HUMBOLDT-UNIVERSITÄT ZU BERLIN

Faculty of Life Sciences
Institute of Psychology

Master Thesis

for the attainment of the academic degree Master of Science

Motivated Responses to a Masculinity Threat in a German Cultural Context

2026-02-09

Submitted by

Widmer, Linus
Matriculation Number: 627542
Email: widmerli@student.hu-berlin.de
Date, Place of Birth: 17.08.1999, Bremen (Germany)

First Examiner: Dr. Ulrich Klocke
Second Examiner: Prof. Dr. Ursula Hess

Abstract

Shame, anxiety, and aggression are common responses by men who feel their masculinity status is threatened (Bosson et al., 2009; Vandello et al., 2008). Recent empirical work suggests that these diverse reactions can be explained by distinct motivational pathways (Stanaland et al., 2023). Although initial evidence supports these associations (Stanaland et al., 2024; Stanaland & Gaither, 2021), the data remain limited and the generalizability of these findings to non-US contexts has not yet been established. The current study tested masculinity threat responses and their motivational pathways using an experimental design in a German cultural context ($N = 196$). It was hypothesized that men in the masculinity-threat condition would exhibit increased aggressive and anxious cognitions, and that these responses would be predicted by pressured and autonomous motivations for masculine behavior, respectively. Contrary to these predictions, aggressive and anxious cognitions did not differ significantly between the experimental groups, nor did pressured or autonomous motivation significantly predict cognitive outcomes. These findings likely reflect limitations in the measurement quality of the cognitive assessments in a German context rather than a true absence of masculinity-threatening effects. Consequently, the results highlight the need for methodologically robust and culturally sensitive instruments to accurately capture the nuances of masculinity threat responses.

Keywords: Fragile Masculinity, Masculinity Threat, Motivation for Masculine Behavior

Contents

1 Introduction

At the core of our being lies our identity: a dynamic, multifaceted self-concept shaped by our personal journeys and the social groups we inhabit. Whether grounded in gender, class, or shared passions, these affiliations influence how we see ourselves, how others perceive us, and the expectations placed upon us. They provide a sense of belonging and purpose, and their norms often serve as guardrails that offer stability among life's unpredictability. Yet, when these guardrails become too narrow, they risk becoming barriers. Meant to protect, they can instead constrain and limit ourselves. Few sets of social norms illustrate this tension as vividly as those surrounding masculinity.

1.1 Masculinity Norms and Their Psychological Cost

Masculinity norms influence the lives of men across their entire lifespan, significantly shaping their behavior and experience. From an early age, boys internalize anti-femininity norms, such as the belief that "boys shouldnt look like girls" (Blakemore, 2003). This early form of gender policing continues into adolescence, where young men report feeling pressured to engage in heterosexual activity (Duckworth & Trautner, 2019). Additional masculinity norms include emotional restriction, self-reliance, the pursuit of achievement and status, aggression, and rejection of homosexuality (Levant et al., 1992). While the content of masculinity norms varies with their cultural background (Lease et al., 2013), these norms consistently shape male behavior and carry significant psychological and social costs. This is illustrated when the narrow range of behaviors that is deemed socially acceptable for men conflicts with a man's personal needs, values, or emotions (O'Neil et al., 1986). For example, attempting to meet masculinity expectations, and potentially failing to do so, is associated with elevated stress (Pleck, 1995) and linked to a range of negative mental health outcomes such as depression and anxiety (Blazina & Watkins Jr, 1996). Moreover, masculinity norms that enforce status-seeking and dominance-related pressures contribute to aggressive behaviors like sexual violence (Rando et al., 1998).

Although the negative consequences of men's gender role conflict are well-documented (O'Neil, 2008), the underlying processes through which this strain develops remain poorly understood. To explain why these costs are so persistent, it is crucial to consider how masculinity functions not just as a set of traits, but as a social status that must be continuously earned and maintained.

1.2 Masculinity as a Precarious Social Status

"Men are made, not born": the idea that masculinity functions as a socially granted status rather than a fixed set of traits has been documented across diverse cultures (Gilmore, 1990). In contrast to womanhood, which is often perceived as a natural, biological outcome, manhood is viewed as a status that must be earned, actively maintained, and socially validated (Vandello & Bosson, 2013). This notion, coined the *precarious manhood thesis*, has been supported by experimental research using gender-identity threats. In these studies, men and women are typically provided with feedback challenging their respective masculinity or femininity. While both groups are subjected to these threats, only men respond with significant psychological and behavioral shifts, such as shame and anxiety (Vandello et al., 2008) or compensatory aggression and dominance (Bosson et al., 2009). These defensive reactions to masculine identity threats have led to the popular characterization of masculinity as *fragile*: the notion that a man's social status as a "real man" is so tenuous that it must be constantly defended and protected (Stanaland et al., 2023). This identity fragility serves as an important link between the psychological strain experienced by the individual and the negative consequences of masculinity norms for broader society. In a series of studies, researchers have found that masculinity threats provoke physical aggression (Bosson et al., 2009), ideological dominance and prejudice toward women (Dahl et al., 2015), denial of social inequities (Weaver & Vescio, 2015), anti-gay prejudice (Brown & Smith, 2023), violence toward gay men (Parrott & Zeichner, 2008), and the tolerance of such violence through inaction (Schermerhorn & Vescio, 2022). These studies highlight the negative

consequences of masculinity norms and emphasize the need for a better understanding of the psychological mechanisms behind this behavior.

A crucial observation in this literature has been the diversity of reactions triggered by masculinity threats. Empirical studies have revealed that both anxiety (Vandello et al., 2008) and aggression (Bosson et al., 2009) are typical responses. These responses are mirrored by the psychological strain experienced in men's gender role conflict, which includes depression and anxiety (Blazina & Watkins Jr, 1996), as well as aggressive behavior (Rando et al., 1998). To address why men respond to threats in such divergent ways, recent research has suggested investigating the underlying motivations for conforming to masculine norms as a potential moderator of these responses.

1.3 A Motivational Framework for Masculinity Threat Effects

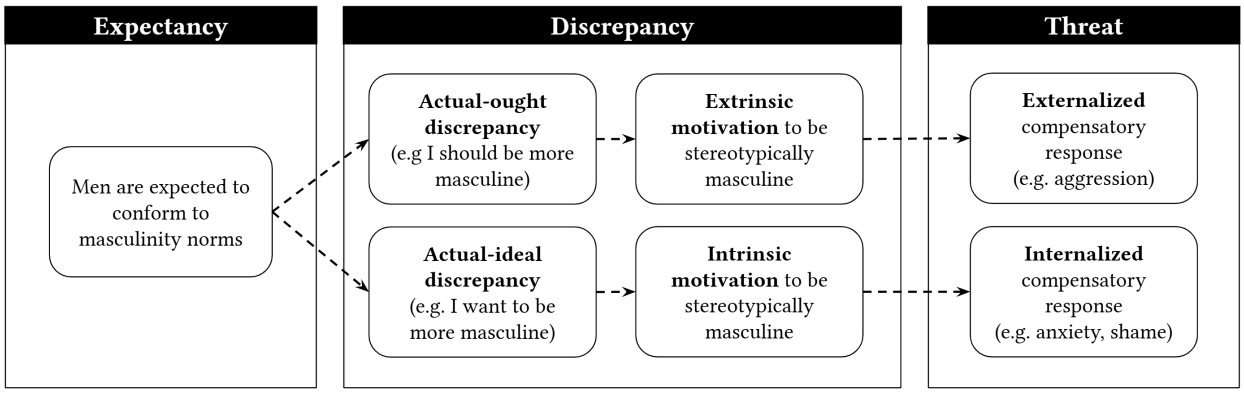
Although stereotypically masculine behaviors may appear uniform on the surface, the motivations behind them can differ significantly. Men may engage in gender-conforming behavior for autonomous reasons, such as when they freely choose and genuinely value such behavior, or for pressured reasons, because they feel compelled by external expectations or fear of rejection (Good & Sanchez, 2010). This distinction between *autonomous* and *pressured motivation* for gender-conforming behavior has recently been applied to research on masculinity threats. In a pioneering study, Stanaland & Gaither (2021) developed the Motivation for Masculine Behavior (MMB) scale to distinguish between these two motivational profiles. Their results demonstrated that men with higher pressured motivation showed significantly more aggressive cognition following a masculinity threat. Crucially, this connection was not explained by men's autonomous motivation, nor by their private regard (i.e., the personal importance they place on being a man). This suggests that the aggressive response is specifically driven by external forces, such as the perceived need to meet social expectations, rather than by intrinsic factors like a personal commitment to masculine values. While Stanaland & Gaither (2021) tested these

motivational predictors for both men and women and induced both masculinity and femininity threats, the aggressive response was only observed and moderated for male participants. This lack of a similar reaction in women further confirms the findings of the precarious manhood thesis, suggesting that the pressure to maintain gender status through compensatory behavior is a uniquely masculine phenomenon.

Building on this evidence, Stanaland et al. (2023) introduced the expectancy-discrepancy-threat (EDT) model of masculine identity (Figure 1) to explain the differences in masculinity threat responses.

Figure 1

The Expectancy-Discrepancy-Threat Model of Masculine Identity



Note. Adapted from Stanaland et al. (2023). Men are socialized to conform to norms of masculinity. When experiencing a masculinity threat, different types of self-discrepancies give rise to distinct motivational processes. An actual-ought discrepancy is thought to trigger extrinsic motivation, often resulting in externalized compensatory behaviors (e.g., aggression). In contrast, an actual-ideal discrepancy is believed to foster intrinsic motivation, leading to an internalized compensatory response (e.g., anxiety).

According to this model, boys and men are socialized into rigid masculinity norms and expectations. Drawing from self-discrepancy theory (Higgins, 1987), the EDT model proposes

that men develop three domains of the self: the actual self (who they believe they are), the ideal self (who they aspire to be), and the ought self (who they feel they should be, based on societal expectations). When men's actual selves do not align with these standards during a masculinity threat, they experience identity discrepancies. Drawing from self-determination theory (Deci & Ryan, 1987), the EDT model suggests that these discrepancies produce distinct motivational responses. An actual-ought discrepancy generates a state of extrinsic motivation to reduce tension, which is associated with externalized responses such as aggression. Conversely, an actual-ideal discrepancy produces a state of intrinsic motivation, which tends to result in internalized reactions such as shame or anxiety. A crucial bridge exists between these situational states and the motivational traits measured by Stanaland & Gaither (2021). While the MMB scale assesses a man's general trait-like tendency to feel pressured or autonomous in his masculinity, the EDT model describes the concrete, state-based motivations that arise in response to a specific discrepancy. The increased aggressive threat response observed by Stanaland & Gaither (2021) can thus be understood through this lens: men with higher pressured motivation as a trait likely experience a greater actual-ought discrepancy when threatened, leading to an externalized, aggressive state. Importantly, men are unlikely to experience only one type of discrepancy; rather, they often experience both forms to different extents, depending on how their underlying motivations interact with a given threat.

1.4 The Present Research

While the EDT model provides a compelling theoretical framework for understanding divergent threat responses, empirical testing of these specific pathways remains in its early stages. The present research addresses this gap by investigating the relationship between men's trait motivations for masculine behavior and their situational responses to masculinity threats. In doing so, this study makes several key contributions to the field.

First, it extends existing research on masculinity threat effects by assessing its cross-cultural generalizability within a German sample, determining if the precarious nature of manhood observed in U.S. contexts holds true in a different cultural landscape. Second, it provides a direct empirical test of the pathways proposed by the EDT model, specifically examining how underlying motivations for conformity shape the psychological and behavioral outcomes of gender-role strain. Third, the present research addresses methodological challenges in the field by exploring implicit measures for assessing masculinity threat responses, which may provide a more nuanced understanding of reactions that are often masked in self-report data.

1.4.1 Cross-Cultural Generalizability of Masculinity Threat Effects

Much of the foundational research on masculinity threat effects and the precariousness of manhood has been conducted in the United States (Bosson et al., 2009; Stanaland & Gaither, 2021; Vandello et al., 2008). This raises concerns about the generalizability of these findings, as the United States —particularly the rural South—is often characterized by an "honor"-based culture that emphasizes the aggressive defense of one's reputation (Cohen et al., 1996). In such contexts, aggressive behavior may serve as a norm-consistent response among men whose masculinity is closely tied to external validation (Jin et al., 2021).

However, the precariousness of manhood appears to be a more universal phenomenon rather than one unique to the United States. This idea was famously proposed by Gilmore (1990), who argued that manhood is a "produced" and "contested" status across nearly all cultures. This anthropological observation has been supported by recent large-scale investigations documenting widespread beliefs in the tenuous nature of masculinity across 62 nations (Bosson et al., 2021). While the intensity of threat responses may scale with broader societal patterns, like gender equality, the underlying precariousness of manhood persists even in egalitarian societies. For instance, men in both Poland and Norway have been found to respond to masculinity threats with increased aggressive cognition (Valved et al., 2021) and reduced support for gender

equality (Kosakowska-Berezecka et al., 2016). This suggests that Germany provides a relevant context for testing these theories, despite its high ranking in gender equality (Rank 7/146; World Economic Forum, 2024). The Norwegian case is particularly instructive: even in a society with a higher equality ranking than Germany (Rank 3/146; World Economic Forum, 2024), masculinity threats still successfully triggered compensatory aggressive cognition (Valved et al., 2021). These findings indicate that structural progress toward gender equality does not necessarily dissolve the psychological association between manhood and social status. Consequently, if manhood remains a precarious status in Germany, a situational threat should activate a masculinity threat effect, characterized by the psychological strain and compensatory pathways described by the EDT model. Accordingly, the following hypotheses are proposed:

- **Hypothesis 1a**: Men exposed to a masculinity threat will exhibit higher levels of aggressive cognition than men who do not receive a masculinity threat.
- **Hypothesis 1b**: Men exposed to a masculinity threat will exhibit higher levels of anxious cognition than men who do not receive a masculinity threat.

