Assessing Gender Essentialist Beliefs Towards Youth Athletes: Development and Preliminary Validation of the Gender Essentialist Beliefs Towards Girls in Sport Scale (GEGS)

Abstract

Despite evidence refuting the legitimacy of inherent and unchanging qualities and the existence of a gender sport continuum, a prevailing gender essentialist ideology persists among coaches. This ideology perpetuates the belief in a binary, hierarchical, and static structure of gender, which imparts higher cultural and social values to men's sport compared to women's and impacts girls' development, participation, and enjoyment in sport. Some interventions have been proposed to address this issue; however, at present, no validated scale exists to assess intervention outcomes related to coaches' gender essentialist beliefs. The aim of the current study was therefore to develop and validate the Gender Essentialist Beliefs Towards Girls in Sport Scale (GEGS), to assess gender essentialism among sport coaches. Content validity was judged by a panel of experts ($N = 4$) and through interviews with coaches ($N = 5$) across various sports and experience levels. Following initial item iteration, the GEGS was assessed with 633 coaches for reliability and validity. The final scale comprised 20 items and showed good reliability (internal consistency, test-retest reliability, inter-item and item-total correlations), validity (convergent validity, differentiation between known groups), and factor structure. We also found partial support for the discriminant validity of the GEGS. Overall, the GEGS provides a valuable tool for addressing gender essentialism in sport coaching and informing coach education programs. Future research should further validate the scale and explore its sensitivity to change in gender essentialist beliefs over time.

Keywords: Coaching, Female Athletes, Gender Essentialism, Gender Stereotypes, Sport.

Introduction

Historically considered a masculine domain, girls and women are entering and engaging in sport in record numbers, allowing them to access the physical, social, and psychological benefits that sport provides.[1] While girls' and women's inclusion and proven achievements in sport are beginning to challenge masculine hegemony, barriers for gender equity in sport persist. Although women athletes have proven their athletic excellence and capabilities within sport environments, girls and women drop out at higher rates than boys and men, and participate less compared to boys and men.[1,2,3] Gender diverse and transgender athletes also find themselves excluded, marginalized, and invalidated in essentialist, binary, and sex-segregated sport environments.[4,5] A significant reason for higher dropout, lower participation rates, and general inequalities among girls and women in sport are due, in part, to gender essentialist viewpoints imbued within sport structures.[6,7,8] In particular, gender essentialist stereotypes espoused by sport coaches and other key social agents within sport often go unchecked and deeply harm girls' and women's sports experiences, contributing to their departure from sports.[6,7,9,10] The purpose of this study was therefore to develop and validate a tool to accurately measure levels of gender essentialism among sport coaches.

Gender Essentialism in Sport

Essentialism is the philosophical belief that things, including people or categories, have a set of inherent, fixed characteristics or "essences" that define what they fundamentally are and naturally unifies its members. According to essentialism, these essences determine the behavior, attributes, and identity of an entity or group and are seen as natural, universal, and unchanging.[11] *Gender essentialism* is a commonplace and pervasive gender ideology, which posits that gender is a binary, immutable, hierarchical, and static structure, wherein women and men have distinct and separate qualities, characteristics, and abilities. This perspective assumes that gender differences are innate, fixed, and universal, rather than

shaped by cultural, social, or environmental influences. According to gender essentialism, women and men are naturally suited for distinct roles in society, often reinforcing traditional gender roles.[7,12,13] Essentialist beliefs prevail, despite evidence from scholars in psychology, neuroscience, and gender studies who refute the legitimacy of inherent and unchanging qualities and argue the existence of a gender sport continuum in which women and men's athletic abilities overlap due to both physical and social factors.[3,14,13] Nonetheless, a sex-segregated sport system, which permeates essentialist stereotypes imparting higher cultural and social values to men's sport compared to women's, persists.[15,16]

In sport, gender essentialism manifests through beliefs of men's athletic superiority compared to women, men's greater suitability for sport leadership positions, and stereotypes of women and men in sport.[6,7,10,15] Gender essentialist stereotypes can also limit athlete choices, experiences, and participation in particular sports through gendered sport typing, where girls are encouraged to engage in sports that are viewed as feminized and socially acceptable, and avoid sports that are perceived as masculine.[17,18] Essentialism can have negative consequences on athletes' sports experiences, such as stereotype threat, discrimination, lower enjoyment, dropout, and poor mental health.[8,12,19,20] Essentialist beliefs also serve to justify gender inequality, erase diversity of athlete experiences, and invalidate the identities of gender diverse and transgender athletes.[4,6,7,21] Although gender essentialism has negative impacts on boys and men in sport, through imposition of narrow standards of idealized masculine performance, essentialist stereotypes have particularly high negative effects on girls' experiences in sport and contribute to their sport dropout.[7,8,22]

The Role of Coaches

Sport coaches deeply impact athletes' experiences in sport, including athletes' mental health, body image, and retention in sport.[23,24,25,26] Coaches are also powerful mediators of societal gender norms and hierarchies in sport. Research on coach discourses has uncovered

how coaches often perceive girls and women in sport through gender essentialist lenses, which in turn perpetuate stereotypes of girls and women as emotional, non-competitive, and inferior athletes compared to boys and men.[9,10,27,28,29] For instance, youth sport coaches often view girls as socially, physically, and psychologically different and deficient compared to boys in sport.[10] Coaches' gender essentialist beliefs can impact their ability to effectively support women athletes' development, success, and sports enjoyment.[29,30] Specifically, gender essentialist coaching practices can limit girls' potential in sport, protect unequal treatment of girls and boys in sport, police the gender performance of those who are viewed as not gender compliant, position girls as deviant to masculine standards, and uphold the hegemony of White, heterosexual, cisgender femininity in sport at the expense of racially diverse, non-cisgender, or non-heterosexual athletes.[2,6,15,27,34]

Despite the risks associated with gender essentialist viewpoints on girls' and women's sports experiences, education and tools to identify and dismantle essentialist and stereotypical coaching practices are limited.[10,26,29,30] Coaches are often unaware and uneducated on topics related to gender and social justice, which allows biases to go undetected and continue to negatively impact athletes.[23,32,33] Within gender studies and psychology research, validated measures to assess gender essentialist beliefs in society exist, such as the Gender Essentialism Scale and the Gender Essentialism Measure.[12,21] However, no sport-specific tool to assess gender essentialist biases currently exists, as these scales measure gender essentialism in general social environments and political contexts (e.g., "Mothers are naturally more sensitive to a baby's feelings than fathers are"[12]; "Men are naturally more suited for leadership roles than women"[21]). Current gender essentialism scales are inadequate to assess gender essentialism in sport contexts, given that sport structures are uniquely positioned as highly masculine-dominated, separated by sex-segregation, rife with biological assumptions of athleticism and sport ability, and historically inaccessible to girls and women. The Gender

