

Promotions:

Q1. Evaluate each item content of BSS using LLM

Q2. Evaluate the dimensionality of BSS using LLM

Q3. Evaluate whether the BSS items function differently across gender using LLM

Claude opus 4

Q1: Evaluate each item content of BSS using LLM

I'll evaluate each item of the Belief in Sexism Shift Scale (BSS) from multiple perspectives, considering their psychometric properties, construct validity, and potential issues.

Item-by-Item Evaluation

Item 1: "In the US, discrimination against men is on the rise."

- **Strengths:** Clear temporal comparison, specific geographic context
- **Weaknesses:** Vague term "discrimination" without specific domains; may capture general perceptions rather than specific beliefs
- **Potential bias:** May activate defensive responses depending on respondent's ideology

Item 2: "Men are not particularly discriminated against." (Reverse-coded)

- **Strengths:** Good reverse-coding item to check response patterns
- **Weaknesses:** "Particularly" is ambiguous - could mean frequency or severity
- **Measurement issue:** Double negative when reverse-coded may confuse interpretation

Item 3: "If anything, men are more discriminated against than women these days."

- **Strengths:** Direct comparative statement, clear position
- **Weaknesses:** Temporal marker "these days" is vague
- **Validity concern:** Assumes discrimination is zero-sum, which may not reflect complex reality

Item 4: "Giving women more rights often requires taking away men's rights."

- **Strengths:** Captures zero-sum belief about gender rights
- **Weaknesses:** "Rights" undefined - legal, social, economic?
- **Conceptual issue:** Conflates rights with privileges or advantages

Item 5: "Under the guise of equality for women, men are actually being discriminated against."

- **Strengths:** Specific claim about hidden discrimination
- **Weaknesses:** Leading language ("guise") may bias responses
- **Framing issue:** Implies deception rather than unintended consequences

Item 6: "In the pursuit of women's rights, the government has neglected men's rights."

- **Strengths:** Specific actor (government) and domain
- **Weaknesses:** Assumes government is primary actor in rights
- **Geographic limitation:** Government structure varies by location

Item 7: "Nowadays, men don't have the same chances in the job market as women."

- **Strengths:** Specific domain (employment), testable claim
- **Weaknesses:** Ignores field-specific variations
- **Empirical issue:** Contradicts most labor statistics

Item 8: "Feminism is about favoring women over men."

- **Strengths:** Taps into perceptions of feminist movement
- **Weaknesses:** Treats feminism as monolithic
- **Validity issue:** Measures perception of feminism, not necessarily sexism beliefs

Item 9: "Feminism does not discriminate against men." (Reverse-coded)

- **Strengths:** Good pairing with Item 8 for consistency check
- **Weaknesses:** Anthropomorphizes feminism as capable of discrimination
- **Logical issue:** Movements vs. individual actors

Item 10: "All in all, men have more responsibilities and fewer benefits."

- **Strengths:** Broad summary statement
- **Weaknesses:** Extremely vague - which responsibilities/benefits?
- **Comparison issue:** No baseline for comparison

Item 11: "In today's society, women can say things that men are not allowed to say."

- **Strengths:** Specific behavioral claim
- **Weaknesses:** "Allowed" is ambiguous - legally, socially?
- **Context missing:** What types of statements?

Item 12: "It is evident that the media is biased against men."

- **Strengths:** Specific institution (media)
- **Weaknesses:** "Evident" assumes objectivity of subjective perception
- **Generalization:** Treats all media as uniform

Item 13: "In today's society, men are often punished for acting manly."

- **Strengths:** Captures masculine norm violations
- **Weaknesses:** "Manly" undefined and culturally variable
- **Ambiguity:** "Punished" could mean many things

Item 14: "All in all, men are well respected in today's society." (Reverse-coded)

- **Strengths:** General status assessment
- **Weaknesses:** May not correlate with other items if respect \neq discrimination
- **Measurement issue:** Respect is multidimensional

Item 15: "While women can use the 'gender-card' to get ahead, men can't."

- **Strengths:** Specific mechanism claim
- **Weaknesses:** Loaded term "gender-card" with negative connotations
- **Asymmetry:** Ignores potential male advantages

Overall Scale Assessment

Strengths:

- Covers multiple domains (legal, social, economic, media)
- Includes reverse-coded items for response bias detection
- Taps into contemporary gender discourse

Weaknesses:

- Heavy use of comparative and zero-sum framing
- Many items use vague or loaded language
- Lacks items about specific, measurable phenomena
- May conflate different constructs (anti-feminism, male victimization, traditional gender roles)

Recommendations:

1. Define key terms more precisely
2. Add items about specific, observable phenomena
3. Balance items to avoid seeming biased
4. Consider subscales for different domains
5. Validate against behavioral measures
6. Test for measurement invariance across gender groups

The scale appears designed to measure beliefs about reverse sexism or male disadvantage, but may benefit from refinement to improve construct validity and reduce potential response bias.