1.4.2 Testing Motivational Predictors of Masculinity Threat Effects

Although the EDT model is theoretically compelling, empirical support for its proposed pathways remains limited. Two studies have found that pressured motivation to be masculine moderates aggressive threat responses, with men higher in pressured motivation showing increased aggressive cognition following a masculinity threat (Stanaland et al., 2024; Stanaland & Gaither, 2021). However, the second pathway proposing that autonomous motivation moderates internalized threat responses such as anxiety remains untested.

Moreover, the model faces conceptual challenges from alternative frameworks that propose a sequential relationship between internalized and externalized threat responses. These frameworks suggest that men may initially experience internal distress (e.g., shame or anxiety), which then escalates into outward aggression (Vescio et al., 2025). In this view, externalized

responses serve as compensation for internal distress. This pattern is rooted in research showing that the expression of anger relieves feelings of discomfort (Jakupcak et al., 2005). These competing perspectives highlight the need for careful empirical testing of the EDT model's predictions.

To explain inter-individual differences in men's responses to masculinity threats, the present research tests the EDT model's predictions. According to the model, pressured motivation is associated with externalized responses: men primarily motivated by external expectations are more likely to experience an actual-ought discrepancy when threatened, which activates extrinsic motivation and manifests as aggression. Conversely, autonomous motivation is linked to internalized responses: men who genuinely value masculine ideals may experience an actual-ideal discrepancy when threatened, resulting in anxiety or shame. Accordingly, the following hypotheses are proposed:

- **Hypothesis 2a**: Pressured motivation to conform to masculinity norms increases the positive effect of a masculinity threat (compared to the control condition) on aggressive cognition.
- **Hypothesis 2b**: Autonomous motivation to conform to masculinity norms increases the positive effect of a masculinity threat (compared to the control condition) on anxious cognition.

1.4.3 Addressing Methodological Limitations

Investigating masculinity threat responses presents a significant methodological challenge: the tendency for men to monitor and regulate their outward behavior to align with gender norms. While prior research has frequently relied on self-report or overt behavioral measures, these can be problematic when studying gender-role strain. Internalized emotional responses, such as shame or anxiety, are often incongruent with dominant masculinity standards and may be consciously downplayed or suppressed in explicit reports. Furthermore, aggressive behavioral

responses may sometimes reflect the enactment of socially learned scripts rather than genuine psychological reactivity.

To address these limitations, the present study focuses on cognitive activation, specifically the implicit accessibility of aggressive and anxious concepts, rather than self-reported states. We employ the Word Fragment Completion Task, an implicit measure in which participants complete word fragments that can be resolved as either target words (e.g., related to aggression or anxiety) or neutral alternatives. This task has been successfully utilized to assess aggressive cognition (Stanaland et al., 2024; Stanaland & Gaither, 2021) and anxious cognition (Vandello et al., 2008) following masculinity threats. By measuring the ease with which these concepts come to mind, the Word Fragment Completion Task provides a window into underlying mental processes that are less susceptible to social desirability bias. This approach allows for a more sensitive and direct examination of how masculinity threats are internally processed, capturing reactions that might otherwise be masked by the pressure to adhere to traditional masculine self-presentation.

2 Methods

To examine responses to masculinity threats and their hypothesized moderators, the study employed a between-subjects experimental design. The primary experimental manipulation involved the administration of either a masculinity threat or a non-threatening control condition. The dependent variables consisted of aggressive and anxious cognition, while pressured and autonomous motivation for masculine behavior were examined as continuous moderators of the masculinity threat response.

2.1 Ethical Considerations and Pre-Registration

Prior to data collection, ethical considerations were addressed, particularly regarding the use of a deceptive experimental manipulation. Ethical approval for this study was granted by the Ethics Committee of the Department of Psychology at Humboldt-Universität zu Berlin (reference number: 2025-41). For transparency, the study's aims and methods were pre-registered on the

Open Science Framework (https://osf.io/n6py4). Any deviations from the pre-registration are justified and reported in the relevant sections.

2.2 Sample and Recruitement

To determine the required sample size, a series of a priori power analyses were conducted using G*Power 3.1 (Cunningham & McCrum-Gardner, 2007). First, power analyses for the main effects of masculinity threat were based on effect sizes reported by Vandello et al. (2008) ($d = 0.55$ & $d = 0.91$) and Stanaland & Gaither (2021) ($d = 0.41$). For independent samples t-tests (one-tailed, $\alpha = .05$, power = .80), an estimated medium effect size of $d \approx 0.55$ indicated a required sample size of approximately 84 participants. Second, a power analysis was conducted for the hypothesized moderation effects. An effect size of $f = 0.17$ was assumed for the interaction between threat and motivation, based on the findings by Stanaland et al. (2024). For a linear multiple regression including three predictors (condition, motivation, and their interaction), this analysis indicated that 208 participants were required to achieve 80% power at $\alpha = .05$. Because the moderation analysis represented the most stringent requirement, the final target sample size was based on this calculation. To account for a potential 10% exclusion rate, the target was set at 229 participants.

To achieve a representative sample, participants were initially recruited via flyers in public spaces such as libraries, gyms, and cafes. Due to a limited response rate from these physical locations, 100 participants were recruited through an online study exchange portal (https://surveycircle.com) whose members are primarily young and academic. In exchange for participation, 1€ was donated to a public organization supporting men's counselling and public interests (*Bundesforum Männer e.V.*).

2.3 Procedure

The procedure was adapted from Stanaland & Gaither (2021) into an online questionnaire and translated into German. Participants completed the session on their own electronic devices,

accessing the survey via a link on the study exchange portal or a QR code provided on recruitment flyers. Upon accessing the survey, participants were informed that the study concerned men's personality, motivation, and performance. Eligibility requirements—identifying as male, possessing native-level proficiency in German, and being at least 18 years of age—were clearly stated, and participants self-confirmed their eligibility and provided informed consent before proceeding.

The session began with the assessment of trait motivations using the MMB scale, where the nine items were presented in a randomized order to prevent sequence effects. Following this, participants completed the Gender Knowledge Test. This test was described as a validated measure of gender-relevant knowledge, designed to provide credible feedback on the participants' masculinity and to facilitate the experimental manipulation. The test items were presented in a fixed, pre-randomized sequence. Upon completion, a brief loading screen was displayed to simulate the real-time calculation of results, thereby increasing the credibility of the subsequent feedback. Participants then received fabricated results that constituted either the masculinity threat or the control condition. To assess the immediate impact of this feedback on cognitive activation, participants completed the Word Fragment Completion Task to measure the accessibility of aggressive and anxious concepts. In the final stage of the session, participants provided demographic information and responded to manipulation check items. Upon completion, a comprehensive debriefing was provided, explaining the study's true purpose and clarifying that the feedback regarding test performance was fabricated and randomly assigned. The rationale for the use of deception was explicitly detailed, and participants were offered the final option to withdraw their data from the study.

2.4 Measures

All translations followed the Translation, Review, Adjudication, Pretesting, and Documentation protocol (Harkness et al., 2010). Complete materials including all item wordings in

German and English, documentation of the translation process and additional statistics are provided in Appendix A.

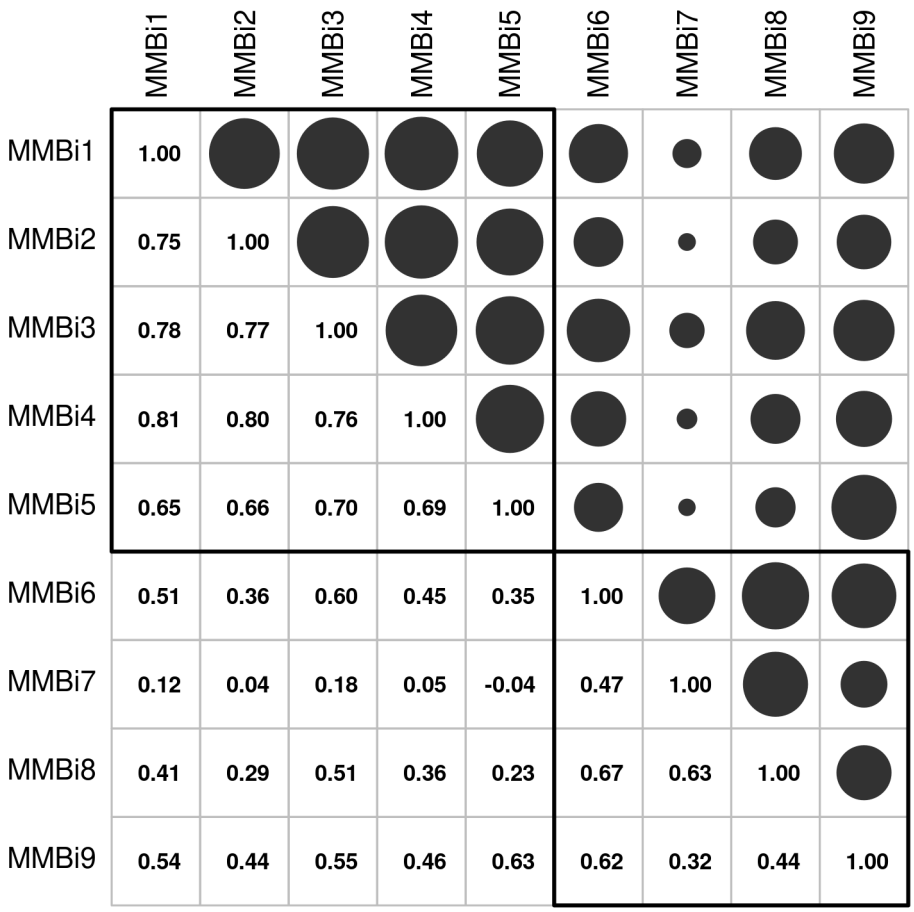
2.4.1 Motivation for Masculine Behavior Scale

The MMB scale was adapted from Stanaland & Gaither (2021) and translated into German. The scale originally consisted of nine items designed to measure participants' underlying motivations for enacting masculine behavior. Responses were recorded on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). To minimize participant suspicion regarding the scale's purpose, the items were presented under the heading "Attitudes Toward Masculinity".

The scale is theorized to comprise two subscales: Pressured Motivation (Items 1–5), reflecting extrinsic social pressure, and Autonomous Motivation (Items 6–9), reflecting intrinsic enjoyment and personal importance. To validate this, the item structure was first assessed via Confirmatory Factor Analysis (CFA). Model fit was evaluated against the criteria established by Hu & Bentler (1999): CFI and TLI $> .95$, RMSEA $< .06$, and SRMR $< .08$. The initial CFA for the hypothesized two-factor model demonstrated poor fit (CFI = .85, TLI = .8, RMSEA = .18, SRMR = .25). Diagnostic inspection of the polychoric correlation matrix (Figure 2) and a subsequent Exploratory Factor Analysis with correlated factors (Table 1) identified the sources of this poor fit. While Items 1–5 showed strong coherence on the Pressured Motivation factor, the Autonomous Motivation factor was less distinct. Notably, Item 9 ("It is important to me not to be feminine") exhibited significant cross-loadings on both factors.

Figure 2

Polychoric Correlation Matrix for Motivation for Masculine Behavior Scale



Note. MMB = Motivation for Masculine Behavior. Rectangles indicate the hypothesized factor structure: Items 1–5 = Pressured Motivation subscale; Items 6–9 = Autonomous Motivation subscale. Larger circles indicate stronger positive correlations.

Table 1

Factor Loadings on the Motivation for Masculine Behavior Scale

Item	Description	Pressured Motivation	Autonomous Motivation
1	In general, I'm masculine because I want others' acceptance and approval.	.83	

Item	Description	Pressured Motivation	Autonomous Motivation
2	In general, I'm masculine because that is what people expect from me.	.90	
3	I'm masculine because I want people to like me.	.79	
4	I'm masculine around other people because that is how others think I should be.	.89	
5	I'm not feminine because people wouldn't like me.	.78	
6	It's important to me to be masculine.		.69
7	I enjoy being masculine.		.74
8	It makes me happy if I'm masculine.		.80
9	It is important to me not to be feminine.	.39	.39

Note. English item versions of MMB Scale. Factor loadings below .30 are suppressed for clarity. Extraction method: Maximum Likelihood with oblimin rotation.

Based on these results, Item 9 was removed. Furthermore, modification indices suggested a residual covariance between Items 7 and 8. This adjustment is conceptually grounded, as both items express positive affect toward masculine identity ("enjoy," "happy") rather than distinct motivational drivers. A final CFA on the refined 8-item scale incorporating this correlated residual yielded a substantially improved fit (CFI = .97, TLI = .95, RMSEA = .1, SRMR = .07). While the TLI and RMSEA remained slightly outside ideal thresholds, the CFI and SRMR satisfied target values. No further modifications were pursued to avoid overfitting and maintain theoretical integrity. Subscale scores were computed as means using the refined structure: Pressured Motivation (Items 1–5) and Autonomous Motivation (Items 6–8). The two subscales showed a moderate positive correlation ($r = .39$). Reliability was assessed using McDonald's omega (ω), indicating excellent reliability for the Pressured Motivation subscale ($\omega = .93$) and acceptable reliability for the Autonomous Motivation subscale ($\omega = .70$).

2.4.2 Masculinity Threat Induction

The masculinity threat was administered by requiring participants to complete a test supposedly measuring their knowledge of gender-specific topics and subsequently providing them with feedback regarding their performance. The task utilized a culturally adapted version of the Gender Knowledge Test, originally developed by Rudman & Fairchild (2004) and updated for European participants by Valved et al. (2021). This test consisted of thirty forced-choice items balanced across stereotypically masculine domains, such as sports and home repair, and feminine domains, such as childcare and fashion. To ensure the credibility of the subsequent bogus feedback, the items were designed with moderate to high difficulty so that participants could not easily estimate their own performance. Each item featured a forced-choice format with one correct and one incorrect response option (e.g., "The paste used for soldering joints is called: gel vs. flux"; "The first company to develop hair coloring was: Clairol vs. L'Oréal").