Essentialism Scale and the Gender Essentialism Measure thus fail to capture the specific intricacies of sport contexts.[12,21] Therefore, a research gap exists relating to a gender essentialism measure in sport. A sport-specific assessment of gender essentialism is necessary to identify gender essentialist beliefs in sport coaching at a variety of levels and environments, detect the extent to which coaches may perpetuate gender biases and stereotypes, and examine the effectiveness of potential educational tools seeking to address and correct gender essentialism in sport structures.[3,7,9,15]

The Current Study

The purpose of this study was to develop and validate the Gender Essentialist Beliefs Towards Girls in Sport Scale, which measures levels of coaches' gender essentialist beliefs towards girls in sport-specific environments. The creation and validation of this scale aims to provide a rigorously tested, evidence-based tool to detect essentialist gender biases about girls among sport coaches, which can subsequently help address how coach development programs can begin to dismantle gender stereotypes, retain girls in sport, and improve environments for not just girls, but for all athletes.

Materials and Methods

Study Design

A combination of research methods was employed to develop and validate the Gender Essentialist Beliefs Towards Girls in Sport Scale (GEGS). Recent recommendations for developing and validating the new scale, as well as when reporting the study findings were followed.[34,35] During the development phase (Phase 1), qualitative data were collected through expert input and cognitive interviews. In the testing phase (Phase 2), quantitative data were obtained via online surveys. The University of Minnesota – Twin Cities Institutional Review Board approved all procedures (ref no. STUDY00017483). Refer to Figure 1 for details on procedures, participant recruitment, and attrition.

[FIGURE 1 NEAR HERE]

Phase 1: Development of Scale Items

Item Development

The development of the initial scale items occurred in four stages. First, we drew from other existing measures on gender, sex, and sexuality to compile a list of items that were potentially relevant to the GEGS (see Appendix A, Supplementary Materials). All items were modified/adapted to the sports context, specifically for youth athletes (i.e., boys and girls). Second, we conducted a thorough review of the existing literature on gender in sport, to extract additional items relevant to this domain. Third, several items were included based on the expertise of the core research team, comprising seasoned researchers in gender, sport psychology, and coaching science, some of whom have personal experiences as athletes and coaches. Specifically, members of the research team have worked as coaches, coach educators, presenters at coaching conferences, and have conducted research on coach discourses about girls and women. This collective experience within the coaching realm inspired many of the items in the first iteration of the GEGS. Fourth, we ensured that items covered all dimensions of gender essentialism, specifically distinctiveness (e.g., "girls are less competitive than boys in sport"), naturalness (e.g., "boys are naturally stronger than girls"), and immutability (e.g., "differences between boys and girls in sport performance are largely determined by genetic predisposition"). The initial version of the GEGS consisted of 56 items, rated on a five-point Likert scale ranging from 1 (*I Strongly Disagree*) to 5 (*I Strongly Agree*). Importantly, we acknowledge that all items in the scale express essentialist viewpoints and conform to notions of a gender binary. Although the aim of the scale is to help identify and eradicate gender essentialist beliefs, most of sport is structured around a sex-segregated gender binary centered around gender essentialism.[7,15] Therefore, items were

created that we believed would align with most coaches' experiences, despite the risk,

likelihood, and irony of the items reinforcing a gender essentialist perspective.

Expert Reviews

Following the initial item generation, a diverse group of experts was assembled to

evaluate the scale's content validity. Seven scholars spanning the fields of sport psychology,

sport sociology, and sport management were identified and invited via email to assess the

first draft of the GEGS. Feedback was received from four of the experts within the timeframe

of this study (four women), who reviewed the initial 56-item scale. Following the expert

reviews, 21 items were deleted due to irrelevance, lack of clarity, and being double-barreled

and eight items were reworded or simplified. Moreover, two items were added related to

transgender athletes, as expert reviewers highlighted the relation of gender essentialism to

discrimination against transgender athletes. The scale post-expert feedback was comprised of

37 items.

Cognitive Interviews

Cognitive interviewing assessed survey questions' clarity and relevance with five

youth and collegiate sport coaches who were recruited via convenience sampling. The sample

comprised two female and three male coaches across basketball, softball, track and field,

soccer, and tennis. One coach identified as Black, while four coaches identified as White.

Coaches completed the scale before taking part in the interviews, followed by retrospective

probing based on the Cognitive Model of Survey Response Process.[35,36] Interviews explored

comprehension, retrieval, judgment, and response processes.

Following the cognitive interviews, an additional seven items were deleted due to

being overly complex, not relevant to the coaches' experiences, or lack of clarity.

Additionally, scale completion instructions were simplified and clarified based on the

coaches' feedback and definitions were added for terms that may be confusing for some

coaches, such as the differences between 'gender' and 'sex', as well as definitions of transgender woman and transgender man. Therefore, the final scale at the end of phase 1 comprised 30 items (Appendix B, Supplementary Materials).

Phase 2: Testing the Scale

Sampling

The sample size was predetermined based on prior recommendations, suggesting 200–300 respondents for factor analysis.[34,35,36,39] Participants were recruited through emails to various sports organizations and university newsletters. Inclusion criteria required participants to be over 18 years old, identify as a coach in any sport, and have coached at any level or age group within the past two years. Surveys were administered twice, one week apart (Time 1 and Time 2). Coaches did not receive financial compensation for taking part in the study.

Measures

Gender Essentialist Beliefs Towards Girls in Sport Scale (GEGS). The GEGS was developed for the purposes of this research. Participants rated 30 items on a five-point Likert scale ranging from 1 (*I Strongly Disagree*) to 5 (*I Strongly Agree*). Example items included: "Girls are less competitive than boys in sports" and "Sports that are associated with grace and beauty are more appropriate for girls (e.g., dance, skating, gymnastics)". Higher scores on the GEGS indicate greater gender essentialist beliefs towards girls in sport.

Gender Essentialism Scale (GES). The GES examines individuals' gender essentialist beliefs[12], comprises 25 items, and is rated on a five-point Likert scale ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). Example items include: "People tend to be either masculine or feminine: there's not much middle ground" and "Women and men are fundamentally different". Higher scores on the GES indicate greater gender essentialist beliefs. The GES has shown good validity and reliability in previous research and in the

current study (Cronbach's α = .903).[12,40] For the purposes of the current research, the GES was used as a measure of *convergent* validity.