Q2: Evaluate the dimensionality of BSS using LLM

Dimensionality Evaluation of the Belief in Sexism Shift Scale (BSS)

Theoretical Dimensionality Analysis

Based on content analysis of the 15 items, I identify several potential underlying dimensions:

1. Male Discrimination/Victimization (Primary dimension)

Items clustering on this dimension:

- Item 1: "discrimination against men is on the rise"
- Item 2: "Men are not particularly discriminated against" (R)
- Item 3: "men are more discriminated against than women"
- Item 5: "men are actually being discriminated against"
- Item 10: "men have more responsibilities and fewer benefits"

2. Zero-Sum Gender Rights Beliefs

Items clustering on this dimension:

- Item 4: "women's rights requires taking away men's rights"
- Item 6: "government has neglected men's rights"
- Item 15: "women can use the 'gender-card'"

3. Anti-Feminism Attitudes

Items clustering on this dimension:

- Item 8: "Feminism is about favoring women"
- Item 9: "Feminism does not discriminate against men" (R)

4. Cultural/Social Disadvantage

Items clustering on this dimension:

- Item 7: "men don't have same chances in job market"
- Item 11: "women can say things men can't"
- Item 12: "media is biased against men"
- Item 13: "men are punished for acting manly"
- Item 14: "men are well respected" (R)

Predicted Factor Structure

Most Likely: Two-Factor Model

Factor 1: General Male Disadvantage Beliefs

- Items 1, 2(R), 3, 5, 7, 10, 11, 12, 13, 14(R), 15
- Captures broad beliefs about men facing discrimination

Factor 2: Ideological Opposition

- Items 4, 6, 8, 9(R)
- Captures beliefs about feminism and gender rights as zero-sum

Alternative: Three-Factor Model

Factor 1: Discrimination Beliefs

- Items 1, 2(R), 3, 5

Factor 2: Systemic Disadvantage

- Items 6, 7, 10, 11, 12, 13, 14(R), 15

Factor 3: Anti-Feminism

- Items 8, 9(R), possibly 4

Potential Dimensionality Issues

1. Conceptual Overlap

Many items tap into similar constructs with slightly different wording:

- Items 1, 3, and 5 all address increasing discrimination
- Items 8 and 9 are nearly redundant about feminism

2. Cross-Loading Candidates

Several items likely load on multiple factors:

- Item 4 (zero-sum rights) could load on both ideological and discrimination factors
- Item 6 (government neglect) bridges systemic and ideological dimensions
- Item 15 ("gender-card") combines discrimination and zero-sum beliefs

3. Weak Reverse-Coded Items

- Item 14 ("well respected") may not correlate strongly with discrimination items
- Item 2 uses ambiguous language ("particularly") that may weaken its relationship

Empirical Predictions

Expected Factor Analysis Results:

1. **Eigenvalues:** First factor likely explains 40-50% of variance
2. **Scree plot:** Probable elbow after 1-2 factors
3. **Factor correlations:** If multidimensional, factors likely correlated .60-.80

Expected Reliability:

- **Total scale:** $\alpha = .85-.92$ (high due to content similarity)
- **Subscales:** $\alpha = .75-.85$ (if separated)

Model Fit Predictions:

- **Unidimensional model:** Acceptable but not optimal (CFI $\sim .90$, RMSEA $\sim .08$)
- **Two-factor model:** Better fit (CFI $\sim .95$, RMSEA $\sim .06$)
- **Bifactor model:** Might fit best with general factor + specific factors

Recommendations for Dimensionality Assessment

1. **Conduct Parallel Analysis** to determine optimal number of factors
2. **Test measurement invariance** across gender groups - dimensionality may differ
3. **Consider bifactor modeling** given strong general factor likely
4. **Examine item response patterns** - some items may show floor/ceiling effects
5. **Test alternative models:**
 - Unidimensional

- Correlated factors (2-3 factors)
- Bifactor
- Higher-order factor model

Conclusion

The BSS likely exhibits a **strong general factor** with possible specific factors. While conceptually multidimensional, empirically it may function as essentially unidimensional due to high interfactor correlations. The scale would benefit from:

- More distinct item content to capture separate dimensions
- Stronger reverse-coded items
- Items that differentiate between related but distinct constructs
- Validation of dimensional structure across different populations

Q3: Evaluate whether the BSS items function differently across gender using LLM

Differential Item Functioning (DIF) Analysis of BSS Across Gender

Overview of Expected Gender-Based DIF

The BSS items are likely to show substantial differential functioning across gender groups due to the inherently gendered nature of the content. Here's a detailed analysis of expected DIF patterns:

Item-Level DIF Predictions

High DIF Expected

Item 1: "In the US, discrimination against men is on the rise"

- **Expected Pattern:** Men more likely to endorse at same latent trait level
- **Type:** Uniform DIF
- **Mechanism:** Direct personal relevance and lived experience differences

Item 3: "If anything, men are more discriminated against than women these days"

- **Expected Pattern:** Strong uniform DIF favoring male endorsement
- **Magnitude:** Large (>0.5 logits difference)
- **Reason:** Fundamentally contradicts many women's experiences

Item 7: "Nowadays, men don't have the same chances in the job market as women"

- **Expected Pattern:** Men higher endorsement
- **Type:** Uniform DIF
- **Note:** Contradicts employment statistics showing male advantages in most sectors

Item 13: "In today's society, men are often punished for acting manly"

- **Expected Pattern:** Complex DIF - depends on interpretation of "manly"
- **Type:** Potentially non-uniform
- **Issue:** Gender groups may interpret "manly" behavior differently

Item 15: "While women can use the 'gender-card' to get ahead, men can't"

- **Expected Pattern:** Strong DIF favoring male endorsement
- **Type:** Uniform DIF
- **Mechanism:** Term "gender-card" likely perceived as more legitimate by men

Moderate DIF Expected

Item 4: "Giving women more rights often requires taking away men's rights"

- **Expected Pattern:** Men more likely to endorse zero-sum framing
- **Type:** Uniform DIF
- **Magnitude:** Moderate (0.3-0.5 logits)

Item 8: "Feminism is about favoring women over men"

- **Expected Pattern:** Gender gap in endorsement
- **Type:** Potentially non-uniform
- **Complexity:** May depend on respondent's exposure to feminist discourse

Item 11: "In today's society, women can say things that men are not allowed to say"

- **Expected Pattern:** Men perceive more speech restrictions
- **Type:** Uniform DIF
- **Context:** Different awareness of gendered communication norms

