The Impact of Gender Nonconformity on Adolescent Peer Evaluation and Its Dynamic Evolution (Postprint)

Authors: Wen Fangfang, Ke Wenlin, Fang Zeming, Wang Yang, Lei Yatian, Zuo Bin

Date: 2023-10-09T00:00:00+00:00

Abstract

Gender constitutes a core component of self-concept and a significant dimension of social categorization. Gender nonconformity refers to the psychological phenomenon wherein individuals exhibit behaviors that are inconsistent with or deviate from the social gender norms associated with their birth sex. In recent years, this phenomenon has become increasingly salient among adolescents. While previous research has demonstrated that gender-nonconforming adolescents encounter challenges in social adaptation, particularly in peer relationships, it has neglected the psychological mechanisms and dynamic evolutionary processes underlying the influence of gender nonconformity on adolescent peer evaluation. To address these gaps, the present study adopts a "phenomenon-explanation-prediction" logical framework, focusing on the scientific question concerning the impact of gender nonconformity on peer evaluation and its evolutionary psychological mechanisms. The overarching research objective is to examine the attributes of gender nonconformity (degree/type) as a point of departure, investigate the influence of gender nonconformity on peer evaluation and its pathway mechanisms, and subsequently explore its dynamic evolution, thereby providing potential intervention pathways for altering negative peer evaluations of gender-nonconforming individuals.

Full Text

The Effects of Gender Nonconformity on Adolescent Peer Evaluation and Related Dynamics

WEN Fangfang[1], KE Wenlin[1], FANG Zeming[2], WANG Yang[1], LEI Yatian[1], ZUO Bin[1]

(1 School of Psychology, Research Center of Social Psychology, Central China Normal University, Key Laboratory of Adolescent Cyberpsychology and Behavior, Ministry of Education, Wuhan 430079, China)
(2 Shanghai Jiao Tong University, School of Medicine, Center for Mental Health, Institute of Psychological and Behavioral Sciences, Shanghai 20030, China)

Abstract: Gender represents a core component of self-concept and a fundamental dimension of social categorization. Gender nonconformity refers to the psychological phenomenon where individuals exhibit gender norms that are inconsistent or incongruent with the social expectations associated with their birth-assigned sex. In recent years, gender nonconformity has become increasingly salient among adolescents. While previous research indicates that gender-nonconforming adolescents face challenges in social adjustment, particularly regarding peer relationships, these studies have largely overlooked the underlying psychological mechanisms and dynamic evolution processes through which gender nonconformity influences adolescent peer evaluation. Addressing these gaps, the present study follows a "phenomenon-explanation-prediction" logical framework, focusing on the scientific question of how gender nonconformity affects peer evaluation and the dynamic psychological mechanisms involved. The overarching research objective is to examine the impact of gender nonconformity—considered across varying degrees and types—on peer evaluation and its mediating pathways. Building upon this foundation, the study further explores the dynamic evolution of these processes, ultimately providing potential intervention strategies for altering negative peer evaluations of gender-nonconforming individuals.

Keywords: adolescents, gender nonconformity, prototype matching, peer evaluation, dynamic evolution
Classification Number: B844

---

1. Problem Statement

Gender, as an innate characteristic, constitutes a core component of self-concept and a crucial dimension of social categorization (Steiner et al., 2022), exerting profound influences on individual psychology, interpersonal relationships, intergroup behavior, and social development. Topics such as gender stereotypes, gender socialization, and gender attitudes have remained central themes in social psychology, developmental and educational psychology, and sociology. Recently, gender nonconformity—defined as individuals exhibiting psychological phenomena inconsistent with the gender norms associated with their birth-assigned sex—has become increasingly prominent among adolescents (Jacques et al., 2022; Qian et al., 2021; van der Miesen et al., 2018). Gender nonconformity typically manifests through adolescents' gender-"atypical" behaviors and interests, including preferences for toys, activities, clothing/hairstyles, and playmates (Cherney & Dempsey, 2010), and represents a violation of culturally defined gender norms

(Lo et al., 2020; Macmullin et al., 2020; Turban & Ehrensaft, 2018).

Current research has yet to fully understand the socio-psychological status and acceptance of gender-nonconforming individuals, with inconsistent findings across studies. Some research suggests that gender-nonconforming individuals may exhibit better psychological health, such as higher well-being among androgynous individuals (Anzani et al., 2022; Cai et al., 2008; Qian et al., 2000). Conversely, other studies find that gender-nonconforming adolescents face behavioral and emotional challenges, including self-harm/suicidality, low psychological well-being, poor peer relationships, maladaptive psychological adjustment, and social maladjustment (Lamer et al., 2022; MacMullin et al., 2020).