Sexual Prejudice in Sport Scale (SPSS). The SPSS assesses attitudes toward lesbian women and gay men in sports.[41] It comprises 19 items and is rated on a seven-point Likert scale ranging from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). The SPSS includes three subscales: (1) Open Rejection, which assesses the blatant prejudice expressed toward lesbian and gay (LG) people (e.g., "LG persons should not be allowed to be trainers/coaches"); (2) Denial of Visibility, which evaluates attitude toward the coming out of LG people (e.g., "I'd feel uncomfortable if LG athletes talked about their sexual orientation openly"); and (3) Gendering Performance, which corresponds to gender stereotypes about the performance/skills of LG people (e.g., "Gay men could not be strong in a combat sport"). Higher scores on the SPSS indicate greater sexual prejudice in sport. The SPSS has shown good validity and reliability in previous research and in the current study (Cronbach's α = .850).[41] For the purposes of the current research, the SPSS was used as a measure of *discriminant* validity.

Social Dominance Orientation Scale (SDOS). The SDOS examines social dominance orientation, or the extent to which one desires that one's in-group dominate and be superior to out-groups and one's degree of preference for inequality among social groups.[42] It comprises 16 items and is rated on a seven-point Likert scale ranging from 1 (*Very Negative*) to 7 (*Very Positive*). Example items include: "To get ahead in life, it is sometimes necessary to step on other groups" and "It's probably a good thing that certain groups are at the top and other groups are at the bottom". Higher scores on the SDOS indicate a higher degree of preference for inequality among social groups. The SDOS has shown good validity and reliability in previous research and in the current study (Cronbach's α = .892).[42,43]

Demographic Information. We gathered the following demographic information: (1) age; (2) gender identity; (3) racial identity; (4) country (and state) of residence; (5) coaching tenure; (6) sport(s) coached; (7) coaching level; (8) athletes' gender; (9) athletes' age; and (10) information about previous training or education around gender or gender stereotypes.

Qualitative Feedback. At the end of the survey, we asked coaches the following question to gain additional feedback about the survey: "Do you have any other thoughts about this survey that you would like to share with us?".

Data Analyses

Data analyses were conducted in SPSS (version 27.0) and AMOS (version 28.0). Initially, analyses were performed to check for normality (skewness and kurtosis $\leq \pm 2.58$). All questionnaires (i.e., GEGS, GES, SPSS, SDOS) showed normal distributions. The dataset was randomly split into approximately equal halves for exploratory factor analysis (EFA) and subsequent confirmatory factor analysis (CFA) to examine the resultant factor structure.

Using recent guidelines for factor analysis[35,44] in order to ascertain the factor structure of the GEGS, initially, a parallel analysis was conducted using the rawpar.sps script developed by O'Connor.[45] Based on permutations of the raw data, 1,000 datasets were generated using a common factor analysis approach, indicating nine eigenvalues, or factors, from the raw data that were above the 95th percentile estimates created by the Monte Carlo simulation. Based on the results of the parallel analysis, an initial EFA was conducted using a Maximal Likelihood (ML) extraction method with Promax rotation and Kaiser normalization, specifying a nine-factor solution. However, in analyzing the results of the EFA, it became evident that five of the nine factors contained too few items (< 3), with one factor comprising only one item. Subsequently, an unrestricted EFA was conducted, considering the Guttman-Kaiser criterion (the number of eigenvalues above 1) and the scree plot to determine how many factors to retain.[35,43,46,47] The minimum factor loading criteria was set to .40.[12,34,35] The

communality of the scale, which indicates the amount of variance in each dimension, was also assessed to ensure acceptable levels of explanation.[35]

Subsequently, model fit was assessed via CFA. Relative and absolute fit indices of the models were computed to determine how many factors to retain and to assess the model fit to the data. The goodness of fit indices included the relative chi-square (χ^2/df: values ≤ 3 and ≤ 2 indicate acceptable and good fit, respectively), the Root Mean Square Error of Approximation (RMSEA 90% CI: values $\leq .08$ and $\leq .06$ indicate acceptable and good fit, respectively), the Comparative Fit Index (CFI: values $\geq .90$ and $\geq .95$ indicate acceptable and good fit, respectively), the Tucker-Lewis Index (TLI: values $\geq .90$ and $\geq .95$ indicate acceptable and good fit, respectively), and the Standardized Root Mean Square Residual (SRMR: values $\leq .10$ and $\leq .08$ indicate acceptable and good fit, respectively).[35,48,44]

Test-retest reliability was assessed using Pearson's correlation coefficient, to evaluate the stability of the scale from Time 1 to Time 2 (one week later); linear regression, to assess whether Time 1 scores on the GEGS predicted GEGS scores at Time 2; and the intraclass correlation coefficient (ICC). Internal consistency was evaluated using Cronbach's alpha ($\alpha \geq .80$ was considered acceptable) and adjusted and non-adjusted item-total correlations. Cohen's guidelines of small ($r \geq .10$), moderate ($r \geq .30$), and large ($r \geq .50$) were used when interpreting correlations.[34,51] The ICC[52] measures the reliability or agreement between symmetrical measurements within a group.[35] ICC values range from 0 to 1, with values below 0.40 indicating poor agreement, 0.40–0.59 indicating fair to moderate agreement, 0.60–0.74 indicating good to very good agreement, and values above 0.75 indicating excellent agreement.[34,52]

Convergent and discriminant validity were assessed by correlating the total score of the GEGS with the total score of the GES and the SPSS, respectively.[11,40] Convergent validity is evidenced by moderate to strong correlations of the total scores of scales that measure

similar or related constructs, while discriminant validity is provided by small to moderate

correlations between scores of scales that measure distinct constructs.[34] We also explored the

relationship between our scale and the SDOS, although no prior hypotheses were made for

this relationship. Furthermore, we conducted a series of independent t-tests to assess

differences in GEGS scores based on known groups, including coach gender (1 = male, 2 =

female) and previous training on gender and/or gender stereotypes (1 = yes, 2 = no). Cohen's

guidelines of small ($d \geq .20$), medium ($d \geq .50$), and large ($d \geq .80$) were used when

interpreting t-tests.[53] Finally, we conducted simple linear regression to predict GEGS scores

based on coaches' age.