Item 12: "It is evident that the media is biased against men"

- **Expected Pattern:** Men more likely to perceive bias
- **Type:** Uniform DIF
- **Factor:** Different media consumption patterns by gender

Low to Moderate DIF Expected

Item 2: "Men are not particularly discriminated against" (R)

- **Expected Pattern:** Women more likely to agree (before reversal)
- **Issue:** "Particularly" ambiguity may function differently

Item 5: "Under the guise of equality for women, men are actually being discriminated against"

- **Expected Pattern:** Moderate DIF
- **Complexity:** "Guise" implies deception, triggering different responses

Item 6: "In the pursuit of women's rights, the government has neglected men's rights"

- **Expected Pattern:** Moderate uniform DIF
- **Factor:** Political orientation may moderate gender differences

Item 10: "All in all, men have more responsibilities and fewer benefits"

- **Expected Pattern:** Complex pattern
- **Issue:** "Responsibilities" and "benefits" interpreted differently by gender

Minimal DIF Expected

Item 9: "Feminism does not discriminate against men" (R)

- **Expected Pattern:** Similar to Item 8 but reverse-coded
- **Note:** May show floor effects for both groups

Item 14: "All in all, men are well respected in today's society" (R)

- **Expected Pattern:** Smallest gender differences
- **Reason:** "Respect" is broad enough to allow agreement across groups

Types of DIF Expected

1. Uniform DIF (Most Common)

- Consistent gender differences across trait levels
- Men systematically more likely to endorse most items
- Largest for items about discrimination and victimization

2. Non-Uniform DIF (Selected Items)

- Gender differences vary by trait level
- Expected for items about feminism and "manliness"
- May show crossover points where patterns reverse

3. Mixed DIF

- Both location and discrimination parameters differ
- Likely for complex items (e.g., Item 13 about "acting manly")

Measurement Implications

Scale-Level Impact

1. **Total scores not comparable** across gender without adjustment
2. **Different factor structures** possible for men and women
3. **Reliability differences:** Scale may be more internally consistent for men

Validity Threats

1. **Construct bias:** Scale may measure different constructs by gender
2. **Predictive bias:** Gender-specific prediction patterns likely
3. **Selection bias:** Cut-scores would disadvantage one gender

Specific Mechanisms Causing DIF

1. **Experiential Differences**
 - Men and women have different lived experiences with discrimination
 - Items trigger different reference points and memories
2. **Social Desirability**
 - Women may face pressure to appear supportive of gender equality
 - Men may feel comfortable expressing grievances
3. **Interpretation Differences**
 - Terms like "discrimination," "rights," "manliness" have gendered meanings
 - Different understanding of what constitutes bias or unfairness
4. **Reference Group Effects**
 - Men compare to other men; women to other women
 - Different baseline expectations for treatment

Recommendations for Addressing DIF

Immediate Steps

1. Conduct IRT-based DIF analysis (e.g., Mantel-Haenszel, IRT likelihood ratio)
2. Test measurement invariance using multi-group CFA
3. Examine item characteristic curves by gender

Scale Modifications

1. **Remove high-DIF items** for cross-gender comparisons

2. **Develop gender-specific norms** if scale used for individual assessment
3. **Create parallel forms** with gender-neutral anchoring items

Analytical Approaches

1. Use **DIF-adjusted scoring** when comparing across gender
2. Report **within-gender standardized scores**
3. Include **gender × item interactions** in analyses

Interpretation Guidelines

1. Avoid direct gender comparisons on raw scores
2. Focus on within-gender patterns and changes
3. Report DIF findings alongside any gender comparisons

Conclusion

The BSS is expected to show **substantial DIF across gender**, making raw score comparisons problematic. The scale likely functions as a different instrument for men versus women, measuring related but distinct constructs. This is not necessarily a flaw—it may accurately reflect genuinely different experiences and perceptions—but it requires careful handling in research and application. Gender-stratified analyses or DIF-adjusted scoring methods are essential for valid use of this scale.

Gemini 2.5pro

Q1: Evaluate each item content of BSS using LLM

Of course. Here is an evaluation of each item in the Belief in Sexism Shift (BSS) Scale.

Overview of the BSS Scale

The Belief in Sexism Shift (BSS) Scale is a psychometric tool designed to measure the perception that societal dynamics have shifted to a point where men are now the primary victims of gender-based discrimination. It assesses beliefs related to a zero-sum view of gender equality, feelings of male disadvantage, and antagonism towards feminism and women's rights initiatives.

A higher score on this scale indicates a stronger belief that men are a disadvantaged group in contemporary society and that gains for women have come at the expense of men.

Evaluation of Individual Scale Items

Here is a breakdown of what each statement aims to measure and the underlying concepts it taps into.

1. In the US, discrimination against men is on the rise.

- **Core Concept:** Perception of a negative trend for men.
- **Evaluation:** This item directly measures the respondent's belief in a changing social landscape that is increasingly hostile to men. It doesn't ask about the absolute level of discrimination but rather its trajectory. High agreement suggests a sense of growing disadvantage and grievance.

2. Men are not particularly discriminated against.*

- **Core Concept:** Denial of widespread discrimination against men. (Reverse-coded)
- **Evaluation:** This is a reverse-coded item. Strong disagreement with this statement indicates a strong belief that men *are* in fact discriminated against. Agreement suggests a view that discrimination against men is not a significant social issue. When scoring the scale, disagreement with this item contributes to a higher overall BSS score.

3. If anything, men are more discriminated against than women these days.

- **Core Concept:** Belief in reversed victimhood.
- **Evaluation:** This is one of the strongest and most direct items on the scale. It posits a complete reversal of the traditionally understood power dynamic. High agreement signifies a belief not just in discrimination against men, but that this discrimination now exceeds that faced by women.