Following the test, participants were shown a diagram depicting a fictitious distribution of scores to anchor the manipulation. The scale ranged from a maximum feminine score of −4 to a maximum masculine score of +4. In the threat condition, participants were informed they had scored −1.83, a value positioned substantially below the alleged male average of +2.24. To reinforce the psychological impact of the threat, this score was accompanied by an explicit statement that their responses were more similar to women than to men. Conversely, participants in the control condition were told they scored +2.18, placing them close to the male average and indicating a typical performance for male participants. In both conditions, the visual distribution graph served to anchor the feedback by highlighting the participant's position relative to the alleged male norm.

2.4.3 Word Fragment Completion Task

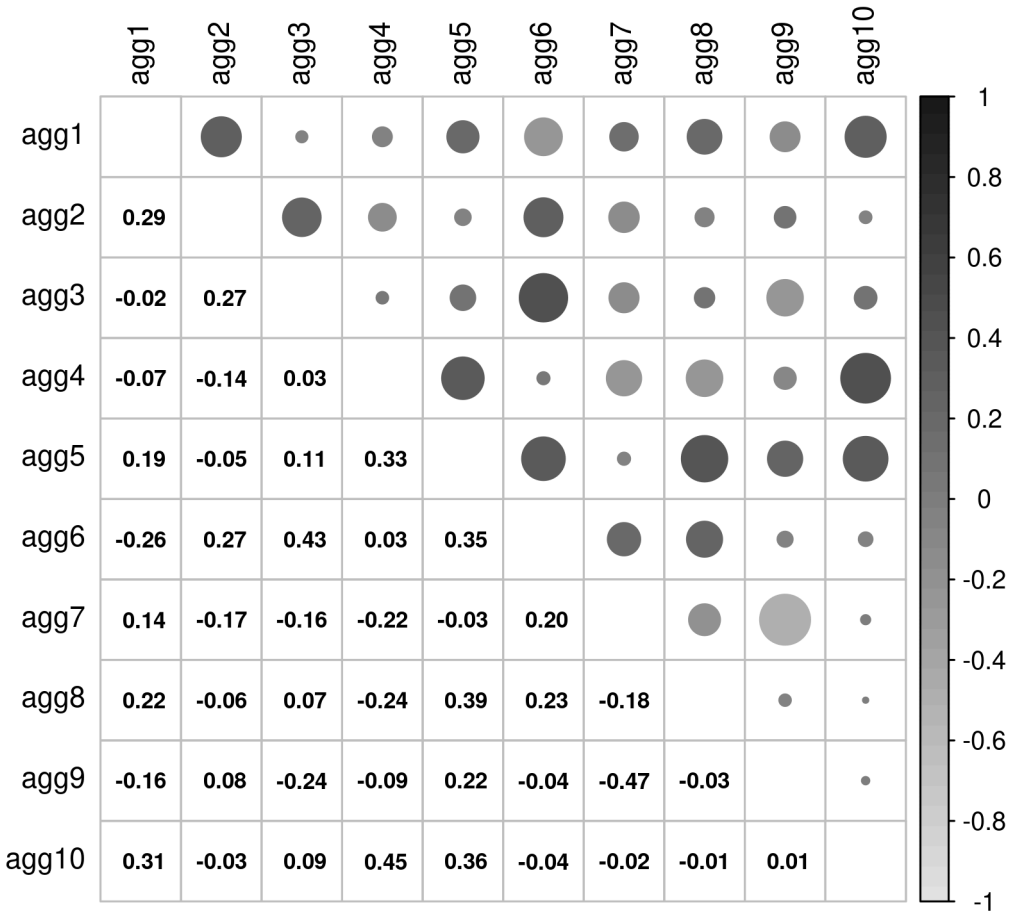
The Word Fragment Completion Task was first developed by Vandello et al. (2008) to implicitly assess aggressive cognition. For the current study, a German version was developed consisting of ten aggressive and ten anxious word fragments. These fragments were created by

identifying German target words semantically related to aggression (e.g., “Wut” [anger]) and anxiety (e.g., “Bangen” [to fear]) using a word association thesaurus (https://wordassociations.net/de/). Potential fragments were generated by removing the first letter of each target word, ensuring each could be completed as either the target word or a neutral alternative (e.g., “_UT” becoming “WUT” [anger] or “HUT” [hat]). To mask the task’s true purpose, it was presented as a measure of “speed of thought.” Word fragments were displayed individually with a seven-second time limit to encourage spontaneous, intuitive responses.

Following data collection, the psychometric properties of the Word Fragment Completion Task were evaluated to assess the reliability of these newly developed items. Diagnostic analyses revealed substantial deficits in the task’s measurement quality. The tetrachoric correlation matrices for both the aggressive (Figure 3) and anxious (Figure 4) subscales showed weak, inconsistent, and occasionally negative correlations among items. For the aggressive items, positive completion rates varied substantially (3.3% to 63.8%; see Table A9), and the Kaiser-Meyer-Olkin (KMO) criterion yielded a Measure of Sampling Adequacy (MSA) of .16—well below the acceptable .60 threshold. Internal consistency was similarly poor ($\alpha = .36$). The anxious items exhibited even greater instability; positive completion rates ranged from 3.4% to 85.7% (Table A11), and the subscale reached an MSA of only .30 with an unacceptable Cronbach’s alpha ($\alpha = .09$).

Figure 3

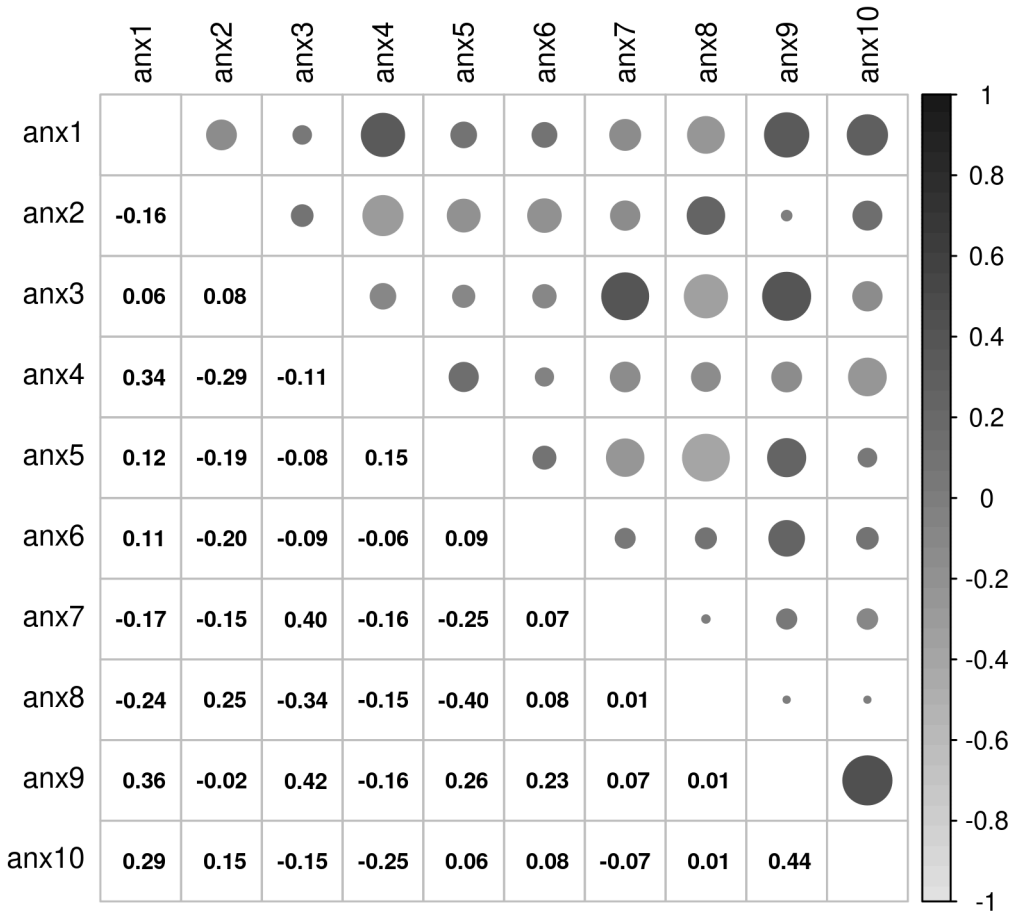
Tetrachoric Correlation Matrix for Aggressive Word Fragment Items



Note. Darker shading indicates stronger positive correlations. Item descriptives are provided in Table A9.

Figure 4

Tetrachoric Correlation Matrix for Anxious Word Fragment Items



Note. Darker shading indicates stronger positive correlations. Item descriptives are provided in Table A11.

2.4.4 Demographics, Manipulation Checks and Suspicion Probe

Demographic data were collected to provide an overview of the sample's characteristics. Participants reported their age, gender, and German proficiency, followed by their highest educational and occupational qualifications. The effectiveness of the experimental manipulation was verified using two check items. First, participants recalled the type of feedback they received on the Gender Knowledge Test on a scale from 1 (*typically feminine*) to 10 (*typically masculine*). Results confirmed that participants correctly interpreted the masculinity feedback, as participants in the threat condition reported lower scores (M = 3.1, SD = 0.9) than those in the

control condition (*M* = 6.2, *SD* = 1). Second, they provided a self-evaluation of their own gender knowledge on the same scale. Consistent with the intended effect, the masculinity feedback was perceived as credible. Participants in the threat condition reported lower scores (*M* = 5.7, *SD* = 1.4) than those in the control condition (*M* = 6.8, *SD* = 1.4). An independent samples *t*-test confirmed this difference was significant, *t*(154) = 4.74, *p* = < .001. To assess suspicion, participants indicated whether they believed the study concerned something other than what was stated. If they responded affirmatively, they were prompted to describe their suspicions. Finally, an open-text field allowed participants to share additional thoughts. Nine participants correctly identified the experimental manipulation of the feedback and were flagged for exclusion during the data cleaning process.

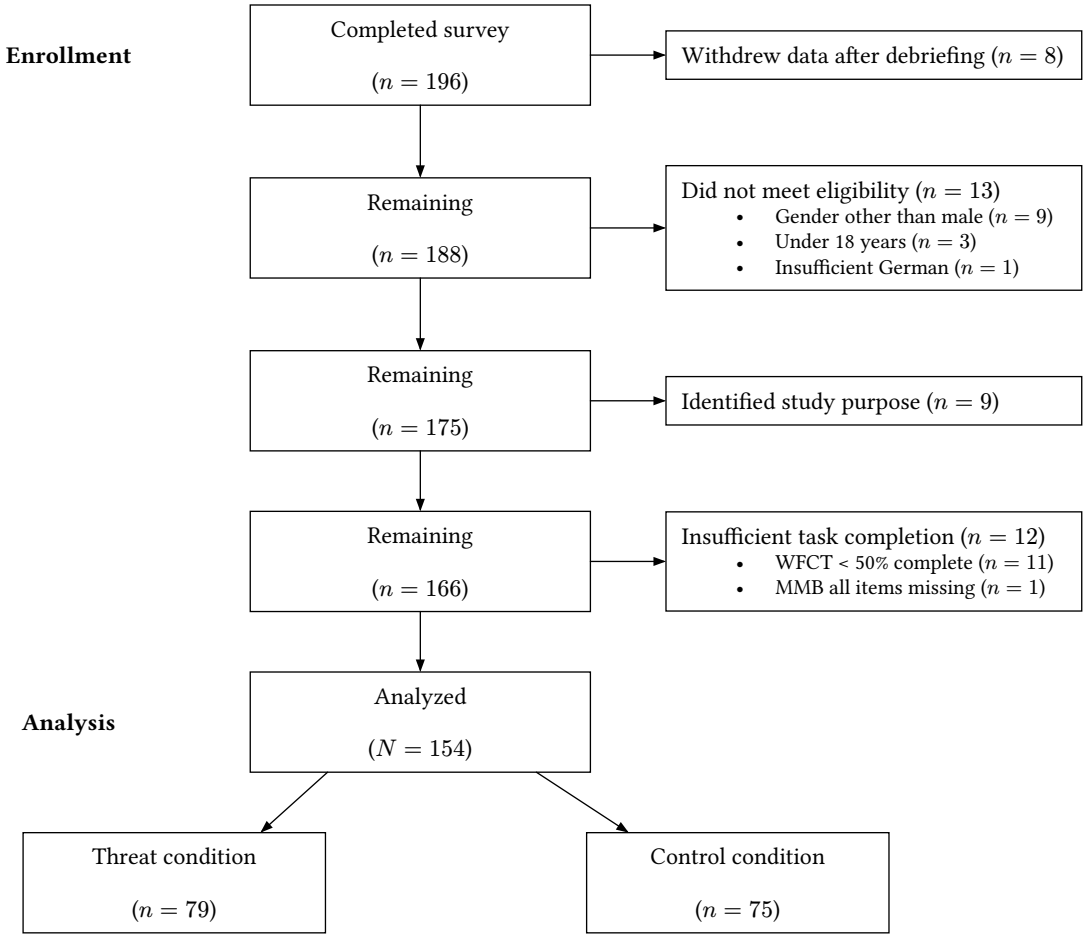
3 Results

3.1 Participants

Data collection was conducted over a two-month period, resulting in an initial sample of 196 participants. Although this fell short of the original target ($N = 229$), collection was concluded due to project time constraints. To ensure data quality, several participants were removed based on pre-registered exclusion criteria. As illustrated in the CONSORT flowchart (Figure 5), exclusions were made for withdrawal of consent, failure to meet eligibility criteria, failed attention checks, and identified suspicion regarding the manipulation. After applying these criteria, 42 participants were removed, yielding a final analytical sample of $N = 154$ ($n_{\text{threat}} = 79$, $n_{\text{control}} = 75$).