Results

Participants

Participant recruitment yielded 820 responses. Responses were removed prior to

analysis due to duplicate data ($n = 17$), failing the reCAPTCHA test ($n = 9$), response

identified as fraudulent or bot ($n = 26$), missing data on primary outcome ($n = 111$), failing

two or more attention checks (out of four; $n = 19$), and failing the consistency check (i.e., age

and year of birth; $n = 5$), leaving 633 participants on the Time 1 survey. For the follow-up

survey (Time 2), of the initial 240 responses, forty responses were removed due to duplicate

data ($n = 5$), response identified as fraudulent or bot ($n = 6$), missing data on primary

outcome ($n = 11$), and inability to match response to Time 1 data ($n = 18$).

The majority of participants in the final sample identified as women ($n = 289$, 53.7%),

White ($n = 443$, 82.3%), and residing in the United States (US; $n = 521$, 97.0%). Participants

ranged in age (18–81 years; $M_{age} = 45.63$, $SD = 13.67$ years) and coaching experience (0–54

years; $M_{experience} = 19.05$, $SD = 12.38$ years). Most coaches reported having received previous

education or training on the topic of gender and/or gender stereotypes ($n = 289$, 54.4%).

Participants worked across a wide range of sports including, but not limited to, yoga,

wrestling, basketball, soccer, volleyball, softball, lacrosse, track and field, ultimate frisbee, fencing, cross country, cheer and dance, tennis, triathlon, swimming, sailing, rowing, running, skiing, baseball, rugby, golf, gymnastics, and ice hockey. The data were randomly split into two independent samples for EFA (sample 1; $N = 320$) and CFA (sample 2; $N = 313$). Follow-up data from the Time 2 survey were matched with the Time 1 survey for participants who completed both assessment time points (sample 3; $N = 200$). Full participant characteristics for each sample are presented in Table 1.

[TABLE 1 NEAR HERE]

Factor Structure and Invariance

Exploratory Factor Analysis

Results of the unrestricted EFA indicated that all communalities were above .40 and there was no cross-loading of items. The size of the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO = .850) revealed that the GEGS items had adequate common variance for factor analysis, and the significance of Bartlett's Test of Sphericity, $\chi^2(n = 276) = 2749.719$, $p < .001$, indicated that the correlation matrix was factorable.[51] The factor solution derived from this analysis yielded seven factors, which accounted for 51.20% of the variation in the data (factor 1 = 10.22%; factor 2 = 17.92%; factor 3 = 8.69%; factor 4 = 4.61%; factor 5 = 3.76%; factor 6 = 3.49%; factor 7 = 2.51%). However, seven items failed to load onto any dimension significantly and were removed from further analysis one by one (items 4, 5, 14, 17, 22, 25, 28). EFA was repeated after excluding these items, showing a seven-factor structure (KMO = .851) that explained a total of 52.71% of the variance among the items (factor 1 = 10.51%; factor 2 = 18.42%; factor 3 = 8.96%; factor 4 = 4.71%; factor 5 = 3.92%; factor 6 = 3.64%; factor 7 = 2.53%). Bartlett's Test of Sphericity proved to be significant, $\chi^2(n = 253) = 2683.091$, $p < .001$, and all communalities were over the required value of .40.

All eigenvalues were over the required criterion of 1 (factor 1 = 6.527; factor 2 = 2.220;
factor 3 = 1.569; factor 4 = 1.519; factor 5 = 1.187; factor 6 = 1.043; factor 7 = 1.014).

However, several factors contained a relatively low number of items. Researchers
have recommended an absolute minimum of three items per factor to allow for further
exploration through CFA and at least five items for future measurement tools.[35] This is
because factors with only two or three items may encounter identification issues in CFA due
to overidentified models. In the current iteration of the GEGS, only factor 1 comprised five
items, factors 2 and 5 comprised four items, factors 6 and 7 comprised three items, and
factors 3 and 4 comprised two items each. Therefore, in line with these recommendations and
previous scales assessing similar constructs, we opted to proceed with a single-factor
solution.[12,35] This solution aligns with theoretical considerations related to gender
essentialism as a complex phenomenon in sport and with other validated scales that examine
gender essentialism in other contexts.[6,7,12] Similar to Skewes et al., a single-factor solution
reflects the GEGS's ability to capture a coherence of gender essentialism in the sport context
and sport coaches' views of girls in sport.[12] A single-factor solution further aligns with the
principles of parsimony and interpretability, allowing for a simplified interpretation of the
construct being assessed.[35] The full scale post-EFA showed high internal consistency
(Cronbach's α = .892) and consisted of 23 items. A summary of the EFA results is presented
in Table 2.

[TABLE 2 NEAR HERE]

Confirmatory Factor Analysis

Thereafter, we performed CFA to confirm the 23-item solution. In the original CFA,
three variables showed loadings < .40 (items 1, 11, and 12) and were therefore removed from
the model. The subsequent model with 20 items was assessed, and initially showed poor
model fit, $\chi^2(170) = 1,035.428$, $p < .001$; $\chi^2/df = 6.091$; TLI = .585; CFI = .629; RMSEA

= .128, p < .001; SRMR = .0960. After reviewing suggested modification indices, covariances were added between the item error terms. The revised model with modification indices showed good model fit, $\chi^2(133) = 259.883$, p < .001; χ^2/df = 1.954; TLI = .922; CFI = .946; RMSEA = .055, p = .185; SRMR = .0528, and was therefore retained. All item loadings were strong, ranging from .40–.69 (see Figure 2).

[FIGURE 2 NEAR HERE]

Reliability

Reliability analyses showed high internal consistency for the 20-item GEGS (α = .883). Item-total correlations revealed significant positive correlations between each item and the total score (rs = .447–.699, p < .01), indicating that higher scores on individual items were associated with a higher total GEGS score. Adjusted item-total correlations revealed a similar pattern (rs = .379–.648, p < .01). Deletion of any one item resulted in marginally lower Cronbach's alpha values (αs = .873–.881). These findings suggest that each item contributes meaningfully to the overall construct measured by the GEGS, supporting its reliability. Further, Time 1 and Time 2 GEGS scores showed a large, significant correlation (r = .861, p < .01; ICC = 0.924, p < .001) and Time 1 scores predicted GEGS scores one week later, $F(199) = 569.303$, p < .001, $R^2 = .742$, indicating high test-retest reliability (see Table 3).

[TABLE 3 NEAR HERE]

Validity

Convergent and Discriminant Validity

A higher GEGS score was strongly associated with higher general gender essentialism as measured by the GES (r = .667, p < .01) with a strong effect, supporting convergent validity. With regards to discriminant validity, our hypothesis was partially supported as we found relatively strong correlations between the GEGS and sexual prejudice in sport as

measured by the SPSS ($r = .527$, $p < .01$), particularly with the open rejection ($r = .644$, $p <$ .01) and gendering performance subscales ($r = .674$, $p < .01$). Finally, we found moderate correlations between the GEGS and social dominance orientation as measured by the SDOS ($r = .420$, $p < .01$) (see Table 3). All correlations remained stable at Time 2 and were consistent across coach gender.