4. Giving women more rights often requires taking away men's rights.

- **Core Concept:** Zero-sum thinking.
- **Evaluation:** This item assesses whether the respondent views rights as a finite resource. A high agreement score indicates a zero-sum worldview, where gender equality is seen as a competitive struggle rather than a process of universal uplift. This is a core tenet of modern sexism beliefs.

5. Under the guise of equality for women, men are actually being discriminated against.

- **Core Concept:** Skepticism of motives and perceived covert discrimination.
- **Evaluation:** This statement frames women's equality movements as deceptive. It implies that the stated goal ("equality") is a cover for the "real" goal of disadvantaging men. Agreement reflects deep distrust of feminist or gender-equity initiatives.

6. In the pursuit of women's rights, the government has neglected men's rights.

- **Core Concept:** Perceived institutional neglect.
- **Evaluation:** This item shifts the focus to institutional actors (the government). It measures the belief that official policies and legal frameworks aimed at helping women have created a blind spot or active disregard for issues affecting men (e.g., fathers' rights, men's health).

7. Nowadays, men don't have the same chances in the job market as women.

- **Core Concept:** Perceived economic and professional disadvantage.
- **Evaluation:** This item localizes the general feeling of discrimination to a specific, high-stakes domain: employment. Agreement suggests a belief that affirmative action, diversity initiatives, or social pressures have tilted the hiring and promotion landscape in favor of women, to the detriment of men.

8. Feminism is about favoring women over men.

- **Core Concept:** Antagonism toward feminism.
- **Evaluation:** This item assesses the respondent's definition of feminism. While academic and activist definitions of feminism center on equality, this statement frames it as a

movement for female supremacy. Agreement indicates a rejection of feminism's stated goals and aligns with common anti-feminist rhetoric.

9. Feminism does not discriminate against men.*

- **Core Concept:** The effect of feminism on men. (Reverse-coded)
- **Evaluation:** As a reverse-coded item, disagreement with this statement indicates a belief that feminism *is* discriminatory towards men. This complements the previous item. A person who believes feminism is about favoring women (Item 8) would logically disagree with this statement. Disagreement here leads to a higher BSS score.

10. All in all, men have more responsibilities and fewer benefits.

- **Core Concept:** Perception of an unfair social contract.
- **Evaluation:** This statement taps into a sense of imbalance in societal roles and rewards. A respondent agreeing with this might feel that men are still expected to fulfill traditional provider/protector roles while receiving fewer of the traditional privileges or modern benefits.

11. In today's society, women can say things that men are not allowed to say.

- **Core Concept:** Perceived double standards in social expression.
- **Evaluation:** This item relates to the "culture wars" and debates over political correctness. Agreement suggests a belief that social norms and potential repercussions for speech are applied unequally, giving women more freedom of expression on sensitive topics (like gender) than men.

12. It is evident that the media is biased against men.

- **Core Concept:** Perceived media bias.
- **Evaluation:** This item measures the belief that major cultural institutions (media) portray men in a negative light. This could refer to depictions of men as foolish, incompetent (e.g., the "doofus dad" trope), or inherently predatory.

13. In today's society, men are often punished for acting manly.

- **Core Concept:** Perceived pathologizing of masculinity.
- **Evaluation:** This taps into the debate around "toxic masculinity." A person agreeing with this statement likely believes that society is becoming hostile to traditional masculine traits (e.g., assertiveness, stoicism, competitiveness), conflating them with negative behaviors and thereby "punishing" men for being men.

14. All in all, men are well respected in today's society.*

- **Core Concept:** Global assessment of men's social status. (Reverse-coded)
- **Evaluation:** The final reverse-coded item. It asks for a holistic judgment on men's social standing. Strong disagreement indicates a belief that men are not respected. This lack of perceived respect is a cornerstone of the grievance measured by the BSS, so disagreement contributes to a higher overall score.

15. While women can use the "gender-card" to get ahead, men can't.

- **Core Concept:** Perceived strategic disadvantage.
- **Evaluation:** This statement posits an unfair tactical advantage for women. The term "gender-card" is colloquial and implies that women can leverage their gender to gain unearned advantages (e.g., in arguments, in HR disputes, in seeking promotions), an option unavailable to men. Agreement shows a belief in systemic, procedural unfairness.

Q2: Evaluate the dimensionality of BSS using LLM

Of course. Here is an evaluation of the dimensionality of the Belief in Sexism Shift (BSS) scale, performed through a conceptual analysis of its items.

Understanding Dimensionality

In psychometrics, **dimensionality** refers to the number of distinct underlying latent constructs that a scale measures.

- A **unidimensional** scale measures only one single, coherent concept. All items tap into the same underlying construct. For example, a scale measuring "height satisfaction" where all questions ask about happiness with one's height would be unidimensional. A single total score is meaningful.
- A **multidimensional** scale measures two or more distinct (though potentially related) underlying constructs.¹ For example, a "job satisfaction" scale might have separate sets of questions measuring "satisfaction with pay," "satisfaction with coworkers," and "satisfaction with work-life balance." These are different dimensions of the overall concept. Using sub-scale scores might be more informative than a single total score.

A definitive answer on dimensionality requires empirical data and statistical procedures like Factor Analysis (EFA/CFA). However, we can perform a conceptual evaluation by analyzing the content of the items to form a strong hypothesis.

Conceptual Evaluation of BSS Dimensionality

Based on an analysis of its item content, the Belief in Sexism Shift (BSS) scale is most likely **intended to be and functions as a unidimensional scale**. The items, while covering different domains (work, media, government, feminism), all converge on a single, overarching latent construct.