Figure 5

CONSORT Flowchart of Participant Exclusions



Note. MMB = Motivation for Masculine Behavior; WFCT = Word Fragment Completion Task

Based on this final sample size, a post-hoc power analysis indicated that the study achieved a power of $1 - \beta = .96$ to detect the expected medium effect ($d = 0.55$) of the masculinity threat. However, for the hypothesized moderation, assuming a small effect size ($f = 0.17$), the power was considerably lower at $1 - \beta = .56$.

The final sample ranged in age from 18 to 69 years (*M* = 30.6, *SD* = 10.9). Participants were predominantly highly educated, with 80.5% holdingat least a general university entrance qualification (*Abitur*), and 60.4% having completed a university degree.

3.2 Statistical Analysis

Data analysis was performed using R version 4.5.2 (R Core Team, 2023). An anonymized dataset and the corresponding analysis code have been made publicly available via GitHub (https://github.com/Linuswidmer/master-thesis).

Hypothesis testing was conducted on the proportion of aggressive and anxious completed word fragments for each participant. The pre-registered analysis plan specified an arcsine square root transformation to address the bounded nature of proportion data (Winer et al., 1971). However, preliminary diagnostics revealed that the distribution of transformed scores was highly skewed, violating the normality assumption required for standard regression. To address this, the analyses employed a generalized linear model (GLM) with a binomial distribution and logit link function. This approach directly models the count of target completions out of total trials, appropriately accounting for the bounded and discrete nature of the outcome without requiring transformation.

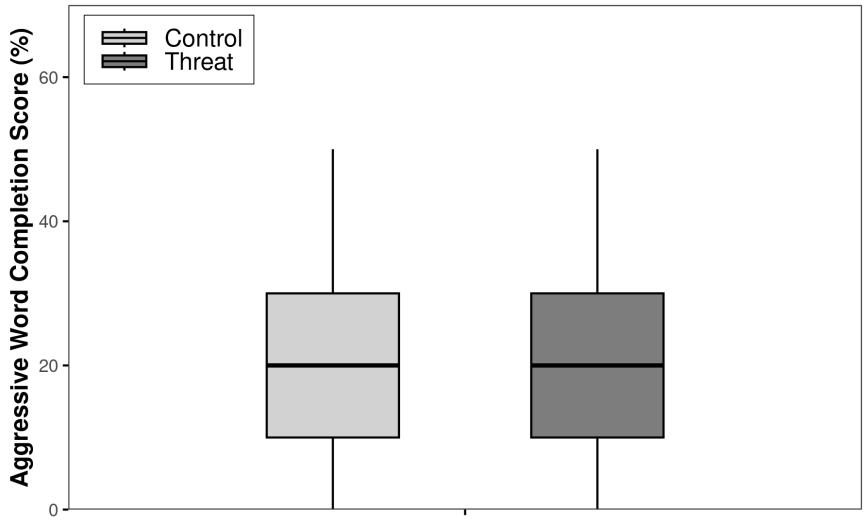
To ensure robustness, several diagnostic checks were performed. Across all models, homogeneity of variance (Levene's test) and residual distribution using DHARMa (Hartig, 2024) were verified, with no significant violations detected. To identify influential observations that could distort the results (Aguinis et al., 2013), a sensitivity analysis was conducted using a Cook's distance threshold of $\frac{4}{n-k}$, where n represents the sample size and k the number of predictors. These cases were removed and the models refitted to ensure the stability and robustness of the primary findings. In all cases, refitting the models without these observations did not alter the significance of the results; therefore, only the results for the full analytical sample are reported.

3.2.1 Hypothesis 1a: Masculinity Threat and Aggressive Cognition

Hypothesis 1a predicted higher aggressive cognition in the threat condition compared to the control condition. Contrary to this prediction, there was no significant difference of aggressive word completion scores in the threat condition (M = 19.5, SD = 12.3) and the control condition (M = 22.8, SD = 12.3; see Figure 6).

Figure 6

Aggressive Word Completion Scores by Threat Condition



Note. Boxplots display the distribution of aggressive word completion scores (%) by experimental condition. The horizontal line represents the median, boxes represent the interquartile range (25% - 75%), and whiskers extend to 1.5 times the interquartile range.

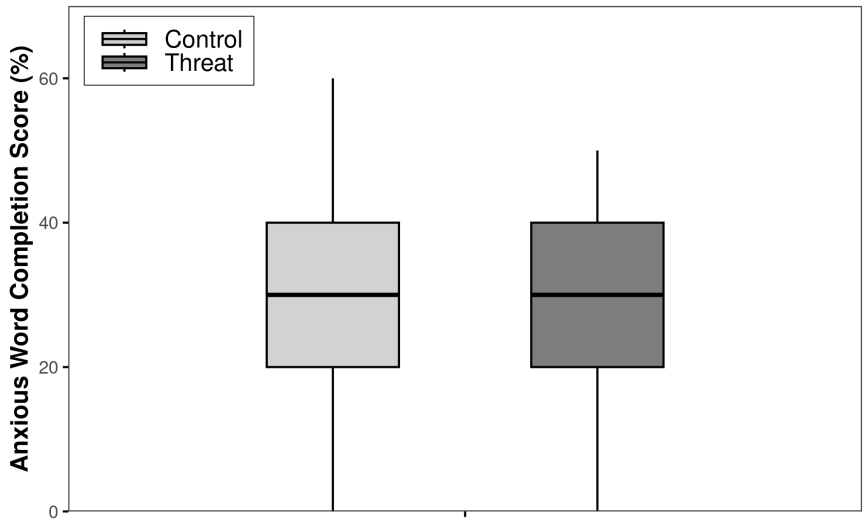
A binomial GLM with threat condition as the predictor and aggressive word completions as the dependent variable revealed no statistically significant effect of condition ($z = -1.59$, $p = .111$).

3.2.2 Hypothesis 1b: Masculinity Threat and Anxious Cognition

Hypothesis 1b predicted higher anxious cognition in the threat condition compared to the control condition. Anxious word completion scores were virtually identical across conditions: threat condition ($M = 28$, $SD = 12.8$) and control condition ($M = 27.6$, $SD = 14.5$; see Figure 7).

Figure 7

Anxious Word Completion Scores by Threat Condition



Note. Boxplots display the distribution of anxious word completion scores (%) by experimental condition. The horizontal line represents the median, boxes represent the interquartile range (25% - 75%), and whiskers extend to 1.5 times the interquartile range.

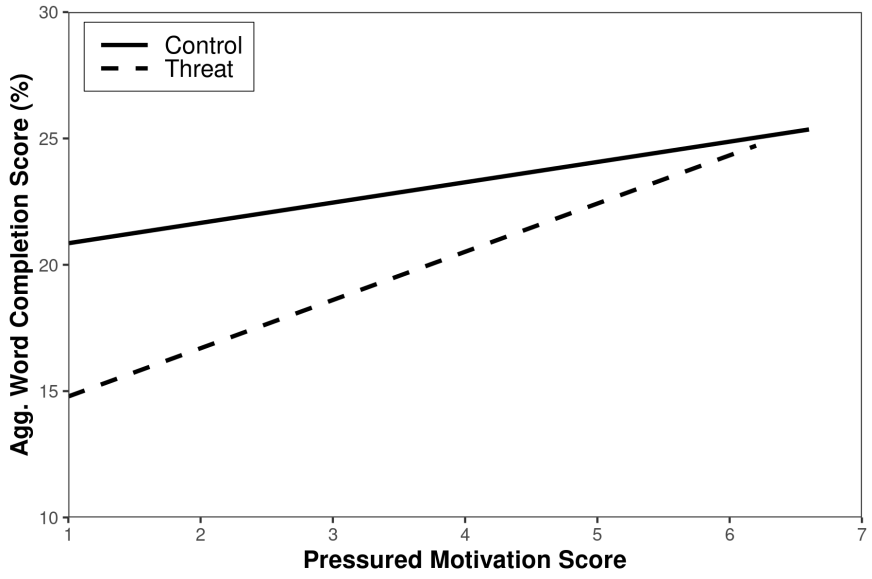
A binomial GLM with threat condition as the predictor and anxious word completions as the dependent variable revealed no statistically significant effect of condition ($z = 0.18$, $p = .859$).

3.2.3 Hypothesis 2a: Pressured Motivation and Aggressive Cognition

Hypothesis 2a predicted that the effect of masculinity threat on aggressive cognition would be moderated by pressured motivation to conform to masculinity norms. Specifically, men with higher pressured motivation were expected to show a stronger increase in aggressive cognition following threat compared to men with lower pressured motivation. Aggressive cognition scores as a function of pressured motivation and experimental condition are shown in Figure 8.

Figure 8

Moderation of Aggressive Word Completion Scores by Pressured Motivation



Note. Aggressive word completion scores (%) by threat condition and pressured motivation

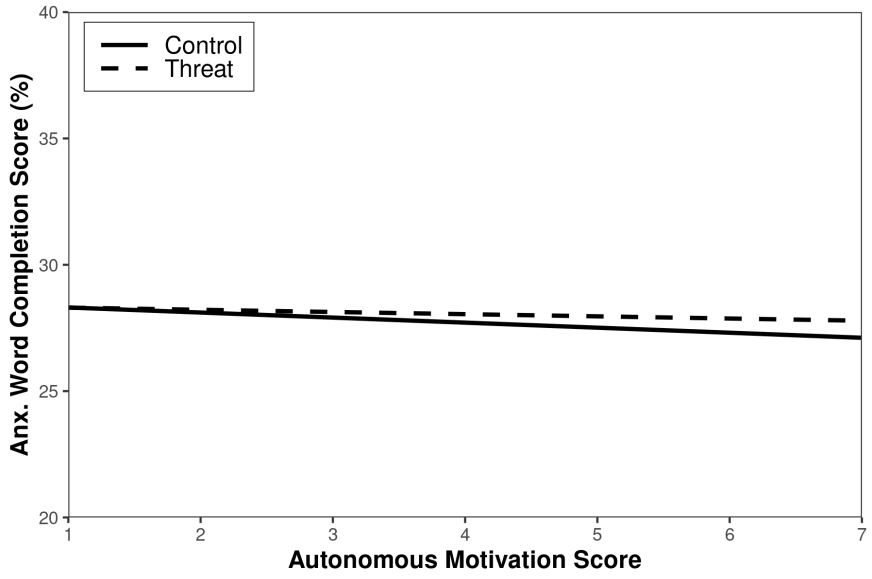
A binomial GLM with aggressive word completions as the dependent variable and threat condition, pressured motivation, and their interaction as predictors was conducted to test this moderation hypothesis. The model revealed no statistically significant main effect of threat condition (z = −1.47, p = .141), no significant main effect of pressured motivation (z = 0.84, p = .403), and no significant interaction effect (z = 0.92, p = .357).

3.2.4 Hypothesis 2b: Autonomous Motivation and Anxious Cognition

Hypothesis 2b predicted that the effect of masculinity threat on anxious cognition would be moderated by autonomous motivation to conform to masculinity norms. Specifically, men with higher autonomous motivation were expected to show a stronger increase in anxious cognition following threat compared to men with lower autonomous motivation. Anxious cognition scores as a function of Autonomous Motivation and experimental condition are shown in Figure 9.

Figure 9

Moderation of Anxious Word Completion Scores by Autonomous Motivation



Note. Anxious word completion scores (%) by threat condition and autonomous motivation.

A binomial GLM with anxious word completions as the dependent variable and threat condition, autonomous motivation, and their interaction as predictors was conducted to test this moderation hypothesis. The model revealed no statistically significant main effect of threat condition ($z = -0.02$, $p = .988$), no significant main effect of autonomous motivation ($z = -0.18$, $p = .855$), and no significant interaction effect ($z = 0.07$, $p = .946$).

4 Discussion

The present study aimed to replicate masculinity threat effects and examine their motivational moderators in a German sample. Contrary to the hypotheses, the results did not reveal the predicted increases in aggressive or anxious cognition following a masculinity threat. Furthermore these relationships were not moderated by pressured or autonomous motivation to behave masculine. While these results contrast with previous findings, the following section examines various factors that may have contributed to the absence of these effects.

4.1 Absence of Masculinity Threat Effects

4.1.1 Experimental Manipulation

The experimental manipulation of masculinity threat appeared to be robust. Manipulation checks confirmed that participants correctly recalled the performance feedback from the Gender Knowledge Test, and those in the threat condition reported a significant decrease in their self-evaluated masculinity compared to the control group. This indicates that the study successfully induced a cognitive appraisal of threat to masculine identity. Consequently, the absence of hypothesized effects is likely not attributable to a failure of the experimental induction itself.

4.1.2 Sample Considerations

The study's ability to detect the consequences of the masculinity threat was influenced by the discrepancy between the planned and achieved sample size. While the research design targeted a final analytical sample of 208 participants to ensure sufficient power for all hypothesized effects, the final sample available for analysis consisted of 154 participants. This shortfall was the result of recruitment reaching 196 participants due to fixed time constraints, coupled with a substantial attrition rate of approximately 22% during the data cleaning process. The high exclusion rate may be attributable to the academic composition of the sample. Research-savvy populations are often more prone to suspicion, and the online format of the study likely contributed to reduced levels of participant engagement. Consequently, the final analytical sample was approximately 26% smaller than the intended target of 208. Nevertheless, post-hoc power analysis indicated that the study still maintained high power ($1 - \beta = .96$) to detect the hypothesized medium-sized effect ($d = 0.55$) of the masculinity threat. While the power was indeed insufficient to detect the smaller moderation effects ($1 - \beta = .56$), the robust power for the main effect suggests that the absence of a primary masculinity threat effect is likely not attributable to the diminished sample size alone.