Differentiation by Known Groups

GEGS scores were significantly higher among male coaches ($M = 2.64$, $SD = 0.55$), compared to female coaches ($M = 2.29$, $SD = 0.65$), $t(253) = 4.622$, $p < .001$, $d = .60$. Further, coaches who had received previous training on gender and/or gender stereotypes reported lower GEGS scores ($M = 2.36$, $SD = 0.64$), compared to coaches with no previous training ($M = 2.55$, $SD = 0.64$), $t(228) = -2.085$, $p = .019$, $d = .64$. Coach age did not predict total GEGS score, $F(258) = 2.148$, $p = .077$, $R^2 = .012$.

Discussion

The current study describes the development and validation of the Gender Essentialist Beliefs Towards Girls in Sport Scale (GEGS), which assesses levels of gender essentialism among sport coaches. The GEGS provides a novel contribution to the literature on gender essentialism in sport, which has highlighted the prevalence of gender essentialism as a structuring element of sport. Further, the GEGS complements existing measures of gender essentialism outside of sport contexts and extends measurement within the context of sport.[3,6,11,20] Sport coaching, practices, discourses, and methodologies are shaped by gender essentialist perspectives and this new measure allows for coaches and sport stakeholders to identify areas in which gender essentialist beliefs are most prevalent.[7,10]

The GEGS was developed based on recent guidelines for scale development and validation.[34,35] The 20-item scale showed good reliability (internal consistency, test-retest reliability, item-total correlations), validity (convergent validity, differentiation between

known groups), and factor loadings. Convergent validity was established by correlating the GEGS with general gender essentialism, as measured by the GES.[12] A higher score on the GEGS was related to higher general gender essentialism in the current sample. This finding provides preliminary support for the ability of the GEGS to tap into the construct of gender essentialism; however, the GEGS is unique in that it captures essentialism within the sport context, unlike the GES. Furthermore, the association between gender essentialism in sport and general gender essentialism suggests that coaches who have gender essentialist tendencies in general are also more likely to bring such attitudes to a sport setting.

We also found partial support for discriminant validity, as the GEGS showed relatively strong positive correlations with sexual prejudice, as measured via the SPSS.[41] Although discriminant validity is usually evidenced by low correlations between scores[34], moderate to strong correlations (e.g., .40 or .50) can be more meaningful when they show the scale can effectively differentiate closely related traits.[35] Therefore, these findings need to be interpreted in light of several considerations. First, no gold-standard measure of discriminant validity currently exists for the GEGS. The SPSS was chosen to explore whether attitudes towards gender and gender identity (as measured by the GEGS) would differ from attitudes towards sexuality (as measured by the SPSS), which are distinct constructs. However, prior research has suggested that prejudice towards sexuality may be associated with gender essentialist beliefs.[55] Second, it is possible that participants conflated gender identity and sexuality. Previous studies exploring differences between these constructs have highlighted that individuals, such as coaches, often confuse and/or equate the concepts of gender, sex, and sexuality.[33] This could partially explain the moderately strong correlations found between the GEGS and the SPSS in the current study. Overall, our findings suggest that coaches who endorsed gender essentialism were also more likely to endorse prejudiced attitudes towards lesbian and gay athletes.

No hypotheses were made comparing GEGS scores to the SDOS, but the results showed that higher scores on the GEGS were moderately correlated with higher scores on the SDOS. In other words, participants who endorsed greater gender essentialist beliefs also reported stronger beliefs in social hierarchies and preference for social inequality. Notably, the GEGS is a unique construct compared to SDOS as the GEGS attempts to measure gender essentialism specifically, while the SDOS measures beliefs of social group dominance and inequalities.

Observed differences between groups supported the validity of the GEGS, exhibiting expected outcomes. Specifically, female coaches and coaches who indicated that they had previously received training on gender and/or gender stereotypes scored significantly lower on the GEGS, compared to male coaches and coaches with no previous training. The finding of female coaches scoring lower on the GEGS is similar to other studies related to assessing gender essentialism, where women were found to show lower levels of essentialism compared to men.[12,21] Qualitative work on gender essentialism in sport has also found how sport stakeholders, including athletes and coaches, can exert agency to challenge and resist essentialist ideologies.[2,7,56] Specifically, research on women coaches' experiences has highlighted how personal experiences of discrimination and marginalization in sport forces women to have to "prove" themselves within a masculinized sports environment and directly challenge gender stereotypes and gender limitations caused by essentialist sport structures.[56,57] Our findings further align with previous research where coaches who engaged with coach education programs related to gender equity exhibited lower levels of gender essentialism.[9,26] However, it is important to note that many coach education programs related to gender equity may inadvertently replicate gender norms, gender essentialism, and stereotypes or protect masculine hegemony in sport.[58,59] Therefore, coach education programs must intentionally challenge essentialist beliefs and take a gender responsive approach to

coaching in order to be most effective in decreasing levels of gender stereotypes in sport coaching, which can foster environments where girls feel safe, valued, and affirmed.[9,10,26,29] Notably, no significant effect was found for age, suggesting that gender essentialism can manifest in coaches across all age groups.

A notable contribution of the GEGS lies in its addition of items related to the inclusion of transgender athletes in sport and its links with gender essentialist beliefs. The inclusion of transgender athletes in sport has been a hotly contested and politicized topic in the last decade.[5,60] Research on coach perceptions of transgender athlete inclusion has found that coaches exhibit a variety of opinions, ranging from staunchly anti-trans and invalidating of trans identities, to uncertain, to firmly inclusive of trans athletes.[32] Furthermore, coaches are less likely to support the sport participation of transgender women compared to transgender men, reflecting a 'protect women's sport' ethos rooted in essentialist stereotypes of women's athletic abilities.[5,32] Although transphobia and essentialism may be distinct concepts, research suggests there are overlaps between transphobia, cisgenderism, and essentialism, as essentialism would uphold a belief in an immutable gender binary.[4,60] Essentialism in sport contexts invalidates transgender identities and leads to discrimination, backlash, and abuse of transgender and gender diverse individuals.[4,5]