This central construct can be defined as: **A grievance-based belief that the societal balance of power has shifted to disadvantage men, driven by a zero-sum view of gender equality and antagonism toward feminism.**

Argument for Unidimensionality

The primary argument for a single dimension is the high degree of conceptual overlap and logical dependence among the items. A person who agrees with one item is highly likely to agree with the others because they all stem from the same core worldview.

Let's group the items thematically to illustrate this convergence:

1. **Core Belief in Reversed Discrimination:**

- 1. In the US, discrimination against men is on the rise.
- 3. If anything, men are more discriminated against than women these days.
- 2. Men are not particularly discriminated against.* (Disagreement points to this belief)
- This is the foundational belief of the construct.

2. **Causal Attribution & Ideological Framing (Zero-Sum & Anti-Feminism):**

- 4. Giving women more rights often requires taking away men's rights. (Zero-sum)
- 8. Feminism is about favoring women over men. (Anti-feminist)

- 5. Under the guise of equality for women, men are actually being discriminated against. (Links feminism directly to male discrimination)
 - 6. In the pursuit of women's rights, the government has neglected men's rights. (Institutional blame)
 - These items explain *why* the respondent believes discrimination against men is happening. They are not separate ideas, but rather the ideological justification for the core belief.
3. **Manifestations in Society (Perceived Consequences):**
- 7. Nowadays, men don't have the same chances in the job market as women. (Economic)
 - 12. It is evident that the media is biased against men. (Cultural)
 - 13. In today's society, men are often punished for acting manly. (Social Norms)
 - 15. While women can use the "gender-card" to get ahead, men can't. (Strategic)
 - These items are specific examples or perceived evidence of the core belief in action. It is unlikely a person would endorse these specific examples without also holding the core belief.

Because these thematic groups are so tightly interwoven, they likely represent facets of a single dimension rather than independent dimensions. Believing that feminism is the cause (Group 2) and that men are punished for being manly (Group 3) are both expressions of the same underlying conviction that men are now a disadvantaged group (Group 1).

Plausible Argument for a Multidimensional Structure

While the scale is likely unidimensional, one could hypothesize a potential multidimensional structure, most likely consisting of two **highly correlated** factors. If forced to split the scale, a plausible two-factor structure might be:

- **Factor 1: Perceived Male Disadvantage & Grievance.** This factor would encompass the direct feelings of discrimination and disadvantage in various domains.
 - Items: 1, 3, 7, 10, 11, 13, 14*, 15
 - This dimension captures the *personal and societal outcomes* of the perceived sexism shift.
- **Factor 2: Ideological Backlash Against Feminism.** This factor would group the more abstract and political beliefs about the causes and agents of the perceived shift, namely feminism and related initiatives.
 - Items: 4, 5, 6, 8, 9*, 12
 - This dimension captures the *ideological justification and blame attribution* for the perceived disadvantage.

Even in this scenario, these two factors would be so strongly related that a person scoring high on one would almost certainly score high on the other. In statistical terms, the correlation between these factors would likely be very high (e.g., $r > 0.70$). When factors are this highly correlated, it is often practical and theoretically justified to treat the scale as unidimensional and use a total score.

Conclusion

Based on a conceptual analysis of the item content, the **Belief in Sexism Shift (BSS) scale is best evaluated as a unidimensional instrument.**²

- **Primary Justification:** All 15 items, despite touching on different topics, are strong and convergent indicators of a single, coherent latent construct: the belief that men have become a disadvantaged and discriminated-against group in a zero-sum gender dynamic.
- **Alternative Possibility:** A multidimensional structure with two highly correlated factors ("Perceived Disadvantage" and "Ideological Backlash") is conceivable but unlikely to represent truly independent constructs. The high correlation would support the use of a single total score.

Therefore, using a single BSS score to represent an individual's overall endorsement of this belief system is a conceptually sound approach.

Q3: Evaluate whether the BSS items function differently across gender using LLM

Of course. Here is a conceptual evaluation of whether the Belief in Sexism Shift (BSS) scale items are likely to function differently across gender.

Understanding the Concept: Differential Item Functioning (DIF)

First, it is crucial to define what "functioning differently" means in this context. The psychometric term for this is **Differential Item Functioning (DIF)**.

DIF occurs when people from different demographic groups (e.g., men and women) with the **same underlying level of the trait being measured** have a different probability of endorsing a specific item.

This is different from a simple difference in average scores. We would *expect* men, on average, to score higher on the BSS scale than women. That is a valid finding about group differences. DIF is a more subtle and problematic issue of **bias**.¹ It means an item might be measuring something other than the intended trait—for example, it could be measuring a construct that is unique to the experience of being a man or a woman, thus making the item unfair.

An evaluation of DIF requires statistical analysis of data from different groups. However, we can perform a conceptual analysis to hypothesize which items are at high risk for DIF and why.

Conceptual Evaluation of DIF Across Gender for BSS Items

My overall evaluation is that several items on the BSS scale are at **high risk for exhibiting Differential Item Functioning** across gender. The scale's content is intrinsically tied to gender identity, lived experience, and social positioning, making it very likely that men and women will interpret the meaning and salience of certain questions differently, even if they share the same overall level of "Belief in Sexism Shift."

Here is a breakdown of items grouped by their likelihood of showing DIF:

High Risk for DIF

These items use language or invoke concepts that are deeply intertwined with gendered experiences and identity, making their interpretation likely to differ between men and women.