One reason for these contradictory findings may lie in differences in how previous studies have conceptualized the psychological representation of gender nonconformity. Prior research has employed a traditional dichotomous approach, positing only a consistent or inconsistent relationship between social gender and biological sex. For example, "gentleness" is traditionally considered a feminine trait; therefore, men exhibiting "gentleness" would be deemed inconsistent with their biological sex and thus characterized as gender-nonconforming. Wen et al. (2020) expanded beyond the traditional "male-or-female" dichotomy by using "perceived gender" to explore the relationship between counter-stereotypicality and peer acceptance. They discovered that acceptance of targets possessing cross-gender traits exhibits an evaluative threshold—beyond a certain degree, the nature of evaluation changes, showing a "rise-cliff" curve that initially increases then decreases with greater counter-stereotypicality. This research demonstrates that although certain traits are traditionally gender-typed, their presence in both sexes can yield positive interpersonal benefits. Drawing from the structural perspective of general and specific factors in personality and ability, we propose that traits like "gentleness" are not exclusively feminine gender properties (GP) but rather basic properties (BP) that exist within gender prototypes for both sexes. In other words, gender-associated descriptive features apply to both men and women, though their evaluative impact differs. For instance, "gentleness" exists as an attribute in both male and female stereotypes, but societal expectations for gentleness in men are weaker than those for women. This theoretical assumption of gender features as basic properties offers a new perspective for understanding gender nonconformity, which can thus be viewed as a cognitive outcome where certain basic properties of an individual exceed the gendered degree set by others or society.

Aligned with our research objectives, this study does not broadly address the diverse socio-psychological effects of gender nonconformity (such as acceptance levels, mental health, social adaptation, and well-being) (MacMullin et al., 2020; Zucker et al., 1995). Instead, we focus specifically on peer evaluation as our key variable. Evaluation refers to the feelings of acceptance and personality judgments that people generate based on limited information, playing an extremely important role in social life (Simon et al., 2020). As a crucial form of social evaluation, peer evaluation describes the acceptance or rejection, welcome or

unwelcome attitudes that group members hold toward an individual. Such attitudes represent a vital pathway for adolescents to understand themselves, others, and their social environment, and constitute a key factor influencing their peer relationships and mental health (Abele et al., 2021; Halimi et al., 2021; Wang et al., 2021; Zhou et al., 2015).

This study builds upon debates in research concerning gender nonconformity' s impact on evaluation, using new theoretical assumptions to explore underlying psychological and cognitive mechanisms. Following a phenomenon-explanation-prediction logic, we have designed three studies across implementation, algorithmic, and computational levels. Study 1 will further formalize the basic property view of gender features, collecting behavioral and neurophysiological evidence from multiple angles at the implementation level to understand the continuous representation of gender features in psychological cognition. Study 2 will, based on the psychological representation of gender features, further explore both the cognitive characteristics of gender prototypes and the motivational characteristics of approach-avoidance, proposing and testing a dual-pathway mechanism of cognition-motivation to dissect the phenomenon at the algorithmic level. Study 3 will employ computational modeling techniques at the computational level to explore the dynamic evolutionary mechanisms of peer evaluation of gender-nonconforming individuals, predicting phenomena not yet addressed in empirical research while attempting to provide interventions for altering negative peer evaluations of gender-nonconforming individuals.

---

2. Literature Review

2.1 (Counter)Gender Stereotypes and Gender Nonconformity

Gender stereotype represents a fundamental form of stereotype research, referring to expectations, requirements, and generalized views regarding the behavior and personality characteristics of men or women (Liu & Zuo, 2006). This research area has advanced alongside broader stereotype research, currently focusing on content dimensions, cognitive processing, and neural mechanisms. (1) Regarding content dimensions, Fiske and colleagues proposed the Stereotype Content Model (SCM) of warmth and competence; Abele and Wojciszke proposed communion and agency dimensions; and Ellemer and colleagues proposed a three-dimensional model of sociability, morality, and competence. These content dimensions provide reference categories for understanding gender nonconformity and evaluation (Abele et al., 2021; Cheng et al., 2012; Zuo, 2015). Research suggests that gender stereotypes comprise aspects including physical appearance, personality traits, role behaviors, and occupational tendencies (Fiske, 2018; Fiske et al., 2007; Deaux & Lewis, 1984). (2) Regarding cognitive processing, researchers have indicated that social categorization and stereotype activation may represent two distinct psychological processes, leading to the proposal of a Two-Stage Model of Stereotype Activation (Zhang & Zuo, 2012). Addition-

ally, some researchers have proposed a goal-based theory of stereotype activation and application, suggesting that stereotypes are activated when their application facilitates goal achievement and suppressed when it hinders goal attainment (Kunda & Spencer, 2003). Following activation, stereotypes may undergo either serial or parallel cognitive processing (Zhang & Wang, 2009). In serial processing, stereotypical information involves only automatic processing, whereas in parallel processing, it involves both automatic and controlled processing. (3) Regarding neural mechanisms, research has found that stereotype-inconsistent contexts elicit more significant N400 in frontal-central regions, while ingroup category word priming evokes larger P600. Neural changes in stereotypes primarily manifest in people's reactions to counter-stereotypical individuals (Wang et al., 2010). Studies show that stereotypical knowledge representations about people and social groups stored in the lateral temporal lobe, particularly the anterior temporal lobe (ATL), are collected in the dorsomedial prefrontal cortex (dmPFC) to support impression formation (i.e., stereotypes). Mitchell et al. (2008) used fMRI to discover that heuristic application of gender stereotypes may rely on the right prefrontal cortex. Furthermore, Endendijk et al.' s (2019) fMRI research found that when parents view children exhibiting counter-stereotypical behavior, brain regions involved in gender stereotype inhibition, behavioral monitoring, response selection, and control—namely the dmPFC and anterior cingulate cortex (ACC)—become activated.