The inclusion of transgender items in the GEGS is justified by the theoretical connections between cisgenderism and essentialism, as well as the harm that essentialist beliefs of trans athletes pose for upholding similarly essentialist stereotypes of cisgender girls and women. Protectionist notions against transgender women replicate essentialist beliefs that athleticism lies on an immutable gender binary and cisgender girls are inferior in sport.[5] In the present study, the two items relating to transgender athletes (i.e., "Transgender girls should not play on girls' sports teams", "Transgender boys should not play on boys' sports teams") showed strong correlations with the total GEGS score, indicating that higher levels of

essentialism are positively associated with greater anti-trans attitudes in sport. Notably, although a thorough qualitative analysis was not conducted, in response to the optional, qualitative, open-ended question "Do you have any other thoughts about this survey that you would like to share with us?", many participants specifically shared their opinions on transgender athlete inclusion. Some responses asserted trans-inclusive opinions, such as "I feel that transgender athletes should be included in sport…At the end of the day, all kids deserve the right to participate regardless of any of these matters" (Participant #260). Others expressed essentialist tones that invalidated trans identities and replicated essentialist stereotypes, such as "If biological males who become females are allowed to compete against biological females it will destroy all progress made in women's sports" (Participant #245). The wide range of opinions expressed voluntarily by participants in an optional open-ended question reflects prior research, highlights how more research is needed to substantiate this association of transphobia, cisgenderism, and essentialism in the sport context, and illustrates the need to develop a new scale that explicitly assesses attitudes towards transgender athletes in sport.[32]

Overall, our scale provides a novel measure to assess the prevalence of gender essentialism among sport coaches and sport practitioners that can be applied to a variety of sport levels and contexts. A scale assessing gender essentialism in sport is helpful to address and remedy coaching opinions and methodologies that might have deleterious effects on athletes' sports experiences, retention, and enjoyment.[9,26] Additionally, the GEGS can help inform and evaluate coach education programs aimed to support coaches, practitioners, and administrators in identifying gaps in the understanding of socio-cultural topics, such as gender equity and gender bias. Improving awareness in areas related to social equity and gender stereotypes can be beneficial in strengthening coach-athlete relationships, retaining

athletes in sport, improving athlete development and well-being, and challenging masculine norms in sports environments.[29,30]

Strengths, Limitations, and Future Directions

The findings of the present study need to be interpreted in light of several strengths and limitations. In terms of strengths, we meticulously followed stringent guidelines for the development and validation of our scale, ensuring its robustness and reliability.[34,35] Additionally, our inclusion of a substantial sample size, in line with established recommendations, bolsters the statistical power of our analyses. Furthermore, the assessment of test-retest reliability affirms the stability of our scale over time.

While our study provides a novel measure to assess gender essentialism towards girls in sport, several limitations warrant acknowledgment. Foremost among these is the absence of an appropriate measure of discriminant validity, which we recognize as a crucial aspect in scale validation. For the purposes of the current study, we proposed the SPSS as a measure of discriminant validity, given inherent differences in the constructs of sexuality and gender. However, our findings showed strong correlations between the GEGS and sexual prejudice in sport as measured by the SPSS. It is plausible that individuals who report higher gender essentialist beliefs are also more likely to exhibit greater levels of sexual prejudice. This is particularly evident given the high correlations between the GEGS and the gendered performance subscale of the SPSS. Second, although two items assessing attitudes towards transgender athletes were included in the GEGS based on expert feedback in the initial phases of scale development, further research is needed to refine and validate these additions. Given how antipathy towards transgender athletes is due to varying types of gender essentialisms, cisgenderism, and transphobia, we acknowledge that two items are unlikely to fully capture the nuance in attitudes towards transgender athletes, particularly in girls' sport. Novel scales are needed to explicitly measure transphobia and trans inclusion in sport settings. Third, our

sample is predominantly composed of White coaches based in the US and includes no trans women or men, which limits the generalizability of our findings to broader populations. Relatedly, given convenience sampling was used to recruit participants, bias may be introduced as respondents self-select to participate. This can potentially exclude participants who have strong gender essentialist beliefs or who are not interested in the topic. Given participants who did not complete the GEGS were removed from analysis, the sample may also be vulnerable to non-response bias. Fourth, this study focuses on the construction and validation of the GEGS but does not demonstrate predictive validity with specific outcomes associated with gender essentialism in sport. We suggest there is opportunity for future research to illustrate the utility of the GEGS in predicting various outcomes in sport (e.g., athlete performance, the coach-athlete relationship).

Moving forward, there are several avenues for future research that stem from the insights gained in this study. Continued testing and validation efforts will further bolster the reliability and validity of our scale, enhancing its utility within the field. As noted above, the GEGS should be tested for predictive validity with associated outcomes related to gender essentialism. For instance, future research could explore if coaches who score high on the GEGS also report more antipathy toward professional women's sport or transgender athlete inclusion, or whether the teams of coaches who score high on the GEGS report less satisfaction with their sports experiences. Further, exploring the invariance of our scale across diverse demographic groups (e.g., coach gender) will shed light on potential variations in gender essentialist attitudes and perceptions. Additionally, future efforts should aim to adapt and validate the GEGS to assess gender essentialism towards different athlete populations (e.g., adult and male athletes). Finally, pre- and post-intervention scores were not assessed in this study, which would determine if the GEGS is sensitive enough to detect changes in gender essentialist beliefs after education or an intervention targeted at coaches. Future

research should assess GEGS scores before and after an intervention, to not only evaluate the effectiveness of the intervention, but to determine the pre- and post-test (predictive) validity of the GEGS.

Conclusions

The Gender Essentialist Beliefs Towards Girls in Sport Scale (GEGS) developed in this study provides a valuable tool for assessing gender essentialism among sport coaches. Our findings support the reliability and validity of the GEGS, demonstrating its potential to identify areas where gender essentialist beliefs are prevalent among coaches. Significant differences between groups supported the validity of the GEGS, with female coaches and those with previous gender training scoring lower on the scale. This underscores the importance of coach education programs in challenging essentialist beliefs. The inclusion of items related to transgender athlete inclusion in the GEGS revealed associations between gender essentialism and anti-trans attitudes. However, further research is needed to validate this connection. Overall, the GEGS is a valuable tool that can be utilized by sport coaches, organizations, administrators, and other related stakeholders to help address and alleviate the ways in which gender essentialism negatively impacts girls' sport experiences. By identifying gender essentialist beliefs in sport, the GEGS can help document and educate coaches and sport administrators on existing biases, which can lead to practical solutions, such as coach education initiatives, that can help challenge stereotypes, improve experiences, and retain girls in sport.[9,25] Future research should continue to refine and validate the scale and explore its sensitivity to changes in gender essentialist beliefs over time.

References

1. Nicole M. LaVoi. Introduction to the Tucker Center research report. In: Nicole M. LaVoi, editor. *Developing physically active girls: An evidence-based multidisciplinary approach*. Minneapolis, MN: The Tucker Center for Research on Girls & Women in Sport; 2018. p. xvii–xxiv.