- **Items 8 & 9 (on Feminism):**
 - 8. Feminism is about favoring women over men.
 - 9. Feminism does not discriminate against men.*

- **Rationale for DIF:** The term "feminism" itself may evoke different cognitive and emotional responses. For a man with a high BSS score, feminism might be perceived as a purely external, antagonistic political force. For a woman with an equally high BSS score (e.g., a woman critical of modern feminism), her relationship with the concept is more complex; it may involve feelings of alienation or betrayal from a movement ostensibly created for her group. This difference in personal relationship to the term, independent of the level of agreement, could cause DIF.
- **Items 13 & 10 (on Masculinity and Social Roles):**
 - 13. In today's society, men are often punished for acting manly.
 - 10. All in all, men have more responsibilities and fewer benefits.
 - **Rationale for DIF:** The concepts of "acting manly" and men's "responsibilities" are experienced directly by men but observed by women. A man's agreement with Item 13 might stem from personal feelings of his behavior being policed or his identity being pathologized. A woman with the same BSS score might agree based on an observation of social discourse (e.g., debates about "toxic masculinity") but lacks the first-person, lived experience. Their reasons for endorsing the item are qualitatively different, which can lead to DIF.
- **Items 3 & 15 (on Comparative Victimhood):**
 - 3. If anything, men are more discriminated against than women these days.
 - 15. While women can use the "gender-card" to get ahead, men can't.
 - **Rationale for DIF:** These items require a direct comparison between genders. For a man to agree, he asserts the primacy of his own group's disadvantage. For a woman to agree, she must psychologically override or minimize the widely discussed disadvantages of her own group in favor of the BSS narrative. The cognitive and social process for a woman to endorse this statement is arguably more complex than for a man, making it a prime candidate for DIF.

Moderate Risk for DIF

These items refer to more general societal phenomena where gendered experiences could still shade interpretation.

- **Items 5 & 12 (on Guise and Media Bias):**
 - 5. Under the guise of equality for women, men are actually being discriminated against.
 - 12. It is evident that the media is biased against men.
 - **Rationale for DIF:** A person's media consumption and interpretation are often influenced by their identity. Men who feel targeted might be more sensitive to perceived anti-male tropes (e.g., the "doofus dad"). Women might be less likely to notice these same tropes or interpret them as harmless. Therefore, a man and a woman with the same latent BSS score might have a different threshold for agreeing that media bias is "evident."

Lower Risk for DIF

These items are more abstract and ideological, and their meaning is less likely to be interpreted differently based on one's gender. The primary difference in response should come from the latent trait itself, not from the item's wording.

- **Items 4 & 6 (on Rights and Government):**

- 4. Giving women more rights often requires taking away men's rights.
- 6. In the pursuit of women's rights, the government has neglected men's rights.
- **Rationale:** These items measure a core tenet of the BSS worldview: zero-sum thinking. This is a more abstract, philosophical belief. While men and women will disagree on its validity, the *meaning* of the statement is fairly stable. Two people (one man, one woman) who are equally high in zero-sum thinking are likely to endorse this item similarly.

Conclusion and Implications

A conceptual analysis strongly suggests that the **BSS scale is at risk for Differential Item Functioning across gender**.

1. **Source of Bias:** The bias likely originates from items that touch upon lived, gender-specific experiences (acting manly), concepts with different personal salience (feminism), and statements requiring a direct comparison of group victimhood.
2. **Impact on Validity:** If significant DIF exists, it compromises the scale's validity for making direct comparisons between men's and women's scores. A difference in the total score between a man and a woman might not solely reflect a difference in their "Belief in Sexism Shift," but also the confounding effect of how their gender influences their interpretation of the questions.
3. **Recommendation:** Before this scale is used to make strong claims about the magnitude of gender differences in this belief, it would be essential for researchers to conduct empirical DIF analyses (e.g., using Item Response Theory or logistic regression methods). If items are found to have DIF, they should either be removed or the scoring should be adjusted using models that can account for the bias.

Chatgpt-o3

Q1: Evaluate each item content of BSS using LLM

Below is an item-by-item content review of the **Belief in Sexism Shift (BSS) Scale**. For each statement I comment on (a) wording/clarity, (b) construct alignment (does it tap "belief that men are now the disadvantaged gender?"), and (c) potential problems (double-barrelled, social-desirability, ambiguity, etc.). Suggestions follow when relevant.

#	Item text	Wording / Clarity	Construct Alignment	Potential Issues & Suggestions
1	"In the US, discrimination against men is on the rise."	Simple, concrete timeframe ("on the rise") may still be interpreted differently (past 5 yrs? lifetime?).	Directly indexes perceived increase in anti-male bias.	Consider specifying a period (e.g., "in the past decade") to standardize reference.
2 (R)	"Men are not particularly discriminated against."	Clear negation; reverse-coded.	Captures baseline belief that men are not disadvantaged.	Double-negative effect when reversed in scoring can confuse respondents. Could rephrase positively and score

3	“If anything, men are more discriminated against than women these days.”	Colloquial “If anything” adds conversational tone; “these days” vague.	Compares perceived direction of inequality.	normally (“Men rarely face discrimination.”). “These days” lacks temporal anchor; delete “If anything” for brevity.
4	“Giving women more rights often requires taking away men’s rights.”	Clear cause-effect framing.	Assesses zero-sum view of gender rights.	“Often” may prompt varied interpretations; consider “necessarily” if zero-sum belief is core.
5	“Under the guise of equality for women, men are actually being discriminated against.”	Slightly dramatic phrasing (“guise”) but understandable.	Measures belief that pro-women initiatives mask anti-male bias.	Tone could prime defensiveness; “guise” may insinuate intent. A more neutral wording: “Some initiatives for women’s equality result in discrimination against men.”
6	“In the pursuit of women’s rights, the government has neglected men’s rights.”	Clear agent (government); “neglected” vs. “violated” implies omission.	Targets perceived policy bias.	Could specify domains (health, education) for concreteness, or keep broad if breadth is desired.
7	“Nowadays, men don’t have the same chances in the job market as women.”	Straightforward; “Nowadays” vague.	Job-market-specific aspect of perceived disadvantage.	Empirical claims may invite factual recall rather than attitude; that is acceptable if belief is the focus.
8	“Feminism is about favoring women over men.”	Personifies feminism; assumes monolithic movement.	Gauges anti-feminist sentiment as indicator of perceived bias.	Some respondents might distinguish strands of feminism; adding “overall” or “in practice” could reduce semantic quibbles.