As a concept opposite to gender stereotypes, counter-gender stereotype refers to individuals who fail to conform to or violate gender stereotype expectations, describing psychological phenomena where gender role behavior contradicts people' s gender stereotypical expectations (Heilman et al., 2004; Liu & Zuo, 2006). For example, a female supervisor's "agentic"behavior violates the "communal"stereotypical expectation for women (Tan et al., 2021). Counter-gender stereotypes primarily focus on attributes/characteristics associated with men and women, emphasizing inconsistency with perceivers'cognitive gender schemas (Cheryan & Markus, 2020). In contrast, gender nonconformity emphasizes characteristics in appearance, behavior, and psychology that are inconsistent with social norms based on birth-assigned sex (e.g., feminine boys or masculine girls) (Adelson, 2012; Jacques et al., 2022; Olson & Enright, 2018). While counter-gender stereotypes emphasize violations of gender schemas and have been studied primarily in work contexts focusing on backlash effects against counter-stereotypical individuals (Phelan & Rudman, 2010; Rudman et al., 2012), gender nonconformity uses birth-assigned sex as a reference, imposing requirements on individuals' spontaneous psychological and behavioral characteristics. Gender nonconformity research has focused primarily on adolescent development, examining its effects on mental health, well-being, peer acceptance, and potential roles of cultural environment, parenting styles, and peer relationships (MacMullin et al., 2020; Jacques et al., 2022; Wang et al., 2021). Given these definitional and empirical differences, gender nonconformity research remains in its early stages. However, considering that previous experiments often used similar manipulation methods for both concepts, counter-gender stereotype research provides an

important foundation for gender nonconformity studies.

Counter-gender stereotypes manifest in two forms: (1) failing to exhibit features that gender stereotypes "should have," and (2) exhibiting features that gender stereotypes "should not have" (Wen et al., 2020). Individuals possessing counter-gender stereotypes receive negative evaluations for violating social expectations (Liu & Zuo, 2006; Xu, 2003). The emergence of counter-gender stereotypical information prompts individuals to spontaneously adopt cognitive and behavioral strategies to maintain their gender stereotypes. Wen et al. (2020) conceptualized counter-gender stereotypes as a continuum from extreme gender stereotyping to gender-common characteristics, proposing a "threshold model of stereotype maintenance" that confirms an inverted U-shaped functional relationship between counter-stereotypicality and individual attitudes (Xu, 2022). This model reveals possible evolutionary mechanisms behind psychological androgyny and provides an important foundation for gender nonconformity research.

2.2 The Relationship Between Gender Nonconformity and Peer Evaluation

Previous research has investigated gender nonconformity in relation to peer preference, peer relationships, self-concept, and mental health, yielding complex results that provide indirect evidence for contradictory findings regarding the relationship between gender nonconformity and peer evaluation. Some studies indicate that people accept targets with moderate levels of cross-gender characteristics more than those with low/high levels (e.g., moderately masculine women receive higher acceptance ratings than low or highly masculine women) (Wen et al., 2020), offering indirect evidence for potentially positive evaluations of gender-nonconforming individuals. Conversely, other research demonstrates that gender nonconformity leads to negative peer evaluations (Nabbijohn et al., 2020; Wang et al., 2021; Zosuls et al., 2016), with adolescents evaluating gender-nonconforming individuals less positively than gender-conforming peers (Blakemore, 2003; MacMullin et al., 2020), and showing greater reluctance to interact with gender-nonconforming peers (Kwan et al., 2019).

These inconsistent results may be influenced by the type and degree of gender nonconformity in the target, as well as the perceiver's own gender. How do these factors affect evaluations of gender-nonconforming peers? Previous research has made initial explorations into this question.

2.3 Boundary Conditions for Gender Nonconformity' s Effect on Peer Evaluation

2.3.1 Degree of Gender Nonconformity Gender nonconformity is not simply a dichotomy of conformity versus nonconformity; different degrees produce different effects on peer evaluation. Forming impressions based on others' behaviors, physical characteristics, and faces represents the most efficient evaluation method (Zhang & Zhu, 2012; Zuo et al., 2019). Previous research indicates that

women exhibiting a certain degree of independence and bravery receive positive evaluations, but excessive independence and bravery elicit negative evaluations and backlash (Bosak et al., 2018). Wen et al. (2020) argue that gender characteristics (e.g., "gentleness," "bravery") exist within people' s gender norms and prototypes, but with different required degrees for men and women. Around the "optimal" degree required for each gender, evaluations of gender-nonconforming targets gradually decrease, showing an inverted U-shaped pattern.