4.1.3 Issues with Dependent Measures

The most plausible explanation for the absence of significant effects concerns the poor measurement quality of the Word Fragment Completion Task. While the manipulation success-

fully induced a threat, the instrument used to capture the resulting cognitive state proved unreliable, as evidenced by extremely poor internal consistency (α as low as .09) and inadequate sampling adequacy (KMO $<$ .30). This lack of reliability likely stems from the inherent difficulty of designing a Word Fragment Completion Task that isolates specific cognitive activation from "lexical noise." In any language, a participant's completion of a fragment is influenced by a multitude of factors unrelated to the experimental threat, such as individual vocabulary size, word frequency, and "orthographic neighborhood density"—the number of alternative words that can be formed from a single fragment.

While it is theoretically possible to construct a valid Word Fragment Completion Task in German, doing so requires balancing these complex lexical variables to ensure that the target word is neither too obvious nor too obscure. In the present study, the high level of measurement error suggests that these uncontrolled factors may have overwhelmed the subtle signal of threat-induced cognitive activation. This mirrors the findings of Kugler et al. (2022), who demonstrated that while German Word Fragment Completion Task can track stable trait aggression, they struggle to capture the acute, state-level shifts in cognition required for experimental research.

4.2 Measuring Masculinity Threat Effects

The psychometric failure of the Word Fragment Completion Task highlights a broader challenge in masculinity research: the difficulty of selecting a measure that effectively balances social desirability bias, construct validity, and the ability to differentiate between distinct threat responses. This necessitates a careful weighing of alternative formats that might better optimize the trade-off between sensitivity and reliability in future studies. The following sections provide an overview of the various measures that can be administered to capture masculinity threat responses.

4.2.1 Self-Report & Attitudinal Measures

Explicit self-report measures of emotional states offer high face validity and are well-established in psychological research as a primary method for capturing subjective experience (e.g. Robinson & Clore, 2002). Several studies have successfully assessed masculinity threat responses through self-reported emotions such as anger, shame, pride, and public discomfort (Valved et al., 2021; Vescio et al., 2025). However, this approach faces two significant limitations. First, it is highly susceptible to social desirability bias: reporting feelings of anxiety or shame may itself be perceived as unmasculine, leading men to underreport these responses. This concern is particularly acute in this context, as masculine identity threat is tied to concerns about appearing insufficiently masculine. Second, self-reports rely on the assumption that individuals possess the necessary introspective awareness to accurately identify their internal states.

An alternative approach involves assessing threat responses through attitudinal outcomes. Masculinity threats have been shown to increase the justification of social inequality (Weaver & Vescio, 2015) or tolerance for sexual violence (Vescio et al., 2025). These measures offer a balance between practical feasibility and construct validity, as attitudes can be assessed through standard questionnaire formats while remaining less transparent than direct emotion reports. However, attitudinal measures may be particularly problematic in highly educated samples, where baseline support of discriminatory attitudes tends to be low. Furthermore, attitudinal measures fail to distinguish between internalized and externalized responses. Because a single reported attitude can stem from either defensive aggression or an underlying state of anxiety, these measures capture the evaluative "output" of a threat rather than the specific cognitive or affective process that triggered it.

4.2.2 Implicit Measures and Associative Tasks

Implicit measures are typically chosen to bypass the social desirability bias and the lack of introspective awareness often found in self-reports. By measuring automatic cognitions, these tasks can theoretically detect genuine threat effects without relying on a participant's willingness

or ability to report them (Vandello et al., 2008). The Word Fragment Completion Task was selected for this study to capture these processes. However, it failed to serve as a reliable measure, highlighting the difficulty of constructing a task sensitive enough to capture experimental effects over the "noise" of individual linguistic differences.

To retain the advantages of implicit testing while avoiding the lexical noise of word fragments, researchers might instead employ associative tasks like the Implicit Association Test (IAT) (Egloff & Schmukle, 2002). Rather than relying on the spontaneous completion of fragments, the IAT measures the relative strength of associations between concepts (e.g., "Self" and "Weakness"). While this format is less influenced by individual vocabulary differences, its validity for capturing acute, state-level shifts following an experimental manipulation is highly questionable. Research has found that implicit formats often fail to register significant changes following experimental inductions (Schmukle & Egloff, 2004). In the specific context of masculinity threat, the extensive length of the IAT presents a critical procedural limitation. The cognitive effort and time required to complete the numerous trials likely allow the short-lived psychological state triggered by the threat to resolve before it can be accurately measured.

4.2.3 Behavioral Measures

Behavioral measures provide perhaps the most direct assessment of threat-relevant outcomes by capturing real-world actions rather than reported feelings. Research has demonstrated masculinity threat effects through proxies such as the voluntary administration of electrical shocks (Fowler & Geers, 2017). This measure effectively bypasses conscious filtration, as participants are often unaware that their actions serve as a defensive response to a psychological threat. While traditional behavioral approaches are often resource-intensive, future research could utilize game-like paradigms to measure masculinity threat responses in a more scalable, online-compatible format.

The primary advantage of these paradigms is their ability to capture a wide range of threat responses simultaneously. For example, by analyzing in-game decision-making, researchers can observe both internalized responses (such as increased caution or avoidance) and externalized responses (such as heightened aggression or risk-taking) within the same task. While these digital environments would introduce their own forms of noise including varying levels of gaming experience, they offer a more naturalistic setting for observing identity-relevant behavior than word-based tasks.

Beyond the task design itself, this format may help address the widespread reliance on "WEIRD" (Western, Educated, Industrialized, Rich, and Democratic) samples in psychological research. Game-based measures likely appeal to a broader and more diverse societal demographic than traditional academic surveys. This shift is of particular interest as these populations may exhibit different motivational profiles, including both autonomous and pressured motivations for masculine behavior. Such an approach could provide a more representative view of masculinity threat responses than the predominantly academic samples that are currently overrepresented in the literature. Finally, because the moderation effects proposed by the EDT model are presumably small, moving toward these scalable behavioral formats would allow researchers to achieve the larger sample sizes necessary to reach adequate statistical power for complex interaction effects.

4.2.4 Recommendations for Future Research

To improve the detection of masculinity threat effects, future research should prioritize validated instruments with established psychometric properties. While developing new tasks is often necessary, these must undergo rigorous validation to ensure they capture the intended constructs rather than external noise. A successful design must also include clear manipulation checks to confirm the threat was perceived.

In practice, explicit measures often provide higher state-sensitivity than implicit tasks and should be prioritized, provided they are refined to account for social desirability. When the goal is

to distinguish between the internalized and externalized pathways proposed by the EDT model, the chosen measures must be capable of clearly separating these distinct responses. Finally, the methodology must match the scale of the research. Because the expected moderation effects are presumably small, researchers should utilize scalable online formats to achieve the large sample sizes necessary for adequate statistical power.

4.3 Validation of the German Motivation for Masculine Behavior Scale

While the study did not observe the hypothesized moderation effects, it successfully established a psychometrically sound German translation of the MMB scale. Unlike the Word Fragment Completion Task, the MMB scale demonstrated adequate variability even within a highly educated sample. This suggests that the participants' academic background did not lead to a conscious rejection of traditional masculinity norms, which would have otherwise resulted in "floor effects." However, the validation process revealed three specific challenges that required careful refinement of the scale's factorial and linguistic structure.

The first challenge concerned the relationship between the two latent factors. While the original model by Stanaland et al. (2023) suggested they might be orthogonal, the German data required allowing Pressured and Autonomous Motivation to correlate to achieve an acceptable fit. This shift from an orthogonal to a correlated model provides a theoretical solution that aligns with the foundational work on motivation for gendered behaviour (Good & Sanchez, 2010). A man's personal ideals (autonomous) do not develop in a vacuum but in constant dialogue with societal expectations (pressured). The internalization of these norms implies an inherent overlap rather than a total separation of motivational sources. The second challenge involved the poor discriminant validity of Item 9 ("It is important to me not to be feminine"). While this item was theoretically designed to load on the Autonomous Motivation factor, it exhibited significant cross-loading onto the Pressured Motivation factor. This pattern mirrors the observations of Stanaland & Gaither (2021) in the original validation and suggests that the "rejection

of femininity" may be a more complex construct. Rather than reflecting a purely autonomous preference, anti-femininity attitudes likely represent a conceptual overlap, driven by a combination of internalized values and perceived external pressures. The final challenge involved the translation of the scale, specifically regarding how to capture the degree to which a person strives for masculinity. In German, a literal translation of "being masculine" ("männlich sein") can be ambiguous. It may be interpreted by participants as a static biological fact rather than a personal value or goal. To address this, the translation employed a strategic mix of trait-based formulations and agentic, behavioral expressions such as "sich männlich verhalten" (behaving masculinely). This balanced approach ensured the items were not seen as mere descriptions of biological status, but rather as reflections of how much an individual values and actively pursues a masculine identity.

These refinements provide a clear roadmap for future research. By employing a correlated two-factor model and a refined item set, researchers can now reliably assess the motivational drivers of masculine behavior in German-speaking populations.

4.4 Theoretical Considerations

While the psychometric limitations of the Word Fragment Completion Task provide a compelling explanation for the null findings, a purely methodological critique may be incomplete. Recent empirical developments in the field offer alternative perspectives on the precarious manhood thesis, suggesting that masculinity threat responses may be more sensitive to specific social and situational contexts than previously assumed. Exploring these broader theoretical directions may provide a more nuanced understanding of why certain threat effects remain elusive and how future research can better capture the conditions under which they emerge.

4.4.1 Expanding the Perspective of Motivational Pathways

One consideration is the potential interplay between the proposed pathways. While the EDT model distinguishes between internalized and externalized responses based on motivation,

these paths may not always function in parallel. Alternative frameworks suggest a more sequential relationship, where a threat first triggers internal distress—such as shame or anxiety—which then motivates externalized aggression as a compensatory mechanism to restore status (Vescio et al., 2025). If aggression acts as a functional relief for an underlying anxious state (Jakupcak et al., 2005), the two pathways are fundamentally linked. This dynamic process would make it difficult for a single measurement point to isolate them as independent outcomes of different motivations.

Beyond the relationship between these pathways, recent evidence suggests that the initial affective experience of gender identity threat may be more symmetric across genders than originally theorized. Wittlin et al. (2024) demonstrated that femininity threats regarding physical appearance evoke significant anxiety and self-esteem deficits in women, mirroring the responses previously observed in men. The fact that earlier studies, such as Bosson et al. (2009), failed to find this effect might be due to several factors. One possibility is that being told one is "similar to men" may not have been appraised as a significant status loss by all female participants. In certain contemporary social or professional contexts, masculine-coded traits like assertiveness are increasingly viewed as desirable, potentially buffering the intended threat of the feedback. However, even when a valid threat is administered and triggers distress, a conceptual tension remains for the EDT model when applied to women. If women primarily respond with internalized anxiety, the model would imply a relative lack of the pressured motivation that typically drives externalized aggression. This appears to stand in contrast to research by Stanaland & Gaither (2021) and Good & Sanchez (2010), who have reliably identified pressured motivation for gendered behavior in women.

This discrepancy suggests that the transition from a perceived "threat" to a specific behavioral response may be influenced by individual control beliefs. According to general stress theory, the quality of a response depends on a person's sense of agency (Lazarus, 1984). Regardless of the

underlying motivation, a person's reaction likely depends on their perceived ability to restore their gender status. When an individual possesses high control beliefs, they may favor engaging in dominant or assertive behaviors designed to re-establish a position of social hierarchy. In this context, the emergence of aggression appears to be contingent upon this perceived agency, serving as a functional, albeit reactive, attempt to alter the social environment and restore standing. Conversely, when control beliefs are absent, leaving the individual feeling fundamentally powerless to change the situation, the response is likely to shift toward helplessness. In this state of perceived incapacity, the motivation to engage in externalized defense diminishes, leading instead to behavioral withdrawal and internalized distress. Incorporating these perceptions of agency may provide a valuable lens for future research to explain why the same motivational pressure leads to active dominance in some social contexts and psychological retreat in others.

4.4.2 Expanding the Perspective of Precarious Manhood

The current results contribute to a growing theoretical body investigating the nature of gender identity and its perceived vulnerability. Traditionally, the precarious manhood thesis has posited a fundamental asymmetry: while womanhood is framed as a biological "given," manhood is viewed as a tenuous social status that must be constantly earned and validated (Vandello & Bosson, 2013). However, the assumption that this vulnerability is unique to men is increasingly being re-evaluated.

As previously discussed, the perceived asymmetry in gender-identity threat responses may stem, in part, from a methodological misalignment in how "threat" is operationalized. In foundational studies, such as those by Bosson et al. (2009), women were often "threatened" by being told they were similar to men. Such a stimulus may not be appraised as a significant status loss by many participants, particularly because masculinity is frequently associated with high social status and institutional power. This highlights a fundamental challenge to the precarious manhood thesis itself: it is possible that the tenuous and elusive nature of masculinity is not

bound to the gender identity per se, but rather to the maintenance of high social status. If masculinity is synonymous with a dominant position in the social hierarchy, the "precariousness" observed may simply be the psychological anxiety inherent in maintaining any privileged status that is socially conferred rather than biologically inherent. From this perspective, any identity linked to power would be inherently precarious, as it requires constant outward proof to justify its continued possession.

Parallel to this question of social status is the question of how gender identity itself is constructed and maintained. These insights invite future research to move beyond traditional binary assumptions and examine gender identity as a multi-dimensional construct. By viewing precariousness as a potential experience across the gender spectrum, empirical research can be better aligned with the performative nature of gender (Butler, 2002). In this view, gender is not a stable, internal essence but an ongoing "doing"—a repetitive social performance that must be constantly maintained. If gender is understood as a continuous act rather than a static state, the risk of its "failure" becomes a shared feature of the human experience rather than a uniquely masculine vulnerability. For instance, a non-binary individual may experience significant identity precariousness when faced with persistent misgendering or social pressures to conform to a binary category. In such cases, the threat does not stem from a failure to be "masculine" or "feminine" in the traditional sense, but from the social invalidation of their specific gender performance. This broader perspective allows the precarious manhood thesis to adapt to a contemporary landscape where all gendered identities are, to some extent, socially precarious.