9 (R)	“Feminism does not discriminate against men.”	Clear; reverse-coded counterpart to Item 8.	Complements Item 8 for balanced wording.	Same conceptual redundancy as 8; acceptable for reliability but watch multicollinearity.
10	“All in all, men have more responsibilities and fewer benefits.”	“All in all” conversational; “responsibilities” broad (family, work?).	Captures global burden/benefit appraisal.	Two constructs (responsibilities and benefits) risk double-barrelledness; could split into two items or focus on one dimension.
11	“In today’s society, women can say things that men are not allowed to say.”	Clear, concrete differential-speech claim.	Measures perceived expressive inequality.	“Not allowed” ambiguous (formal prohibition vs. social sanction). Could rephrase “without facing negative consequences”.
12	“It is evident that the media is biased against men.”	“It is evident” assumes obviousness; may pressure agreement.	Probes media-bias belief.	Remove “It is evident” to reduce demand: “The media is biased against men.”
13	“In today’s society, men are often punished for acting manly.”	“Acting manly” vague (aggression? stoicism?); “punished” strong.	Addresses norms penalizing masculinity.	Define or give example (e.g., showing traditional masculinity) to improve clarity; “punished” → “criticized” or “penalized” if physical punishment is not intended.
14 (R)	“All in all, men are well respected in today’s society.”	Straightforward; reverse-coded.	Assesses perceived male status.	“Well respected” may vary across contexts; still acceptable.
15	“While women can use the ‘gender-card’ to get ahead, men can’t.”	Informal idiom “gender-card” may be unfamiliar to some.	Captures belief in strategic advantage for women.	Provide quotation marks as done; still slang-like—consider “While women can leverage gender to get ahead...”. Also double-barrelled

(“women can” vs. “men can’t”)—could split.

General observations & recommendations

1. **Temporal vagueness**—phrases like “*these days*”, “*nowadays*”, “*today’s society*” appear in 6 items. Decide whether vague recency is acceptable; otherwise anchor to a period (e.g., “in the past five years”) for uniform interpretation.
2. **Reverse-coded items**—Items 2, 9, 14 increase method variance and help detect acquiescence, but double-negatives can confuse respondents when using 7-point agree/disagree. Consider rewriting them in the same direction and reverse-scoring silently (e.g., positive wording but still keyed opposite).
3. **Colloquialisms / idioms**—“*guise*,” “*gender-card*,” and “*acting manly*” add color but may reduce clarity or cross-cultural equivalence if you intend non-US samples. Neutral, plain language usually boosts validity.
4. **Single focus per item**—Items 4, 10, 15 each embed dual concepts (rights trade-off; responsibilities vs. benefits; women’s advantage & men’s disadvantage). Splitting those could yield cleaner factor loadings.
5. **Social-desirability cueing**—Some wording (“It is evident...”) or strong claims (“punished”) could nudge agreement or trigger reactance. Moderating intensity can lower measurement error without diluting meaning.

Implementing these modest edits will tighten wording consistency, reduce ambiguity, and likely improve the BSS’s content validity while retaining its theoretical core: the perception that gender inequality has “flipped” to disfavor men.

Q2: Evaluate the dimensionality of BSS using LLM

What do the data say about BSS’s dimensionality?

Source	Sample(s)	Key analytic results	Take-away
Original scale development (Zehnter et al., 2021)	4 US samples, N = 666	<ul style="list-style-type: none">• One-factor CFA: CFI = .93, TLI = .92, RMSEA = .10• Mis-fit traced to the three reverse-worded items• Bi-factor CFA (general BSS + “reverse-wording” factor) improved fit dramatically (CFI = .97, RMSEA = .08)	A single substantive factor accounts for the attitudes; the extra factor is method variance from item wording.
Italian validation (Manganelli et al., 2023)	2 Italian subsamples, N ≈ 687	<ul style="list-style-type: none">• EFA (PCA) → one component with strong loadings (KMO = .95,	Same unidimensional

Summary across studies	(see also PubMed abstract)	Bartlett's $\chi^2 = 5067$, $p < .001$ • CFA replicated 1-factor solution with good global fit indices (CFI > .95, SRMR < .05) Authors repeatedly describe the BSS as “a concise, one-dimensional, 15-item measure.”	structure in another language/culture. Converging evidence for one general factor.
------------------------	----------------------------	--	---

Statistical consensus: All empirical work so far supports a *single substantive latent trait*—belief that sexism has “flipped” to disadvantage men—with a *minor, orthogonal method factor* capturing reverse-keyed wording.

What does an LLM-based content review add?

When the 15 items are grouped purely by semantics, three thematic clusters emerge:

1. Perceived anti-male discrimination

Items 1, 3, 5, 6, 7, 11, 12, 13

2. Zero-sum / rights-trade-off beliefs

Items 4, 10, 15

3. Anti-feminist framing

Items 8, 9 (R), 2 (R), 14 (R)

Yet the clusters share so much conceptual overlap (all centre on men as the new victims) that their boundaries are fuzzy; inter-cluster semantic similarity is high. Hence, even a language-model clustering collapses quickly into one broader theme, mirroring the empirical single-factor solutions.