2.3.2 Type of Gender Nonconformity Classifications of counter-gender stereotypes provide reference for categorizing types of gender nonconformity. Previous research distinguishes between descriptive and prescriptive counter-gender stereotypes (Rudman et al., 2012). Descriptive stereotypes reflect the typicality of a gender characteristic among men/women in general, while prescriptive stereotypes reflect what a culture (e.g., nation, ethnicity, region) believes men/women should possess, including both engaging in "should-not-do" behaviors and failing to do "should-do" behaviors (Heilman, 2012). Counter-gender stereotypes violate perceivers' cognitive gender schemas, which can be either descriptive (organizing knowledge about gender) or prescriptive (dictating how others should behave). Gender nonconformity, from the actor's perspective, involves concrete violations of prescriptive requirements based on birth-assigned sex. Drawing from prescriptive counter-gender stereotype classifications, this study categorizes gender nonconformity into: (1) With the Other-gender Feature (WOF), and (2) Lack the Same-gender Feature (LSF). Limited research suggests differential evaluative reactions to these two types, with people rejecting LSF but potentially accepting WOF (Heilman & Okimoto, 2007), though most studies have focused only on WOF. This study systematically examines how both WOF and LSF types affect adolescent peer evaluation.

2.3.3 Gender Differences Gender serves as an important moderating variable in how gender nonconformity affects peer evaluation, including target gender, perceiver gender, and their interaction (Nielson et al., 2022). Gender norms are stricter for males than females throughout socialization (Salvati et al., 2021). Boys are encouraged to enact "male roles,"and face harsher criticism from parents and peers than girls when they deviate from these roles (Bosson & Michniewicz, 2013). For example, boys who challenge gender stereotypes (compared to girls) hold lower social status among peers and receive more negative evaluations (Nabbijohn et al., 2020). Relatedly, research indicates that boys (compared to girls) have a narrower and less flexible range of acceptable gender behaviors (Green et al., 2004). Boys with high gender nonconformity receive poorer community peer evaluations (MacMullin et al., 2020) and experience greater peer rejection (Braun & Davidson, 2017); gender nonconformity pressure also triggers more school misconduct in boys than girls (Heyder et al., 2021). The interaction between perceiver and target gender appears in findings that children evaluate opposite-sex gender-nonconforming peers more leniently than same-sex peers, though this may be domain-specific (Kwan et al., 2019). The gender differences

and mechanisms in how gender nonconformity affects peer evaluation warrant systematic investigation.

2.4 Psychological Mechanisms of Gender Nonconformity' s Effect on Peer Evaluation

The gender stereotype maintenance model and Status Incongruity Hypothesis (SIH) suggest that people resist gender-nonconforming individuals to maintain self-esteem. Therefore, perceivers' evaluations are influenced by cognitive resources and motivational processing (Rudman et al., 2012). Based on the lazy brain theory, some research has used the Efficient Coding Principle to explain the psychological mechanisms of gender nonconformity (Barlow, 1961). This principle posits that people encode and represent input stimuli in the most efficient, resource-saving manner, most commonly through sparse coding—encoding only what differs from existing prototypes (Olshausen & Field, 1996). The better the match between input information and prototypes, the more concise the coding, and this conciseness produces pleasure. If gender norms (stereotypes) are viewed as prototypes and image descriptions as input stimuli, then individuals who better conform to gender norms will be more positively evaluated.

During impression evaluation, people "match" target characteristics with category prototype features and categorize targets as group members showing the best match on characteristic sets (Wang, 1999). Some research suggests that information processing involves self-group bias in prototype information (Wen et al., 2020), and that individuals resist information inconsistent with gender stereotypes because prototype consistency reduces cognitive resource consumption (Wilder et al., 1996; Chen, 2021). Other research indicates that perceivers actively engage in self-regulation to rationalize information violating gender norms, thereby reducing the impact of gender norms on their behavior, and that approach or avoidance intentions elicited by information valence affect evaluative responses (Allidina, 2018; Zhang et al., 2012). People evaluate gender-nonconforming individuals with different valence information differently, possibly influenced by promotion orientation and preventive orientation intentions (Moss-Racusin & Rudman, 2010). Thus, the psychological mechanisms underlying peer evaluation of gender-nonconforming adolescents may be closely related to perceivers' approach-avoidance intentions toward targets.

2.5 Limitations of Existing Research

Analysis of existing research on gender nonconformity' s effect on peer evaluation reveals several limitations despite accumulated paradigms and theoretical achievements. First, previous studies have primarily employed a traditional dichotomous approach of gender conformity versus nonconformity. Second, they have lacked investigation into the cognitive-motivational pathway mechanisms of prototype matching and approach-avoidance intentions underlying gender nonconformity' s effect on adolescent peer evaluation. Third, previous research has

adopted a static perspective, neglecting the dynamic generation and evolution of gender nonconformity.

Based on domestic and international findings and relevant theories, this study addresses these limitations by focusing on the psychological essence of gender nonconformity—the basic property view. By moving beyond binary gender categories and examining the relationship between basic properties and gender attributes, we aim to explain both the static effects and dynamic changes in the psychological mechanisms through which gender nonconformity influences peer evaluation.

