nekmat, E., & Gonzenbach, W. J. (2013). Multiple opinion climates in online forums: Role of website source reference and within-forum opinion congruency. *Journalism and Mass Communication Quarterly*, 90(4), 736–756. <https://doi.org/10.1177/1077699013503162>
- Noelle-Neumann, E. (1993). *The spiral of silence: Public opinion—our social skin*. The University of Chicago Press.
- Outten, H. R., & Lawrence, M. E. (2020). Intergroup threat and heterosexual cisgender women's support for policies regarding the admittance of trans women at a women's college. *Social Sciences*, 9, Article 208. <https://doi.org/10.3390/socsci9110208>
- Outten, H. R., Lee, T., & Lawrence, M. E. (2019). Heterosexual women's support for trans-inclusive bathroom legislation depends on the degree to which they perceive trans women as a threat. *Group Processes & Intergroup Relations*, 22(8), 1094–1108. <https://doi.org/10.1177/1368430218812660>
- Pain, R., Cygnus Support, & O'Neil, S. (2021). *One of the lasses: Trans inclusion and safety in abuse services*. <https://www.ncl.ac.uk/mediav8/gps/files/One%20of%20the%20Lasses%20Report.pdf>
- Pearce, R., Erikainen, S., & Vincent, B. (2020). TERF wars: An introduction. *Sociological Review*, 68(4), 677–698. <https://doi.org/10.1177/0038026120934713>
- Perkins, H., Craig, D., & Perkins, J. (2011). Using social norms to reduce bullying: A research intervention among adolescents in five middle schools. *Group Processes & Intergroup Relations*, 14(5), 703–722. <https://doi.org/10.1177/1368430210398004>
- Perryman, M., Davis, C., & Hull, S. (2018). Perceived community acceptance of same-sex marriage: Persuasive press, projection, and pluralistic ignorance. *International Journal of Public Opinion Research*, 30(2), 305–315. <https://doi.org/10.1093/ijpor/edx003>
- Portelinha, I., & Elcherth, G. (2016). From marginal to mainstream: The role of perceived social norms in the rise of a far-right movement. *European Journal of Social Psychology*, 46(6), 661–671. <https://doi.org/10.1002/EJSP.2224>
- Prentice, D., & Miller, D. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64(2), 243–256. <https://doi.org/10.1037/0022-3514.64.2.243>
- Pringle, R., & Denison, E. (2022). Examining World Rugby's transgender ban and the perspectives of cisgender women who play rugby in England, Canada and Australia. In A. D. Greey & H. J. Lenskyj (Eds.), *Justice for trans athletes: Challenges and struggles* (pp. 75–94). Emerald Group Publishing Ltd. <https://doi.org/10.1108/978-1-80262-985-920221006>
- Rapinoe, M. (2021, March 28). Megan Rapinoe: Bills to ban transgender kids from sports try to solve a problem that doesn't exist. *The Washington Post*. <https://www.washingtonpost.com/opinions/2021/03/28/megan-rapinoe-transgender-kids-sports-ban/>
- Richard, N., & Wright, S. C. (2010). Advantaged group members' reactions to tokenism. *Group Processes & Intergroup Relations*, 13(5), 559–569. <https://doi.org/10.1177/1368430210362227>
- Ryan, C., & Reicher, S. (2019). An analysis of the nature and use of promigrant representations in an antideportation campaign. *Political Psychology*, 40(3), 583–598. <https://doi.org/10.1111/pops.12526>
- Sargent, R. (2021). *Pluralistic ignorance of attitudes toward mental health services among college students* (Publication No. 1329). [Doctoral dissertation, Syracuse University]. <https://surface.syr.edu/etd/1329>
- Sargent, R., & Newman, L. S. (2021). Pluralistic ignorance research in psychology: A scoping review of topic and method variation and directions for future research. *Review of General Psychology*, 25(2), 163–184. <https://doi.org/10.1177/1089268021995168>
- Schönbrodt, F. D., & Humberg, S. (2023). *RSA: An R package for response surface analysis* (version 0.10.6). <https://cran.r-project.org/package=RSA>
- Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of self-categorization. *Self and Identity*, 1(4), 353–376. <https://doi.org/10.1080/152988602760328012>
- Schumann, S., & Klein, O. (2015). Substitute or stepping stone? Assessing the impact of low-threshold online collective actions on offline participation. *European Journal of Social Psychology*, 45(3), 308–322. <https://doi.org/10.1002/ejsp.2084>
- Serano, J. (2018, October 17). *Putting the transgender activists versus feminists debate to rest*. Medium. <https://juliaserano.medium.com/putting-the-transgender-activists-versus-feminists-debate-to-rest-e18075df03af>
- Shanock, L. R., Baran, B. E., Gentry, W. A., Patterson, S. C., & Heggestad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming limitations of difference scores. *Journal of Business and Psychology*, 25(4), 543–554. <https://doi.org/10.1007/s10869-010-9183-4>
- Sokoloski, R., Markowitz, E., & Bidwell, D. (2018). Public estimates of support for offshore wind energy: False consensus, pluralistic ignorance, and partisan effects. *Energy Policy*, 112, 45–55. <https://doi.org/10.1016/j.enpol.2017.10.005>
- Stangor, C., Sechrist, G. B., & Jost, J. T. (2001). Changing racial beliefs by providing consensus information. *Personality and Social Psychology Bulletin*, 27(4), 486–496. <https://doi.org/10.1177/0146167201274009>
- Steiner, L., Fejes, F., & Petrich, K. (1993). Invisibility, homophobia and heterosexism: Lesbians, gays and the media. *Critical Studies in Mass Communication*, 10(4), 395–422. <https://doi.org/10.1080/15295039309366878>
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of inter-group conflict. In W. G. Austin, & S. Worchel (Eds.), *The social psychology of inter-group relations* (pp. 33–47). Brooks/Cole.