Practical implications for researchers

Question	Recommendation
----------	----------------

How to score?	Compute a single mean (or sum) score across all 15 items.
Worried about the reverse-worded items?	Use a bi-factor CFA or compute omega-hierarchical to show that the general factor explains most score variance; or simply drop the three reverse items if fit problems persist.
Want subscales?	Any sub-dimensional scoring (e.g., “media bias” vs. “rights trade-off”) should be exploratory and sample-specific. Current evidence does not justify separate subscales.
Future validation steps	<ul style="list-style-type: none"> • Run parallel analysis before EFA—expect only one eigenvalue > 1. • Compare 1-factor vs. 3-factor CFA; the latter should show very high latent correlations (> .80), signalling redundancy. • Test measurement invariance across gender, age, and culture; prior studies have found full invariance across gender.

Bottom line: Both traditional psychometrics and LLM-based content clustering converge on the conclusion that the Belief in Sexism Shift Scale is essentially **unidimensional**. Treat it as a single scale, while accounting for minor wording artefacts if you use latent-variable modelling.

Q3: Evaluate whether the BSS items function differently across gender using LLM

Do BSS items operate the same way for women and men?

Short answer:

Across two independent validations (U.S. development study and an Italian replication) the Belief in Sexism Shift Scale shows *strong* evidence of **configural and metric invariance**—the factor structure and loadings are equal for women and men. At the intercept level, the U.S. sample met full scalar invariance, whereas the Italian sample required freeing the three reverse-worded items (2, 9, 14) to reach acceptable fit, indicating **minor, wording-related DIF**. No content-driven DIF large enough to threaten the overall score has been documented.

1. Empirical evidence to date

Study (Sample)	Invariance Tests	Verdict
Zehnter et al., 2021 (N = 666, U.S.)	Multi-group CFA by gender: configural → metric → scalar → strict.	Full invariance—no gender DIF detected.

	$\Delta CFI \leq .01$ at every step; Table 8 reports acceptable fit at the scalar level.	
Morando et al., 2023 (Italian, N = 538)	Configural & metric passed. Full scalar failed; freeing intercepts of items 2, 9, 14 (reverse-keyed) produced $\Delta CFI = .006$.	Partial scalar invariance— tiny intercept DIF confined to the three reverse items (likely a method effect).

Interpretation. The latent construct behaves equivalently for men and women; score comparisons are valid once you (a) model a wording factor or (b) allow partial scalar invariance for the three reverse items.

2. LLM-based content inspection: which items *might* show gender-specific functioning?

Item #	Surface Features That Could Provoke Gender-Specific Interpretation	Empirical Flag?	DIF-Risk Rating
1	Focuses on anti-male discrimination trend; angle is clear, no negation.	None	Low
2 (R)	Negation + reverse coding; may tax verbal working memory differently.	Intercept DIF in Italian sample	Moderate
3	Direct “men vs. women” comparison; semantic symmetry.	None	Low
4	Zero-sum framing (“take away men’s rights”); might resonate more with men’s group- defence motives.	None yet	Moderate
5	“Under the guise...” implies hostile intent; may trigger threat reactions more in men.	None	Moderate

6	Blames government; broad policy frame understood by both genders.	None	Low
7	Job-market claim; women may recall wage-gap data, men may recall diversity policies—possible assimilation bias.	None	Moderate
8	Defines feminism as pro-female; women's in-group attachment to feminism could lower agreement.	None	Moderate
9 (R)	Reverse wording plus direct statement about feminism; same issues as #8 but reverse-coded.	Intercept DIF (Italian)	Moderate–High
10	Double-barrel (“more responsibilities & fewer benefits”); women may relate responsibilities to childcare differently.	None	Moderate
11	Speech-privilege claim; could be interpreted via personal anecdotes, tuition differs by gender.	None	Moderate
12	Media bias claim; media consumption differs by gender but not strongly tied to wording complexity.	None	Low
13	“Punished for acting manly” vague; men may	None	Moderate

	think of social sanctions, women of toxic masculinity.		
14 (R)	Reverse wording about respect; flagged in Italian sample.	Intercept DIF	Moderate–High
15	Idiom “gender-card”; familiarity may differ across gender and age groups.	None	Moderate

*Risk ratings combine semantic analysis with empirical flags.

3. Practical recommendations for researchers

1. **For most work, treat BSS as gender-invariant.** Full or partial invariance has been replicated cross-culturally; latent mean comparisons are defensible once partial scalar constraints are allowed for items 2, 9, 14.

2. Model or mitigate the reverse-worded items.

Option A (latent modelling): include a method factor or free their intercepts.

Option B (scoring): drop the three items; reliability remains $\alpha \approx .90$.

3. When you need item-level DIF diagnostics:

Use **ordinal logistic regression or IRT-LR DIF** with gender as grouping variable.

Flag an item if both the McFadden pseudo- $R^2 > .02$ and the χ^2 test is significant after Benjamini–Hochberg correction.

4. **If studying non-binary participants**, re-run invariance with ≥ 3 groups; semantic wording is still binary (“men vs. women”), so partial DIF is plausible.

5. Future refinements:

Rephrase the reverse items in a positive direction (e.g., “Men are generally treated fairly”) and key them negatively—this should reduce wording-related DIF.

Pilot alternative phrasing for idiomatic 15 to ensure equal comprehension across gender and age cohorts.

Current psychometric evidence shows **no substantive content-based DIF** across gender for the BSS; only the **negatively worded** items exhibit small intercept shifts that are easily modelled. Thus, researchers can compare men's and women's BSS scores with confidence—just handle those reverse-coded items with a little extra care.