4.5 Conclusion

The present study aimed to replicate masculinity threat effects and test predictions of the EDT model in a German sample. None of the four hypotheses were supported: neither main effects of threat on aggressive or anxious cognition nor moderation by pressured or autonomous motivation reached statistical significance. These null findings are attributed to inadequate

measurement rather than absence of the underlying phenomena, given the poor psychometric properties of the German Word Fragment Completion Task. Despite these limitations, the study makes two positive contributions. First, the manipulation checks demonstrate that the masculinity threat paradigm can successfully induce perceived threats in German samples, providing a foundation for future studies with better outcome measures. Second, the validated German translation of the MMB scale enables future research on masculine motivation in German-speaking populations. Future research on masculinity threat effects should prioritize validated dependent measures with established psychometric properties. Word Fragment Completion Task, while theoretically attractive for their implicit nature, may not transfer reliably across languages and should be thoroughly validated before use in substantive hypothesis testing. Alternative approaches, including self-report measures, behavioral paradigms, and physiological indicators, each carry their own limitations but may provide more reliable assessments of threat-induced cognition and affect. The continued development and validation of masculinity threat paradigms remains an important goal for understanding how gender identity shapes psychological responses to social evaluation.

References

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301. https://doi.org/10.1177/1094428112470848

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*(9), 411–419. https://doi.org/10.1023/a:1023574427720

Blazina, C., & Watkins Jr, C. E. (1996). Masculine gender role conflict: Effects on college men's psychological well-being, chemical substance usage, and attitudes towards help-seeking. *Journal of Counseling Psychology*, *43*(4), 461–465. https://doi.org/10.1037/0022-0167.43.4.461

Bosson, J. K., Jurek, P., Vandello, J. A., Kosakowska-Berezecka, N., Olech, M., Besta, T., Bender, M., Hoorens, V., Becker, M., Timur Sevincer, A., & others. (2021). Psychometric properties and correlates of precarious manhood beliefs in 62 nations. *Journal of Cross-Cultural Psychology*, *52*(3), 231–258. https://doi.org/10.1177/0022022121997997

Bosson, J. K., Vandello, J. A., Burnaford, R. M., Weaver, J. R., & Arzu Wasti, S. (2009). Precarious manhood and displays of physical aggression. *Personality and Social Psychology Bulletin*, *35*(5), 623–634. https://doi.org/10.1177/0146167208331161

Brown, T. L., & Smith, N. G. (2023). Experimentally testing the impact of status threat on heterosexual men's use of anti-gay slurs: A precarious manhood and coalitional value perspective. *Current Psychology*, *42*(14), 11891–11901. https://doi.org/10.1007/s12144-021-02489-7

Butler, J. (2002). *Gender Trouble: Tenth Anniversary Edition (2nd ed.)*. Routledge. https://doi.org/10.4324/9780203902752

Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the southern culture of honor: An" experimental ethnography.". *Journal of Personality and Social Psychology*, *70*(5), 945–960. https://doi.org/10.4324/9780429501142-4

Cunningham, J. B., & McCrum-Gardner, E. (2007). Power, effect and sample size using GPower: practical issues for researchers and members of research ethics committees. *Evidence-Based Midwifery*, *5*, 132–137. https://link.gale.com/apps/doc/A172050741/HRCA?u=anon~712f907f&sid=googleScholar&xid=b307aea6

Dahl, J., Vescio, T., & Weaver, K. (2015). How threats to masculinity sequentially cause public discomfort, anger, and ideological dominance over women. *Social Psychology*, *46*(4), 242–254. https://doi.org/10.1027/1864-9335/a000248

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, *53*(6), 1024–1037. https://doi.org/10.1037/0022-3514.53.6.1024

Duckworth, K. D., & Trautner, M. N. (2019). Gender goals: Defining masculinity and navigating peer pressure to engage in sexual activity. *Gender & Society*, *33*(5), 795–817. https://doi.org/10.1177/0891243219863031

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, *83*(6), 1441–1455. https://doi.org/10.1037//0022-3514.83.6.1441

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fowler, S. L., & Geers, A. L. (2017). Does trait masculinity relate to expressing toughness? The effects of masculinity threat and self-affirmation in college men. *Psychology of Men & Masculinity*, *18*(2), 176–186. https://doi.org/10.1037/men0000053

Gilmore, D. D. (1990). *Manhood in the making: Cultural concepts of masculinity*. Yale University Press. https://doi.org/10.2307/3317781

Good, J. J., & Sanchez, D. T. (2010). Doing gender for different reasons: Why gender conformity positively and negatively predicts self-esteem. *Psychology of Women Quarterly*, *34*(2), 203–214. https://doi.org/10.1111/j.1471-6402.2010.01562.x

Harkness, J. A., Villar, A., & Edwards, B. (2010). Translation, adaptation, and design. *Survey Methods in Multinational, Multiregional, And Multicultural Contexts*, 115–140. https://doi.org/10.1002/9780470609927.ch7

Hartig, F. (2024). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. https://cran.r-project.org/package=DHARMa

Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological Review*, *94*(3), 319–340. https://doi.org/10.1037/0033-295x.94.3.319

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, *26*(3), 329–367. https://doi.org/10.1177/0049124198026003003

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jakupcak, M., Tull, M. T., & Roemer, L. (2005). Masculinity, Shame, and Fear of Emotions as Predictors of Men's Expressions of Anger and Hostility. *Psychology of Men & Masculinity*, *6*(4), 275–284. https://doi.org/10.1037/1524-9220.6.4.275

Jin, Y., Sun, C., Wu, J., An, J., & Li, J. (2021). Precarious manhood and its effects on aggression: The role of cultural script. *Journal of Interpersonal Violence*, *36*(9–10), NP5521–NP5544. https://doi.org/10.1177/0886260518800312

Kosakowska-Berezecka, N., Besta, T., Adamska, K., Jaśkiewicz, M., Jurek, P., & Vandello, J. A. (2016). If my masculinity is threatened I won't support gender equality? The role of agentic

self-stereotyping in restoration of manhood and perception of gender relations. *Psychology of Men & Masculinity*, *17*(3), 274. https://doi.org/10.1037/men0000016

Kugler, D., Ascone, L., Pohlmann, K., Weichenberger, M., & Kühn, S. (2022). Construct Validation of a German Version of the Word Completion Task for Aggression [Unpublished Manuscript]. *Universitätsklinikum Hamburg-Eppendorf*. https://doi.org/10.31234/osf.io/3dh5r

Lazarus, R. S. (1984). *Stress, appraisal, and coping* (Vol. 445). Springer.

Lease, S. H., Montes, S. H., Baggett, L. R., Sawyer, R. J., Fleming-Norwood, K. M., Hampton, A. B., Ovrebo, E., Çiftçi, A., & Boyraz, G. (2013). A cross-cultural exploration of masculinity and relationships in men from Turkey, Norway, and the United States. *Journal of Cross-Cultural Psychology*, *44*(1), 84–105. https://doi.org/10.1177/0022022111432293

Levant, R. F., Hirsch, L. S., Celentano, E., Cozza, T. M., & others. (1992). The male role: An investigation of contemporary norms. *Journal of Mental Health Counseling*, 325–337.

O'Neil, J. M. (2008). Summarizing 25 years of research on men's gender role conflict using the Gender Role Conflict Scale: New research paradigms and clinical implications. *The Counseling Psychologist*, *36*(3), 358–445. https://doi.org/10.1177/0011000008317057

O'Neil, J. M., Helms, B. J., Gable, R. K., David, L., & Wrightsman, L. S. (1986). Gender-Role Conflict Scale: College men's fear of femininity. *Sex Roles*, *14*(5–6), 335–350.

Parrott, D. J., & Zeichner, A. (2008). Determinants of anger and physical aggression based on sexual orientation: An experimental examination of hypermasculinity and exposure to male gender role violations. *Archives of Sexual Behavior*, *37*, 891–901. https://doi.org/10.1007/s10508-007-9194-z

Pleck, J. H. (1995). *The gender role strain paradigm: An update.*

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Rando, R. A., Rogers, J. R., & Brittan-Powell, C. S. (1998). Gender role conflict and college men's sexually aggressive attitudes and behavior. *Journal of Mental Health Counseling*, *20*(4), 359–370. https://doi.org/10.1007/bf00287583

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934

Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: the role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, *87*(2), 157–176. https://doi.org/10.1037/0022-3514.87.2.157

Schermerhorn, N. E., & Vescio, T. K. (2022). Perceptions of a sexual advance from gay men leads to negative affect and compensatory acts of masculinity. *European Journal of Social Psychology*, *52*(2), 260–279. https://doi.org/10.1002/ejsp.2775

Schmukle, S. C., & Egloff, B. (2004). Does the Implicit Association Test for assessing anxiety measure trait and state variance?. *European Journal of Personality*, *18*(6), 483–494. https://doi.org/10.1002/per.525

Stanaland, A., & Gaither, S. (2021). "Be a man": The role of social pressure in eliciting men's aggressive cognition. *Personality and Social Psychology Bulletin*, *47*(11), 1596–1611. https://doi.org/10.31234/osf.io/5wxtf

Stanaland, A., Gaither, S., & Gassman-Pines, A. (2023). When is masculinity "fragile"? An expectancy-discrepancy-threat model of masculine identity. *Personality and Social Psychology Review*, *27*(4), 359–377. https://doi.org/10.31234/osf.io/fgbk9

Stanaland, A., Gaither, S., Gassman-Pines, A., Galvez-Cepeda, D., & Cimpian, A. (2024). Adolescent boys' aggressive responses to perceived threats to their gender typicality. *Developmental Science*, *27*(6). https://doi.org/10.1111/desc.13544

Valved, T., Kosakowska-Berezecka, N., Besta, T., & Martiny, S. E. (2021). Gender belief systems through the lens of culture — Differences in precarious manhood beliefs and reactions to masculinity threat in Poland and Norway. *Psychology of Men & Masculinities*, *22*(2), 265–276. https://doi.org/10.1037/men0000331

Vandello, J. A., & Bosson, J. K. (2013). Hard won and easily lost: A review and synthesis of theory and research on precarious manhood. *Psychology of Men & Masculinity*, *14*(2), 101–121. https://doi.org/10.1037/a0029826

Vandello, J. A., Bosson, J. K., Cohen, D., Burnaford, R. M., & Weaver, J. R. (2008). Precarious manhood. *Journal of Personality and Social Psychology*, *95*(6), 1325–1339. https://doi.org/10.4135/9781446269930.n8

Vescio, T. K., Schermerhorn, N. E., Lewis, K. A., Yamaguchi-Pedroza, K., & Loviscky, A. J. (2025). Masculinity threats sequentially arouse public discomfort, anger, and positive attitudes toward sexual violence. *Personality and Social Psychology Bulletin*, *51*(1), 96–109. https://doi.org/0.1177/01461672231179431

Weaver, K. S., & Vescio, T. K. (2015). The justification of social inequality in response to masculinity threats. *Sex Roles*, *72*, 521–535. https://doi.org/10.1007/s11199-015-0484-y

Winer, B. J., Brown, D. R., Michels, K. M., & others. (1971). *Statistical principles in experimental design* (Vol. 2). Mcgraw-hill New York. https://doi.org/10.1037/11774-000

Wittlin, N. M., Lafrance, M., Dovidio, J. F., & Richeson, J. A. (2024). US cisgender women's psychological responses to physical femininity threats: Increased anxiety, reduced self-esteem. *Journal of Experimental Social Psychology*, *110*, 104547. https://doi.org/10.1016/j.jesp.2023.104547

World Economic Forum. (2024). *The Global Gender Gap Report 2024*. World Economic Forum. https://www.weforum.org/publications/global-gender-gap-report-2024/

Declaration of Independence

Eigenständigkeitserklärung

Ich versichere, dass ich die vorliegende schriftliche Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe, alle Ausführungen, die anderen Schriften wörtlich oder sinngemäß entnommen wurden, kenntlich gemacht sind und die Arbeit in gleicher oder ähnlicher Form noch nicht für andere Prüfungen verwendet wurde sowie keiner anderen Prüfungsbehörde vorgelegen hat.

Ich habe alle Stellen, die dem Wortlaut oder dem Sinn nach (inkl. Übersetzungen) anderen Werken entnommen sind, unter genauer Angabe der Quelle (einschl. des World Wide Web sowie anderer elektronischer Datensammlungen) deutlich als Entlehnung kenntlich gemacht. Dies gilt auch für angefügte Zeichnungen, bildliche Darstellungen, Skizzen und dergleichen.

Zusätzlich versichere ich, dass ich beim Einsatz von IT-/KI-gestützten Werkzeugen diese Werkzeuge in der unten genannten „Übersicht verwendeter Hilfsmittel„ mit ihrem Produktnamen und der Versionsnummer, meiner Bezugsquelle (z.B. URL) und Angaben zur Nutzung vollständig aufgeführt sowie die Checkliste wahrheitsgemäß ausgefüllt habe. Davon ausgenommen sind diejenigen IT-/KI-gestützten Schreibwerkzeuge, die von meinem zuständigen Prüfungsbüro bis zum Zeitpunkt der Abgabe meiner Arbeit als nicht anzeigepflichtig eingestuft wurden („Whitelist“). Bei der Erstellung dieser Arbeit habe ich durchgehend eigenständig und beim Einsatz IT-/KI-gestützter Schreibwerkzeuge steuernd gearbeitet.

Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung gemäß der fachspezifischen Prüfungsordnung und/oder der Fächerübergreifenden Satzung zur Regelung von Zulassung, Studium und Prüfung der Humboldt-Universität zu Berlin (ZSP-HU) eingeleitet wird.