---

3. Research Design

Following a "phenomenon-explanation-prediction" logical framework, this study focuses on the scientific question of how gender nonconformity affects peer evaluation and its evolutionary psychological mechanisms. The overall objective is to use attributes of gender nonconformity (different degrees/types) as an entry point to examine its impact on peer evaluation and underlying pathways, and to explore dynamic evolution toward providing possible interventions for changing negative evaluations of gender-nonconforming individuals. Specific objectives include: (1) constructing a basic property view of gender features in gender nonconformity and providing neurophysiological evidence; (2) revealing dual cognitive-motivational pathway mechanisms of prototype matching and approach-avoidance intentions; and (3) investigating the dynamic evolutionary mechanisms of gender nonconformity's effect on peer evaluation. The theoretical model is illustrated in Figure 1 [Figure 1: see original paper].

To achieve these three objectives, this study systematically examines the effects and evolutionary psychological mechanisms of gender nonconformity on peer evaluation across progressively advancing levels: implementation, algorithmic, and computational. The research comprises three aspects: (1) multi-form, fine-grained examination of gender nonconformity' s impact on peer evaluation, constructing the basic property view and providing behavioral and neurophysiological evidence at the implementation level; (2) revealing cognitive-motivational dual-pathway mechanisms at the algorithmic level; and (3) exploring dynamic evolutionary mechanisms at the computational level using computational modeling. The overall framework is shown in Figure 2 [Figure 2: see original paper].

3.1 Study 1: The Basic Property View of Gender Features in Gender Nonconformity

Study 1 aims to explore, through two sub-studies using behavioral and neurophysiological indices, what effects gender nonconformity has on peer evaluation at the phenomenological level, thereby constructing and refining the basic property view of gender features in gender nonconformity. This perspective builds on Wen et al.' s (2020) findings that traits like "gentleness" and "resoluteness"

are not exclusive gender properties under traditional dichotomous thinking, but rather basic properties (BP) existing in both gender prototypes, differing only in degree. For instance, "gentleness" is a characteristic possessed by both men and women, but gender norms demand it less in men than in women. Study 1 formalizes this description mathematically: any gender image can be represented by a high-dimensional vector $\mathbf{x} = [x_1, x_2, ..., x_n]$, where each dimension corresponds to a gender feature and each value represents the degree of that stereotypical feature. Here, each dimension' s value $x_i \in [0, 1]$ is obtained by normalizing a 7-point Likert scale (Wen et al., 2020). For example, if dimension 1 represents "gentleness," dimension 2 "resoluteness," and dimension 3 "intelligence," a male image might be expressed as $[0.4, 0.8, 0.4]$. Based on this definition, we can express male and female norms separately. For instance, male norms across these dimensions might be $\mathbf{p}_m = [0.3, 0.9, 0.5]$, while female norms might be $\mathbf{p}_f = [0.8, 0.2, 0.55]$.

To clarify, Figure 3 [Figure 3: see original paper] classifies "gentleness," "resoluteness," and "intelligence" using both traditional dichotomous and the new "basic property view" approaches. For typical gender features ("resoluteness" and "gentleness"), the two classification methods show no obvious difference. However, for less representative/atypical features like "intelligence," substantial differences emerge. Under traditional dichotomy, this term would be categorized as feminine alongside the typical feature "gentleness," whereas the basic property view predicts classification uncertainty that reflects feature atypicality.

Study 1 includes two sub-studies. Study 1a tests the "basic property view"of gender features and its derived definition of gender nonconformity degrees. Using classification tasks, it explores whether participants' reaction times during peer evaluation correlate with uncertainty predictions made by the basic property view. Study 1b further employs event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI) to examine electrophysiological and brain region changes in adolescent peer evaluation caused by different degrees and types of gender nonconformity.

Study 1 advances previous research in two ways: First, it manipulates gender nonconformity using both WOF and LSF types, tested across multiple modalities including faces, behaviors, and traits. Second, it employs three different methods to operationalize degrees of gender nonconformity: (1) participants directly report stimulus targets' masculinity/femininity typicality on 7-point scales, with negative values defining nonconformity degree; (2) participants rate masculinity/femininity degrees, using Cohen' s d between two normal distributions to define nonconformity degree; and (3) participants directly select the gender ("male" or "female") corresponding to gender features, using negative selection frequency as the nonconformity index.

3.2 Study 2: Dual-Pathway Mechanisms of Gender Nonconformity' s Effect on Peer Evaluation

Study 1 obtained an inverted U-curve by manipulating gender nonconformity degree. Study 2 aims to explain this curve' s formation and uncover its psychological mechanisms. Based on previous theories, we propose a linear cognitive-motivational dual-pathway model affecting peer evaluation y:

$$y = \beta_0 + \beta_1 \cdot (-d(\mathbf{x}, \mathbf{p})) + \beta_2 \cdot v + \epsilon$$

where y is the peer evaluation variable, $\mathbf{x}$ represents the current person image, $\mathbf{p} \in \{\mathbf{p}_m, \mathbf{p}_f\}$ represents gender norms (with $\mathbf{p}_m$ as male norms and $\mathbf{p}_f$ as female norms), $\beta_0, \beta_1, \beta_2$ are model parameters, d is a distance metric, and v is feature valence.