2. Allison R. Privileging difference: Negotiating gender essentialism in U.S. women's professional soccer. *Sociology of Sport Journal* 2021 Jun 1;38(2):158–66.

3. Kane MJ. Resistance/transformation of the oppositional binary: Exposing sport as a continuum. *Journal of Sport and Social Issues* 1995 May;19(2):191–218.

4. Afroozeh MS, Phipps C, Afrouzeh A, Mehri A, Alipour Asiri Z. "The spectators ask, is it a boy or a girl? What is it?": Cultural cisgenderism and trans men's sporting experiences in Iran. *International Review for the Sociology of Sport* 2023 Dec;58(8):1221–40.

5. Greey AD, Lenskyj H, editors. *Justice for trans athletes: challenges and struggles*. 1st edition. United Kingdom: Emerald Publishing; 2023.

6. Allison R. *Kicking center: gender and the selling of women's professional soccer.* New Brunswick, New Jersey: Rutgers University Press; 2018.

7. Messner M. Gender ideologies, youth sports, and the production of soft essentialism. *Sociology of Sport Journal* 2011 Jun;28(2):151–70.

8. Women in Sport. *Sport, Stereotypes, and Stolen Dreams: Why Girls Still Feel Like They Don't Belong in Sport*. Women in Sport; 2023. Available from: https://womeninsport.org/resource/sport-stereotypes-and-stolen-dreams/

9. Goorevich A, Boucher C, Schneider J, Silva-Breen H, Matheson EL, Tinoco A, et al. Acceptability and preliminary efficacy testing of a web-based coach development program addressing gender essentialism among coaches of adolescent girls. *International Sport Coaching Journal* 2023;1–13.

10. Goorevich A, LaVoi NM. Essentially different or equally the same: Uncovering sport coach discourses about coaching girls. *Sports Coaching Review* 2024 Jan 29;1–26.

11. Greene S. Biological Determinism and Essentialism. In: Naples NA, editor. Companion to Feminist Studies [Internet]. 1st ed. Wiley; 2020 [cited 2024 Sep 19]. p. 13–34. Available from: https://onlinelibrary.wiley.com/doi/10.1002/9781119314967.ch2

12. Skewes L, Fine C, Haslam N. Beyond Mars and Venus: The role of gender essentialism in support for gender inequality and backlash. Lozano S, editor. *PLoS ONE* 2018 Jul 24;13(7):e0200921.

13. Lorber J. Believing is seeing: Biology as ideology. *Gender & Society* 1993 Dec;7(4):568–81.

14. Hyde JS. The gender similarities hypothesis. *American Psychologist* 2005 Sep;60(6):581–92.

15. Allison R, Love A. "We All Play Pretty Much the Same, Except. . .": Gender-Integrated Quidditch and the Persistence of Essentialist Ideology. *Journal of Contemporary Ethnography* 2022 Jun;51(3):347–75.

16. Love A, Kelly K. Equity or Essentialism?: U.S. Courts and the Legitimation of Girls' Teams in High School Sport. *Gender & Society* 2011 Apr;25(2):227–49.

17. Jakubowska H, Byczkowska-Owczarek D. Girls in Football, Boys in Dance. Stereotypization Processes in Socialization of Young Sportsmen and Sportswomen. *QSR* 2018 Aug 28;14(2):12–28.

18. Plaza M, Boiché J, Brunel L, Ruchaud F. Sport = Male… But Not All Sports: Investigating the Gender Stereotypes of Sport Activities at the Explicit and Implicit Levels. *Sex Roles* 2017 Feb;76(3–4):202–17.

19. Felton L, Jowett S. "What do coaches do" and "how do they relate": Their effects on athletes' psychological needs and functioning. *Scandinavian Med Sci Sports* 2013;23(2).

20. Sue DW, Capodilupo CM, Torino GC, Bucceri JM, Holder AMB, Nadal KL, et al. Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist* 2007;62(4):271–86.

21. Lee KY, Reis HT, Rogge RD. Seeing the World in Pink and Blue: Developing and Exploring a New Measure of Essentialistic Thinking about Gender. *Sex Roles* 2020;83(11–12):685–705.

22. Adams A. Humour, masculinities and youth sport coaching: 'Good morning, ladies!' *Sport, Education and Society* 2020;25(4):463–74.

23. Ferguson HL, Swann C, Liddle SK, Vella SA. Investigating Youth Sports Coaches' Perceptions of Their Role in Adolescent Mental Health. *Journal of Applied Sport Psychology* 2019;31(2):235–52.

24. Langan E, Blake C, Lonsdale C. Systematic review of the effectiveness of interpersonal coach education interventions on athlete outcomes. *Psychology of Sport and Exercise* 2013 Jan;14(1):37–49.

25. Newman TJ, Lower-Hoppe LM, Burch M, Paluta LM. Advancing positive youth development-focused coach education: contextual factors of youth sport and youth sport leader perceptions. *Managing Sport and Leisure* 2021 Jul 4;26(4):326–40.

26. Schneider J, Matheson EL, Tinoco A, Gentili C, White P, Boucher C, et al. Body confident coaching: a pilot randomized controlled trial evaluating the acceptability of a web-based body image intervention for coaches of adolescent girls. *Journal of Applied Sport Psychology* 2023;1–26.

27. de Haan D, Knoppers A. Gendered discourses in coaching high-performance sport. *International Review for the Sociology of Sport* 2020 Sep;55(6):631–46.

28. Gosai J, Jowett S, Rhind DJA. Coaching Through a "Gender Lens" May Reveal Myths That Hinder Female Athletes: A Multistudy Investigation. *International Sport Coaching Journal* 2022 May 1;9(2):222–33.

29. Norman L. Is There a Need for Coaches to Be More Gender Responsive? A Review of the Evidence. *International Sport Coaching Journal* 2016 May;3(2):192–6.

30. Jones L, Mills J, Avner Z. Learning to problematise 'the way things are' when coaching female athletes. In: Cope E, Partington M, editors. *Sports Coaching*. 1st ed. Routledge; 2019. p. 135–45.

31. Bolter ND, Lucas SM. Female and male youth sport coaches' perceptions of teaching sportspersonship to female and male athletes. *International Journal of Sports Science & Coaching* 2018 Oct;13(5):619–28.

32. Arnold ML, Chambers K, Petrie TA. NCAA coaches' perceptions of the inclusion of transgender athletes: a qualitative analysis. *Sport, Education and Society* 2023 Nov 23;1–17.