Berlin, 09.02.2026 

Ort, Datum, Unterschrift

Übersicht verwendeter Hilfsmittel

Ausfüllhinweis: Bitte geben Sie an, welche der folgenden Programme Sie im Rahmen Ihrer Arbeit verwendet haben. Bitte tragen Sie dazu bei jedem Programm ein, wie Sie dieses genutzt haben. Tragen Sie dazu die jeweilige Ziffer in die Spalte „Nutzung" ein. Es können mehrere Ziffern je Programm eingetragen werden. Sollten Sie ein Programm genutzt haben, welches nicht in der Tabelle aufgeführt ist, ergänzen Sie dieses bitte unter Angabe der Nutzungsart.

Nutzung:

0. Nicht verwendet
1. Generierung von Ideen/Brainstorming
2. Literaturrecherche
3. Übersetzung von Texten
4. Zusammenfassen von Quellen
5. Inhalte auf andere Art und Weise erklären lassen (z. B. Konstrukte, methodische Vorgehensweisen, Analysen)
6. Erstellen von Textabschnitten, welche als Vorlage dienen
7. Überarbeitung von eigenen Textelementen (bitte Seitenzahl der Arbeit angeben)
8. Auswertung von Daten (z. B. Schreiben von Code, Definieren der passenden Auswertungsmethoden, Erstellen von Abbildungen)
9. Rechtschreibprüfung
10. Visualisierungen zu illustrativen oder dekorativen Zwecken

Checkliste: IT-/KI-gestützte Schreibwerkzeuge:

Programm (Bitte geben Sie auch weitere Programme an, die Sie verwendet haben, um Ihre Arbeit zu schreiben.)	**falls relevant: betroffene Abschnitte** (Bitte geben Sie Seitenzahlen an, sofern es nicht auf das gesamte Dokument zutrifft.)	**Nutzung** (Bitte Zahl eintragen – siehe oben)
ChatGPT	1, 2, 3, 4, 5, 6, 7, 8, 9	
Elicit.org		

Grammarly		
GitHub Copilot		
Perplexity AI		
ChatPDF		
DeepL	3	
BingChat		
Gamma		
Humanata AI		
Keenious		
Monica		
Claude	1, 2, 3, 4, 5, 6, 7, 8, 9	
Gemini	6, 7, 9	

Sofern Sie sich unsicher sind, ob ein Programm unter „künstliche Intelligenz“ fällt, tragen Sie es ein. Es erwachsen Ihnen keine Nachteile durch die Nennung des Programms. Sollten Sie weitere Programme verwendet haben, können Sie diese in die leeren Zeilen eintragen.

Whitelist

Folgende Programme müssen nicht aufgelistet oder bewertet werden. Diese Programme können ohne weitere Angaben genutzt werden:

- ✓ Microsoft Office, LaTeX, OpenOffice, iWork
- ✓ Google Scholar
- ✓ Datenbanken der Universitätsbibliothek
- ✓ Literaturverwaltungsprogramme (Zotero, Endnote, Mendeley, etc.)

Ggf. weitere Erklärungen:

Appendix A

A Materials and Measures Supplement

The following sections contain the original materials used in the study.

Study Introduction

The following text was displayed to participants at the beginning of the study.

> **Willkommen zur Studie!**
>
> Schön, dass Sie an dieser Online-Studie teilnehmen möchten – herzlich willkommen!
>
> Diese psychologische Untersuchung wird im Rahmen einer Masterarbeit am Institut für Psychologie der Humboldt-Universität zu Berlin durchgeführt. Ziel ist es das Zusammenspiel von Persönlichkeit, Motivation und Leistung bei Männern besser zu verstehen.
>
> Die Studie umfasst folgende Abschnitte:
>
> - Ein Fragebogen zu persönlichen Einstellungen zum Thema Männlichkeit
> - Ein geschlechtsspezifischer Wissenstest
> - Feedback zum geschlechtsspezifischen Wissenstest
> - Ein kurzer Test zur Denkgeschwindigkeit
> - Abschließende demographische Angaben
>
> Bitte bearbeiten Sie die Umfrage an einem ruhigen Ort, an dem Sie sich gut konzentrieren können. Die Teilnahme dauert etwa 10 bis 15 Minuten.
>
> **Wichtige Hinweise:**
>
> Die Studie richtet sich ausschließlich an Männer …
>
> … mit muttersprachlichen Deutschkenntnissen …
>
> … die mindestens 18 Jahre alt sind.
>
> Sollten diese Kriterien nicht auf Sie zutreffen, bedanken wir uns dennoch herzlich für Ihr Interesse.
>
> Die Teilnahme ist freiwillig. Sie können die Befragung jederzeit abbrechen – ohne Nachteile.

> Falls Sie die Teilnahme beenden möchten, nutzen Sie bitte den Button „Studie abbrechen“, der Ihnen nach der Einwilligung auf allen Seiten des Fragebogens zur Verfügung steht.
> Bitte beachten Sie, dass die Teilnahme nicht vergütet wird. Mit jedem vollständig ausgefüllten Fragebogen werden jedoch 1€ an ein Beratungsangebot für Jungen und Männer des Bundesforum Männer e.V. gespendet.
> Falls Sie die Umfrage mit SurveyCircle ausfüllen, erhalten Sie den Code auf der letzten Seite.
> Vielen Dank für Ihre Unterstützung!
> Im ersten Schritt bitten wir um Ihre **Einwilligung zur Teilnahme**.

English translation: Welcome to the study! This psychological study is conducted as part of a master's thesis at the Institute of Psychology, Humboldt University of Berlin. The aim is to better understand the interplay of personality, motivation, and performance in men. The study includes: a questionnaire on personal attitudes toward masculinity, a gender knowledge test, feedback on the test, a brief cognitive speed test, and demographic questions. Participation takes approximately 10–15 minutes. Eligibility requires identifying as male, native German proficiency, and being at least 18 years old. Participation is voluntary and can be discontinued at any time. For each completed questionnaire, 1€ is donated to a men's counseling organization.

Motivation for Masculine Behavior

Table A1 presents the nine items of the Motivation for Masculine Behavior scale with their original English wording and German translations.

Table A1

Motivation for Masculine Behavior Scale Items

Item	Subscale	English (Original)	German (Translation)
1	P	In general, I'm masculine because I want others' acceptance and approval.	Im Allgemeinen verhalte ich mich männlich, weil ich die Akzeptanz und Anerkennung anderer möchte.

Item	Subscale	English (Original)	German (Translation)
2	P	In general, I'm masculine because that is what people expect from me.	Allgemeinen bin ich männlich, weil das von mir erwartet wird.
3	P	I'm masculine because I want people to like me.	Ich verhalte mich männlich, weil ich möchte, dass man mich mag.
4	P	I'm masculine around other people because that is how others think I should be.	Ich verhalte mich in Gegenwart anderer männlich, um ihre Erwartungen zu erfüllen.
5	P	I'm not feminine because people wouldn't like me.	Ich verhalte mich nicht weiblich, weil ich glaube, dass mich die Leute sonst nicht mögen würden.
6	A	It's important to me to be masculine.	Es ist mir wichtig, männlich zu sein.
7	A	I enjoy being masculine.	Ich bin gerne männlich.
8	A	It makes me happy if I'm masculine.	Es macht mich glücklich, mich männlich zu verhalten.
9	A	It is important to me not to be feminine.	Es ist mir wichtig, mich nicht weiblich zu verhalten

Note. P = Pressured Motivation subscale; A = Autonomous Motivation subscale. Items rated on a 7-point Likert scale (1 = "Disagree Strongly" to 7 = "Agree Strongly"). The original English Items are from Stanaland & Gaither (2021).

Table A2 presents descriptive statistics for the Motivation for Masculine Behavior scale.

Table A2

Descriptive Statistics for Motivation for Masculine Behavior Scale

Item	German Item	*M*	*SD*
1	Im Allgemeinen verhalte ich mich männlich, weil ich die Akzeptanz und Anerkennung anderer möchte.	3.55	1.78
2	Im Allgemeinen bin ich männlich, weil das von mir erwartet wird.	3.75	1.74
3	Ich verhalte mich männlich, weil ich möchte, dass man mich mag.	3.35	1.68
4	Ich verhalte mich in Gegenwart anderer männlich, um ihre Erwartungen zu erfüllen.	3.51	1.74

Item	German Item	*M*	*SD*
5	Ich verhalte mich nicht weiblich, weil ich glaube, dass mich die Leute sonst nicht mögen würden.	2.97	1.68
6	Es ist mir wichtig, männlich zu sein.	4.06	1.79
7	Ich bin gerne männlich.	5.29	1.49
8	Es macht mich glücklich, mich männlich zu verhalten.	4.32	1.68
9	Es ist mir wichtig, mich nicht weiblich zu verhalten.	3.26	1.77

Note. N = 154. Items 1–5 = Pressured Motivation subscale; Items 6–9 = Autonomous Motivation subscale. Items rated on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

The transfer of the MMB scale from English to German was conducted following the Translation, Review, Adjudication, Pretesting, and Documentation protocol (Harkness et al., 2010).

Translation

Two independent native German, fluent English translators were given the original scale items and were instructed to find the most appropriate translation focusing on natural and semantically correct phrasing. Table A3 and Table A4 present the independent translations and translator comments from this initial translation phase.

Table A3

Motivation for Masculine Behavior Scale Translation: Translator 1

Item	English (Original)	German Translation	Comments
1	In general, I'm masculine because I want others' acceptance and approval.	Im Allgemeinen bin ich männlich, weil ich von anderen akzeptiert und anerkannt werden will.	„Acceptance and approval" wurde mit „akzeptiert und anerkannt" übersetzt, um sowohl emotionale als auch soziale Aspekte der Zustimmung wiederzugeben. „Im Allgemeinen" klingt natürlich.

Item	English (Original)	German Translation	Comments
2	In general, I'm masculine because that is what people expect from me.	Im Allgemeinen bin ich männlich, weil die Leute das von mir erwarten.	Die Formulierung ist direkt und entspricht der Alltagssprache. „Die Leute" ist eine geläufige, allgemein gehaltene Übersetzung für „people".
3	I'm masculine because I want people to like me.	Ich bin männlich, weil ich möchte, dass man mich mag.	Die Konstruktion mit „dass man mich mag" nutzt das unpersönliche „man", was im Deutschen üblich ist, um eine allgemeine Aussage über „people" zu treffen.
4	I'm masculine around other people because that is how others think I should be.	Ich verhalte mich in der Nähe anderer männlich, weil sie denken, dass ich so sein sollte.	„Verhalte mich männlich" betont das Verhalten (performative Männlichkeit)
5	I'm not feminine because people wouldn't like me.	Ich bin nicht feminin, weil ich glaube, dass mich die Leute dann nicht mögen würden.	„Ich glaube" wurde eingefügt, um die subjektive Wahrnehmung zu betonen – im Deutschen klingt es so natürlicher. „Dann" verstärkt die kausale Beziehung.
6	It's important to me to be masculine.	Es ist mir wichtig, männlich zu sein.	Sehr direkte und idiomatische Übersetzung – stilistisch wie in Fragebögen oder formellen Aussagen gebräuchlich.
7	I enjoy being masculine.	Ich genieße es, männlich zu sein.	Klare, wörtliche und gut klingende Übersetzung. „Es genießen" ist im Deutschen idiomatisch.

Item	English (Original)	German Translation	Comments
8	It makes me happy if I'm masculine.	Es macht mich glücklich, wenn ich männlich bin.	Wörtlich, aber idiomatisch korrekt. Die Satzstruktur ist im Deutschen üblich und gut verständlich.
9	It is important to me not to be feminine.	Es ist mir wichtig, nicht feminin zu sein.	Direkte Negation, die in dieser Form im Deutschen völlig natürlich klingt. Wichtig war, „nicht feminin" klar und nicht überbetont darzustellen.

Note. Items 1–5 = Pressured Motivation subscale; Items 6–9 = Autonomous Motivation subscale.

Table A4

Motivation for Masculine Behavior Scale Translation: Translator 2

Item	English (Original)	German Translation	Comments
1	In general, I'm masculine because I want others' acceptance and approval.	Im Allgemeinen bin ich männlich, weil ich die Akzeptanz und Zustimmung anderer möchte.	'Acceptance and approval' wird im Deutschen als 'Akzeptanz und Zustimmung' übersetzt.
2	In general, I'm masculine because that is what people expect from me.	Im Allgemeinen bin ich männlich, weil das von mir erwartet wird.	'That is what people expect from me' wird im Deutschen als 'weil das von mir erwartet wird' übersetzt.
3	I'm masculine because I want people to like me.	Ich bin männlich, weil ich möchte, dass die Leute mich mögen.	Keine besonderen Anmerkungen.
4	I'm masculine around other people because that is how others think I should be.	Ich bin in Gegenwart anderer männlich, weil andere	'Around other people' wird im Deutschen als 'in Gegenwart anderer' übersetzt.

Item	English (Original)	German Translation	Comments
		denken, dass ich so sein sollte.	
5	I'm not feminine because people wouldn't like me.	Ich bin nicht weiblich, weil die Leute mich dann nicht mögen würden.	Keine besonderen Anmerkungen.
6	It's important to me to be masculine.	Es ist mir wichtig, männlich zu sein.	Keine besonderen Anmerkungen.
7	I enjoy being masculine.	Ich genieße es, männlich zu sein.	Keine besonderen Anmerkungen.
8	It makes me happy if I'm masculine.	Es macht mich glücklich, wenn ich männlich bin.	Keine besonderen Anmerkungen.
9	It is important to me not to be feminine.	Es ist mir wichtig, nicht weiblich zu sein.	Keine besonderen Anmerkungen.

Note. Items 1–5 = Pressured Motivation subscale; Items 6–9 = Autonomous Motivation subscale.