For the cognitive "prototype matching" pathway, we define the negative distance between any person image $\mathbf{x}$ and its gender norm as $-d(\mathbf{x}, \mathbf{p})$. If the difference between $\mathbf{x}$ and $\mathbf{p}$ is small, then $-d(\mathbf{x}, \mathbf{p})$ is relatively large, indicating high prototype matching and low gender nonconformity. If the difference is large, $-d(\mathbf{x}, \mathbf{p})$ is small, indicating low prototype matching and high gender nonconformity. Here, gender nonconformity represents the psychological phenomenon of prototype matching mechanisms. Feature valence is introduced as a second motivational pathway. Model testing occurs in two phases: Phase 1 reuses Study 1 data to fit the model, ensuring quantitative accuracy of the dual-pathway mechanism. Phase 2 experimentally tests whether these mechanisms actually operate in peer evaluation.

Study 2 comprises three sub-studies. Study 2a explores the role of "prototype matching" cognitive activation, first examining the priority of gender prototype activation in gender nonconformity' s effect on peer evaluation, then using a delayed matching paradigm to measure reaction differences to prototype versus non-prototype faces based on normative coding models. Study 2b investigates "approach-avoidance intention" motivational activation, first using Go/No-Go paradigms to test potential associations between different valences (positive/negative) of typical male/female trait words and approach-avoidance intentions, then exploring the mediating role of approach-avoidance intentions. Study 2c integrates these perspectives using stimulus-response compatibility tasks and situational experiments to examine potential interactions between cognitive and motivational mechanisms.

3.3 Study 3: Dynamic Evolution of Gender Nonconformity' s Effect on Peer Evaluation

Study 3 extends previous findings to the temporal dimension at a dynamic level. Three sub-studies will use reinforcement learning paradigms to simulate gender prototype learning processes and examine how gender nonconformity affects peer evaluation across different learning stages (duration). The fundamental

scientific question is: Why do people have different degree requirements for normative attributes of men and women? How do these degrees form? Our hypothesis posits important roles for social learning. To support this, we will use reinforcement learning frameworks to simulate this process.

Study 3 includes three sub-studies. Study 3a examines the formation of gender nonconformity, investigating whether social learning theory based on reinforcement learning can explain its effects on peer evaluation. Using multi-trait image comprehensive evaluation, it explores the weight of trait typicality and word frequency on peer evaluation, and how typicality changes affect prototype acquisition. Study 3b builds on this using implicit association test principles and affect misattribution procedures, employing post-priming reinforcement learning paradigms to examine the dynamic role of motivation/valence in cognitive prototype formation and change. Study 3c uses more ecologically valid field experiments (individual reading information reward tasks and social interaction feedback tasks) to discuss the possibility of social learning of gender nonconformity prototypes in realistic social contexts, revealing dynamic evolutionary mechanisms and providing intervention pathways for altering negative peer evaluations.

---

4. Theoretical Construction and Research Significance

Theoretically, this study addresses key scientific questions: Where does gender cognition originate? How do gender prototypes form? Can they be gradually formed and changed through learning? We propose a basic property view of gender features to understand gender nonconformity, construct a cognitive-motivational dual-pathway mechanism for its effect on peer evaluation, and use model-based reinforcement learning to simulate prototype formation and dynamic evolution. The basic property view breaks through traditional dichotomous approaches in gender research and evaluation, offering theoretical novelty that helps explain contradictory previous findings. The dual-pathway mechanism reveals psychological mechanisms from a static perspective, while the algorithmic and implementation levels derived from reinforcement learning' s "maximizing reward" computational theory help reveal and verify dynamic evolutionary mechanisms. These theoretical models and validations innovate and enrich theoretical knowledge in social psychology and developmental psychology regarding (counter)gender stereotypes and adolescent gender development. This study represents only a starting point; we believe the theoretical framework can explain more gender-related social psychological phenomena.

Practically, our findings can provide management and educational insights for promoting adolescent gender development, peer relationships, and mental health from a gender perspective. For instance, by revealing psychological mechanisms through which different types and degrees of gender nonconformity affect peer evaluation, we offer possible improvement methods and motivational interven-

tion pathways for addressing school bullying (peer rejection and bullying of gender-nonconforming individuals). Using reinforcement learning to compute and simulate dynamic evolutionary mechanisms provides important practical value for media, schools, families, and society to potentially associate gender-nonconforming adolescents with positive attributes, thereby promoting positive peer evaluation and interpersonal relationships.

References

Cai, H., Huang, X., & Song, H. (2008). The relationship model between gender roles and subjective well-being: A test based on Chinese college students. *Acta Psychologica Sinica*, 40(4), 474–486.

Chen, L. (2021). *Stereotypes: Formation and Change*. China Renmin University Press.

Cheng, J., Guan, J., & Wang, X. (2012). Consensus-based discrimination and stereotypes: A case study of migrant workers and urban residents. *Chinese Journal of Clinical Psychology*, 20(4), 543–546.

Liu, X., & Zuo, B. (2006). Psychological mechanisms of gender stereotype maintenance. *Advances in Psychological Science*, 14(3), 456–461.