- Tankard, M., & Paluck, E. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review*, 10(1), 181–211. <https://doi.org/10.1111/sipr.12022>
- Tankard, M., & Paluck, E. (2017). The effect of a supreme court decision regarding gay marriage on social norms and personal attitudes. *Psychological Science*, 28(9), 1334–1344. <https://doi.org/10.1177/0956797617709594>
- Thomas, J., Ursell, A., Robinson, E., Aveyard, P., Jebb, S., Herman, C., & Higgs, S. (2017). Using a descriptive social norm to increase vegetable selection in workplace restaurant settings. *Health Psychology*, 36(11), 1026–1033. <https://doi.org/10.1037/hea0000478>
- Travers, A., & Deri, J. (2011). Transgender inclusion and the changing face of lesbian softball leagues. *International Review for the Sociology of Sport*, 46(4), 488–507. <https://doi.org/10.1177/1012690210384661>
- Turner, J. C., & Reynolds, K. J. (2001). The social identity perspective in intergroup relations: Theories, themes and controversies. In R. Brown, & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 133–152). Blackwell.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.
- Vallentin-Holbech, L., Rasmussen, B. M., & Stock, C. (2018). Effects of the social norms intervention the GOOD life on norm perceptions, binge drinking and alcohol-related harms: A cluster-randomised controlled trial. *Preventive Medicine Reports*, 12, 304–311. <https://doi.org/10.1016/j.pmedr.2018.10.019>
- van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives.

- Psychological Bulletin*, 134(4), 504–535. <https://doi.org/10.1037/0033-2909.134.4.504>
- van Zomeren, M., Leach, C., & Spears, R. (2010). Does group efficacy increase group identification? Resolving their paradoxical relationship. *Journal of Experimental Social Psychology*, 46, 1055–1060. <https://doi.org/10.1016/j.jesp.2010.05.006>
- Verkuyten, M., & Reijerse, A. (2008). Intergroup structure and identity management among ethnic minority and majority groups: The interactive effects of perceived stability, legitimacy, and permeability. *European Journal of Social Psychology*, 38(1), 106–127. <https://doi.org/10.1002/ejsp.395>
- White, K., Hogg, M., & Terry, D. (2002). Improving attitude-behavior correspondence through exposure to normative support from a salient ingroup. *Basic and Applied Social Psychology*, 24(2), 91–103. https://doi.org/10.1207/S15324834BASP2402_2
- Willem, C., Platero, R., & Tortajada, I. (2022). Trans-exclusionary discourses on social media in Spain. In T. Krijnen, P. G. Nixon, M. D. Ravenscroft, & C. M. Scarcelli (Eds.), *Identities and intimacies on social media* (1st ed., pp. 185–200). Routledge. <https://doi.org/10.4324/9781003250982-15>
- Wright, S. C. (2001). Restricted intergroup boundaries tokenism, ambiguity, and the tolerance of injustice. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations* (pp. 223–254). Cambridge.
- Yamin, P., Fei, M., Lahlou, S., & Levy, S. (2019). Using social norms to change behavior and increase sustainability in the real world: A systematic review of the literature. *Sustainability*, 11, Article 5847. <https://doi.org/10.3390/su11205847>
- Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2020). The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychological Bulletin*, 146(2), 118–149. <https://doi.org/10.1037/bul0000218>
- Zhang, X. X., Zheng, J., Liu, L., Zhao, X., & Sun, X. M. (2014). The effect of group boundary permeability on intergroup prejudice: The case of rural-to-urban migrants in China. *Journal of Pacific Rim Psychology*, 8(2), 53–61. <https://doi.org/10.1017/prp.2014.7>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.