33. Scheadler TR, Anderson-Butcher D, Bates S. College coaches' engagement in advocacy for lesbian, gay, bisexual, transgender, queer, intersex, asexual, and other gender and sexual minorities (LGBTQIA+) rights and racial justice: Experiences, facilitators, and barriers. *Journal of Diversity in Higher Education* 2024 Feb 8; Available from: https://doi.apa.org/doi/10.1037/dhe0000559

34. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health* 2018 Jun 11;6:149.

35. Vitoratou S, Uglik-Marucha N, Hayes C, Pickles A. An introductory comprehensive guide for assessing measurement tool quality: The Contemporary Psychometrics (ConPsy) Checklist [Internet]. 2023 [cited 2024 Apr 29]. Available from: https://osf.io/t2pbj

36. R. Tourangeau. Cognitive Sciences and Survey Methods. In: T. B. Jabine, M. L. Straf, J. M. Tanur, R. Tourangeau, editors. *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington D.C.: National Academy Press; 1984. p. 73–100.

37. Willis GB. Analysis of the cognitive interview in questionnaire design. Oxford: Oxford university press; 2015. (Understanding qualitative research).

38. Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. In: Kazdin AE, editor. Methodological issues and strategies in clinical research (4th ed) [Internet]. Washington: American Psychological Association; 2016 [cited 2024 May 1]. p. 187–203. Available from: https://content.apa.org/books/14805-012

39. Comrey AL. Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology* 1988;56(5):754–61.

40. Schudson ZC, Van Anders SM. Gender/sex diversity beliefs: Scale construction, validation, and links to prejudice. Group Processes & Intergroup Relations 2022 Jun;25(4):1011–36.

41. Baiocco R, Pistella J, Salvati M, Ioverno S, Lucidi F. Sexual Prejudice in Sport Scale: A New Measure. *Journal of Homosexuality*. 2020 Mar 20;67(4):489–512.

42. Pratto F, Sidanius J, Stallworth LM, Malle BF. Social dominance orientation: A personality variable predicting social and political attitudes. Journal of Personality and Social Psychology 1994 Oct;67(4):741–63.

43. Aichholzer J, Lechner CM. Refining the Short Social Dominance Orientation scale (SSDO): A validation in seven European countries. *J Soc Polit Psych* 2021 Sep 16;9(2):475–89.

44. Wood ND, Akloubou Gnonhosou DC, Bowling JW. Combining Parallel and Exploratory Factor Analysis in Identifying Relationship Scales in Secondary Data. Marriage & Family Review. 2015 Jul 4;51(5):385–95.

45. O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behavior Research Methods, Instruments, & Computers. 2000 Sep;32(3):396–402.

46. Hendrickson AE, White PO. PROMAX: A QUICK METHOD FOR ROTATION TO OBLIQUE SIMPLE STRUCTURE. *British Journal of Statistical Psychology* 1964 May;17(1):65–70.

47. Yeomans KA, Golder PA. The Guttman-Kaiser Criterion as a Predictor of the Number of Common Factors. *The Statistician* 1982 Sep;31(3):221.

48. Bollen KA. *Structural equations with latent variables*. New York: John Wiley & Sons; 1989.

49. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 1999 Jan;6(1):1–55.

50. Kline RB. *Principles and practice of structural equation modeling*. Fifth edition. New York London: The Guilford Press; 2015. 494 p. (Methodology in the social sciences).

51. Cohen J. A power primer. *Psychological Bulletin* 1992;112(1):155–9.

52. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin. 1979;86(2):420–8.

53. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Routledge; 2013.

54. Tabachnick BG, Fidell LS. *Using multivariate statistics*. Seventh edition. New York, NY: Pearson; 2019.

55. Vanessa Hettinger. Reconceptualizing the role of essentialism in attitudes toward gays and lesbians: The intersection of gender and sexual orientation. [Internet]. University of South Florida; 2014. Available from: https://core.ac.uk/reader/154470560

56. Norman L. Bearing the Burden of Doubt: Female Coaches' Experiences of Gender Relations. *Research Quarterly for Exercise and Sport* 2010 Dec;81(4):506–17.

57. Norman L, Rankin-Wright A. Surviving rather than thriving: Understanding the experiences of women coaches using a theory of gendered social well-being. *International Review for the Sociology of Sport* 2018 Jun;53(4):424–50.

58. Allen KR, Kaestle CE, Goldberg AE. More Than Just a Punctuation Mark: How Boys and Young Men Learn About Menstruation. *Journal of Family Issues* 2011 Feb;32(2):129-56.

59. Norman L. The impact of an "equal opportunities" ideological framework on coaches' knowledge and practice. *International Review for the Sociology of Sport* 2016 Dec;51(8):975–1004.

60. Gleaves J, Lehrbach T. Beyond fairness: the ethics of inclusion for transgender and intersex athletes. *Journal of the Philosophy of Sport* 2016 May 3;43(2):311–26.

61. Diehl AB, Stephenson AL, Dzubinski LM, Wang DC. Measuring the invisible: Development and multi-industry validation of the Gender Bias Scale for Women Leaders. *Human Resource Dev Quarterly* 2020 Sep;31(3):249–80.

62. Haslam N, Rothschild L, Ernst D. Essentialist beliefs about social categories. *British J Social Psychol* 2000 Mar;39(1):113–27.

63. Tezare, K. Gender Biaz Quiz [Internet]. Commonwealth of Learning; 2015. Available from: http://hdl.handle.net/11599/1740

64. Kerr, P. S., Holden, R. R. Development of the Gender Role Beliefs Scale (GRBS). *Journal of Social Behavior and Personality* 1996;11(5):3–16.

65. Nurlu Ö. Developing a Teachers Gender Stereotype Scale toward Mathematics. *International Electronic Journal of Elementary Education* 2017 Dec 30;10(2):287–99.

66. Voulgaridou I, Kokkinos CM. The Relational aggression scale (RAS): Psychometric properties of a newly developed measure of relational aggression. *European Review of Applied Psychology*. 2018 Jan;68(1):11–22.

67. Crick NR. The Role of Overt Aggression, Relational Aggression, and Prosocial Behavior in the Prediction of Children's Future Social Adjustment. *Child Development* 1996 Oct;67(5):2317.

68. Hill DB, Willoughby BLB. The Development and Validation of the Genderism and Transphobia Scale. *Sex Roles* 2005 Oct;53(7–8):531–44.

69. Billard TJ. Attitudes Toward Transgender Men and Women: Development and Validation of a New Measure. *Front Psychol* 2018 Apr 3;9:387.