Qian, M., Zhang, G., Luo, S., & Zhang, X. (2000). Development of the Chinese Sex-Role Inventory (CSRI) for college students. *Acta Psychologica Sinica*, 32(1), 99–104.

Wang, P. (1999). A review of social cognition research on stereotypes. *Psychological Science*, 22(4), 342–345.

Wang, P., Yang, Y., & Zhao, L. (2010). Activation effects of stereotypes: Behavioral and ERP evidence. *Acta Psychologica Sinica*, 42(5), 607–617.

Wen, F., Zuo, B., Ma, S., & Xie, Z. (2020). Ingroup bias in face recognition. *Advances in Psychological Science*, 28(7), 1164–1171.

Xu, D. (2003). Research on gender effects of gender stereotypes. *Psychological Science*, 26(4), 741–742.

Xu, Y. (Ed.). (2022). *Social Psychological Research*. East China Normal University Press.

Zhang, X., & Wang, P. (2009). Mechanisms of stereotype threat: Advances in cognitive neuroscience research. *Chinese Journal of Special Education*, (11), 75–79.

Zhang, X., & Zuo, B. (2012). A two-stage model of stereotype activation based on face perception. *Acta Psychologica Sinica*, 44(9), 1189–1201.

Zhang, X., Xuan, Y., & Fu, X. (2012). The effect of emotional valence on approach-avoidance responses. *Advances in Psychological Science*, 20(7), 1023–

1030.

Zhang, Y., & Zhu, L. (2012). Impression formation based on faces: A neuroscience perspective. *Advances in Psychological Science*, 20(7), 1031–1039.

Zhou, Z., Sun, X., Zhao, D., Tian, Y., & Fan, C. (2015). Research on the development of peer relationships. *Psychological Development and Education*, 31(1), 62–70.

Zuo, B. (2015). *Stereotype Content and Form*. Central China Normal University Press.

Barlow, H. (1961). Possible principles underlying transformations of sensory messages. *Sensory Communication*, 1, 217–234.

Blakemore, J. E. O. (2003). Children' s beliefs about violating gender norms: Boys shouldn' t look like girls, and girls shouldn' t act like boys. *Sex Roles*, 48, 411–419.

Bosak, J., Kulich, C., Rudman, L., & Kinahan, M. (2018). Be an advocate for others, unless you are a man: Backlash against gender-atypical male job candidates. *Psychology of Men & Masculinity*, 19(1), 156–165.

Bosson, J. K., & Michniewicz, K. S. (2013). Gender dichotomization at the level of ingroup identity: What it is, and why men use it more than women. *Journal of Personality and Social Psychology*, 105(3), 425–442.

Braun, S. S., & Davidson, A. J. (2017). Gender (non)conformity in middle childhood: A mixed methods approach to understanding gender-typed behavior, friendship, and peer preference. *Sex Roles*, 77(1), 16–29.

Cherney, I. D., & Dempsey, J. (2010). Young children' s classification, stereotyping and play behaviour for gender neutral and ambiguous toys. *Educational Psychology*, 30(6), 651–669.

Cheryan, S., & Markus, H. R. (2020). Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review*, 127(6), 1022–1052.

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology*, 46(5), 991–1004.

Endendijk, J. J., Smit, A. K., van Baar, A. L., & Bos, P. A. (2019). Boys' toys, girls' toys: An fMRI study of mothers' neural responses to children violating gender expectations. *Biological Psychology*, 148, 107776.

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27(2), 67–73.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83.

Green, V. A., Bigler, R., & Catherwood, D. (2004). The variability and flexibility of gender-typed toy play: A close look at children' s behavioral responses to counterstereotypic models. *Sex Roles*, 51, 371–386.

Halimi, M., Davis, S. N., & Consuegra, E. (2021). The power of peers? Early adolescent gender typicality, peer relations, and gender role attitudes in Belgium. *Gender Issues*, 38, 210–237.

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32(1), 113–135.

Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communality deficit. *Journal of Applied Psychology*, 92(1), 416–427.

Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology*, 89(3), 416–427.

Heyder, A., van Hek, M., & van Houtte, M. (2021). When gender stereotypes get male adolescents into trouble: A longitudinal study on gender conformity pressure as a predictor of school misconduct. *Sex Roles*, 84, 61–75.

Jacques, K. P., Feinstein, B. A., Darling, A. K., & Humphreys, K. L. (2022). An analogue study investigating differential parenting of gender conforming and nonconforming boys. *Archives of Sexual Behavior*, 51, 3569–3581.

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, 129(4), 522–544.

Kwan, K. M. W., Shi, S. Y., Nabbijohn, A. N., Macmullin, L. N., Vanderlaan, D. P., & Wong, W. I. (2019). Children' s appraisals of gender nonconformity: Developmental pattern and intervention. *Child Development*, 91(4), e780–e798.

Lamer, S. A., Dvorak, P., Biddle, A. M., Pauker, K., & Weisbuch, M. (2022). The transmission of gender stereotypes through televised patterns of nonverbal bias. *Journal of Personality and Social Psychology*, 123(6), 1315–1335.

Lo, S. Y., King, J. T., & Lin, C. T. (2020). How does gender stereotype affect the memory of advertisements? A behavioral and electroencephalography study. *Frontiers in Psychology*, 11, Article 1580.

MacMullin, L. N., Aitken, M., Nabbijohn, A. N., & VanderLaan, D. P. (2020). Self-harm and suicidality in gender-nonconforming children: A Canadian community-based parent-report study. *Psychology of Sexual Orientation and Gender Diversity*, 7(1), 76–90.

Mitchell, J. P., Ames, D. L., Jenkins, A. C., & Banaji, M. R. (2008). Neural correlates of stereotype application. *Journal of Cognitive Neuroscience*, 21(3), 594–604.

Moss-Racusin, C. A., & Rudman, L. A. (2010). Disruptions in women' s self-promotion: The Backlash Avoidance Model. *Psychology of Women Quarterly*, 34(2), 186–202.

Nabbijohn, A. N., MacMullin, L. N., Kwan, K. M. W., Santarossa, A., Peragine, D. E., Wong, W. I., & VanderLaan, D. P. (2020). Children' s bias in appraisals of gender-variant peers. *Journal of Experimental Child Psychology*, 196, 104865.

Nielson, M. G., Rogers, A. A., & Cook, R. E. (2022). Nuanced longitudinal effects of domains of perceived gender similarity on adolescent peer victimization. *Sex Roles*, 86(9-10), 559–575.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583), 607–609.

Olson, K. R., & Enright, E. A. (2018). Do transgender children (gender) stereotype less than their peers and siblings? *Developmental Science*, 21(4), e12606.

Phelan, J. E., & Rudman, L. A. (2010). Reactions to ethnic deviance: The role of backlash in racial stereotype maintenance. *Journal of Personality and Social Psychology*, 99(2), 265–281.

Qian, M., Wang, Y., Wong, W. I., Fu, G., Zuo, D. P., & Vanderlaan, D. P. (2021). Correction to: The effects of race, gender, and gender-typed behavior on children' s friendship appraisals. *Archives of Sexual Behavior*, 50(3), 1073–1074.

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48(1), 165–179.

Salvati, M., Passarelli, M., Chiorri, C., Baiocco, R., & Giacomantonio, M. (2021). Masculinity threat and implicit associations with feminine gay men: Sexual orientation, sexual stigma, and traditional masculinity. *Psychology of Men & Masculinities*, 22(4), 649–668.

Simon, J. C., Styczynski, N., & Gutsell, J.N. (2020). Social perceptions of warmth and competence influence behavioral intentions and neural processing. *Cognitive, Affective & Behavioral Neuroscience*, 20, 265–275.

Steiner, T. G., Vescio, T. K., & Adams Jr, R. B. (2022). The effect of gender identity and gender threat on self-image. *Journal of Experimental Social Psychology*, 101(5), 104335.

Tan, X., Zuo, B., Wen, F., Xie, Z., & Song, S. (2021). Fear of backlash moderates female senior executives' communion (but not agency) as compared to female lecturers. *Frontiers in Psychology*, 12, 520590.

Turban, J. L., & Ehrensaft, D. (2018). Research review: Gender identity in youth: Treatment paradigms and controversies. *Journal of Child Psychology*

and Psychiatry, 59(12), 1228–1243.

van der Miesen, A. I., Nabbijohn, A. N., Santarossa, A., & VanderLaan, D. P. (2018). Behavioral and emotional problems in gender-nonconforming children: A Canadian community-based study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57(7), 491–499.

Wang, Y., Qian, M., Nabbijohn, A. N., Wen, F., Fu, G., Zuo, B., & VanderLaan, D. P. (2021). Culture influences the development of children' s gender-related peer preferences: Evidence from China and Thailand. *Developmental Science*, 25(4), e13221.

Wen, F., Zuo, B., Wang, Y., Wu, Y., Fang, Z., & Ma, S. (2020). The (continuous) nature of perceived gender counter-stereotype: A threshold model of gender stereotype maintenance. *Archives of Sexual Behavior*, 49(7), 2511–2530.

Wilder, D. A., Simon, A. F., & Faith, M. (1996). Enhancing the impact of counterstereotypic information: Dispositional attributions for deviance. *Journal of Personality and Social Psychology*, 71(2), 276–287.

Zosuls, K. M., Andrews, N. C. Z., Martin, C. L., England, D. E., & Field, R. D. (2016). Developmental changes in the link between gender typicality and peer victimization and exclusion. *Sex Roles*, 75(5–6), 243–256.

Zucker, K. J., Wilson-Smith, D. N., Kurita, J. A., & Stern, A. (1995). Children' s appraisals of sex-typed behavior in their peers. *Sex Roles*, 33(11–12), 703–725.

Zuo, B., Wen, F., & Wu, Y. (2019). Sex differences in mate retention and mate quality enhancement: The effect of facial sexual dimorphism cues on willingness to introduce a new friend to one' s partner. *Archives of Sexual Behavior*, 48(6), 1785–1794.

Note: Figure translations are in progress. See original paper for figures.