Review and Adjudication

The independent translations highlighted two key considerations for the review phase. First, both translators used the German word *männlich* rather than the loanword *maskulin*, suggesting that *männlich* (and correspondingly *weiblich* for feminine) sounds more natural in everyday German. The terms *maskulin* and *feminin* are more commonly used in German to describe physical appearance rather than behavioral traits. Second, the translators predominantly used stative phrasing (*ich bin männlich*; "I am masculine") rather than agentic phrasing (*ich verhalte mich männlich*; "I behave masculinely").

These observations were discussed with the thesis supervisor and other students during the review phase. Two decisions were made for the final translations. First, the translations retained *männlich* and *weiblich* rather than *maskulin* and *feminin*, as these terms are more familiar in German when describing personal characteristics. Second, a mix of static and agentic phrasings was adopted in the final version. The purely statuc phrasing *ich bin männlich* carries

a connotation of "being a man" in German, which may imply an innate, ascribed property rather than an enacted behavior. To balance fidelity to the original English items with the recognition that masculinity involves behavioral choices, the final translations incorporated both phrasings. This approach is consistent with prior work on motivation to conform to gender norms by Good & Sanchez (2010), which also employed varied item phrasings.

Pre-Test

A pre-test was conducted with nine male participants to evaluate the translated scale. No changes were made to the translations based on this pre-test. Participants reported that although they understood the items, it was not easy for them to provide self-judgments about their motivations for being masculine. This feedback suggests that the construct may be inherently difficult to introspect upon, which is consistent with the implicit nature of motivational processes.

Masculinity Threat Induction

The masculinity threat was administered by requiring participants to complete a test (Gender Knowledge Test) supposedly measuring their knowledge of gender-specific topics and subsequently providing them with feedback regarding their performance.

Gender Knowledge Test

The transfer of the Gender Knowlesge Test items (orignally from (Rudman & Fairchild, 2004), used in a culturally adapted version by (Valved et al., 2021)) from English to German was conducted following the Translation, Review, Adjudication, Pretesting, and Documentation protocol (Harkness et al., 2010). Table A5 presents the 30 items of the Gender Knowledge Test with their original English wording and German translations.

Table A5

Gender Knowledge Test Items

Item	Type	English (Original)	German (Translation)
1	M	A motorcycle engine turning at 8000 rpms generates an exhaust sound at (4000 rpms vs. 8000 rpms)	Ein Motorrad mit 8000 Umdrehungen pro Minute erzeugt ein Auspuffgeräusch bei (4000 U/min vs. 8000 U/min)
2	M	To help an engine produce more power you should (inject the fuel vs. reduce displacement)	Um die Leistung eines Motors zu erhöhen, sollte man (den Kraftstoff einspritzen vs. den Hubraum verringern)
3	M	In nature, the best analogy for a spark plug is (solar fire vs. lightning)	In der Natur entspricht eine Zündkerze am ehesten (Sonnenfeuer vs. Blitz)
4	M	Karate originated in martial arts developed in (Japan vs. China)	Karate entstand aus Kampfkünsten, entwickelt in (Japan vs. China)
5	M	Soldiers in WWII often used what type of guns? (Gatling vs. Tommy)	Welche Art von Gewehren wurde im Zweiten Weltkrieg oft von Soldaten verwendet? (Gatling vs. Tommy)
6	M	The groove inside the barrel of a revolver is (spiraled vs. smooth)	Die Rillen im Lauf eines Revolvers sind (spiralförmig vs. glatt)
7	M	If you need to replace the tank ball in a toilet, ask for a (flapper vs. ball cock)	Wenn man den Schwimmer im Spülkasten austauschen muss, fragt man nach (einer Absperrklappe vs. einem Kugelhahn)
8	M	The paste used for soldering joints is called (gel vs. flux)	Die Paste, die man zum Löten verwendet, heißt (Gel vs. Flussmittel)
9	M	Hugh Hefner first published Playboy magazine in (1963 vs. 1953)	Hugh Hefner veröffentlichte das erste Playboy-Magazin im Jahr (1963 vs. 1953)
10	M	Arnold Schwarzenegger killed more people in which film? (True Lies vs. Total Recall)	In welchem Film tötete Arnold Schwarzenegger mehr Menschen? (True Lies vs. Total Recall)
11	M	By Olympic rules, boxing gloves for all weight classes weigh (12 ounces vs. 10 ounces)	Laut Olympischen Regeln wiegen Boxhandschuhe in allen Gewichtsklassen (12 Unzen vs. 10 Unzen)

Item	Type	English (Original)	German (Translation)
12	M	When punching someone, you should aim your fist (a foot beyond target vs. directly at target)	Beim Zuschlagen sollte die Faust worauf gerichtet sein? (hinter das Ziel vs. direkt auf das Ziel)
13	M	When punching someone, the majority of the force comes from (the speed of your fist vs. your upper arm and shoulder)	Beim Zuschlagen kommt die meiste Kraft von (der Geschwindigkeit der Faust vs. dem Oberarm und der Schulter)
14	M	What's the best way to deflect a punch? (use the forearm vs. use hand)	Was ist der beste Weg, einen Schlag abzuwehren? (mit dem Unterarm zum Block vs. mit der Hand fangen)
15	M	When ramming a car to disable it, you should aim for the (rear passenger's tire vs. front driver's tire)	Wenn man ein Auto rammt, um es außer Gefecht zu setzen, sollte man worauf zielen? (hinteren Beifahrerreifen vs. vorderen Fahrerreifen)
16	F	You wear Manolo Blahniks on your (head vs. feet)	Man trägt Manolo Blahniks am (Kopf vs. Fuß)
17	F	Botox temporarily erases wrinkles by (skin hydration vs. muscle paralysis)	Botox glättet Falten vorübergehend durch (Hautbefeuchtung vs. Muskelentspannung)
18	F	The first company to develop hair coloring was (Clairol vs. L'Oreal)	Die erste Firma, die Haarfärbemittel entwickelte, war (Clairol vs. L'Oréal)
19	F	The TV show "Sex in the City" popularized which drink? (Cosmopolitan vs. Manhattan)	Welches Getränk machte die Sendung "Sex and the City" populär? (Cosmopolitan vs. Manhattan)
20	F	Children typically start to teethe when they are (over vs. under) 1 year old?	In welchem Alter beginnen Kinder typischerweise mit dem Zahnen? (über 1 Jahr vs. unter 1 Jahr)
21	F	Toilet training should start around the age of (36 months vs. 12 months)	In welchem Alter sollte das Toilettentraining beginnen? (36 Monate vs. 12 Monate)

Item	Type	English (Original)	German (Translation)
22	F	Children should not be given which medication? (ibuprofen vs. aspirin)	Welches Medikament sollte Kindern nicht gegeben werden? (Ibuprofen vs. Aspirin)
23	F	How many cups of water does it take to cook 1 cup of rice? (2 vs. 3)	Wie viele Tassen Wasser braucht man, um 1 Tasse Reis zu kochen? (2 Tassen vs. 3 Tassen)
24	F	Leftovers can be safely kept at room temperature for up to (4 hours vs. 2 hours)	Wie lange können Reste bei Zimmertemperatur bedenkenlos aufbewahrt werden? (4 Stunden vs. 2 Stunden)
25	F	If you don't have baking powder, you substitute baking soda plus (salt vs. cream of tartar)	Wenn kein Backpulver vorhanden ist, ersetzt man es durch Natron und (Salz vs. Weinstein)
26	F	A roux is best described as a (sauce vs. cake)	Eine Mehlschwitze ist am besten zu beschreiben als (Soße vs. Kuchen)
27	F	Compared to men, women need more (iron vs. zinc)	Im Vergleich zu Männern benötigen Frauen mehr (Eisen vs. Zink)
28	F	Which of these contains a natural mood enhancer? (chocolate vs. caviar)	Welches dieser Lebensmittel enthält einen natürlichen Stimmungsaufheller? (Schokolade vs. Kaviar)
29	F	During pregnancy, morning sickness usually occurs in which trimester? (second vs. first)	Während der Schwangerschaft tritt morgendliche Übelkeit normalerweise in welchem Trimester auf? (zweites vs. erstes)
30	F	Exercises that improve a woman's sex life are called (Kegel's vs. Pilates)	Übungen, die das Sexualleben einer Frau verbessern, heißen (Kegel-Übungen vs. Pilates)

Note. M = Masculine-stereotyped knowledge; F = Feminine-stereotyped knowledge. Correct answers are underlined in the original study materials. Items adapted from Valved et al. (2021). Items were presented in a fixed randomized order.

Translation. Two independent native German, fluent English translators were given the original scale items and were instructed to find the most appropriate translation focusing on natural and semantically correct phrasing. Table A6 and Table A7 present the independent translations and translator comments.

Table A6

Gender Knowledge Test Translation: Translator 1

Item	Type	English (Original)	German Translation	Comments
1	M	A motorcycle engine turning at 8000 rpms generates an exhaust sound at (4000 rpms vs. 8000 rpms)	Ein Motorrad mit 8000 Umdrehungen pro Minute erzeugt einen Auspuffton bei (4000 U/min vs. 8000 U/min)	Technisches Wissen, stereotypisch männlich; 'U/min' als gebräuchliche Abkürzung.
2	M	To help an engine produce more power you should (inject the fuel vs. reduce displacement)	Um einem Motor zu mehr Leistung zu verhelfen, sollte man (den Kraftstoff einspritzen vs. den Hubraum verringern)	Fachbegriffe wie 'Kraftstoff einspritzen' und 'Hubraum' kulturgerecht übersetzt.
3	M	In nature, the best analogy for a spark plug is (solar fire vs. lightning)	In der Natur entspricht eine Zündkerze am ehesten (Sonnenfeuer vs. Blitz)	'Sonnenfeuer' ist eine poetische, aber verständliche Entsprechung im Deutschen.
4	M	Karate originated in martial arts developed in (Japan vs. China)	Karate stammt ursprünglich aus Kampfkünsten, die in (Japan vs. China) entwickelt wurden	Bekanntes kulturelles Wissen; DE kennt Unterschiede zwischen Japan/China.
5	M	Soldiers in WWII often used what type of guns? (Gatling vs. Tommy)	Soldaten im Zweiten Weltkrieg benutzten häufig welche Art von Waffe? (Gatling vs.	'Tommy Gun' als 'Thompson' bekannt; 'Gatling' weniger gebräuchlich.

Item	Type	English (Original)	German Translation	Comments
			Thompson-Maschinenpistole)	
6	M	The groove inside the barrel of a revolver is (spiraled vs. smooth)	Die Rillen im Lauf eines Revolvers sind (spiralförmig vs. glatt)	‘Spiralförmig’ und ‘glatt’ beschreiben den Lauf anschaulich im Deutschen.
7	M	If you need to replace the tank ball in a toilet, ask for a (flapper vs. ball cock)	Wenn man den Schwimmer im Spülkasten austauschen muss, fragt man nach einem (Membranventil vs. Schwimmerventil)	‘Ball cock’ wird im Deutschen anders benannt – Sanitärsprache angepasst.
8	M	The paste used for soldering joints is called (gel vs. flux)	Die Paste, die man zum Löten verwendet, heißt (Gel vs. Flussmittel)	‘Flussmittel’ ist in technischer Sprache üblich – passt für Löt-Kontext.
9	M	Hugh Hefner first published Playboy magazine in (1963 vs. 1953)	Hugh Hefner veröffentlichte das erste Playboy-Magazin im Jahr (1963 vs. 1953)	Jahreszahlen und historische Fakten sind universell verständlich.
10	M	Arnold Schwarzenegger killed more people in which film? (True Lies vs. Total Recall)	In welchem Film tötete Arnold Schwarzenegger mehr Menschen? (True Lies vs. Total Recall)	Arnold Schwarzenegger bekannt – keine Lokalisierung nötig.
11	M	By Olympic rules, boxing gloves for all weight classes weigh (12 ounces vs. 10 ounces)	Laut Olympischen Regeln wiegen Boxhandschuhe in allen Gewichtsklassen (12 Unzen vs. 10 Unzen)	Unzen gelassen wegen Regelkontext – könnte metrisch ergänzt werden.
12	M	When punching someone, you should aim your fist (a foot beyond target vs. directly at target)	Beim Zuschlagen sollte man die Faust auf (einen Fuß hinter das Ziel vs. direkt auf das Ziel) richten	Ein ‘Fuß’ ist ca. 30 cm – kulturell weniger gebräuchlich.

Item	Type	English (Original)	German Translation	Comments
13	M	When punching someone, the majority of the force comes from (the speed of your fist vs. your upper arm and shoulder)	Beim Zuschlagen kommt die meiste Kraft von (der Geschwindigkeit der Faust vs. dem Oberarm und der Schulter)	Anatomisches und sportliches Wissen – direkt übertragbar.
14	M	What's the best way to deflect a punch? (use the forearm vs. use hand)	Was ist der beste Weg, einen Schlag abzuwehren? (Unterarm zum Blocken vs. Hand zum Fangen nutzen)	Kampfkunstkontext – verständlich, keine kulturelle Anpassung nötig.
15	M	When ramming a car to disable it, you should aim for the (rear passenger's tire vs. front driver's tire)	Wenn man ein Auto stoppen will, sollte man auf (hinteren Beifahrerreifen vs. vorderen Fahrreifen) zielen	Auto-Action-Wissen; eher aus Filmkontext – Übersetzung bleibt erhalten.
16	F	You wear Manolo Blahniks on your (head vs. feet)	Man trägt Manolo Blahniks am (Kopf vs. Fuß)	Markenbezug klar; 'Kopf' vs. 'Fuß' bleibt humorvoll.
17	F	Botox temporarily erases wrinkles by (skin hydration vs. muscle paralysis)	Botox glättet Falten vorübergehend durch (Hautbefeuchtung vs. Muskelentspannung)	'Muskelentspannung' statt 'Paralyse' – klingt natürlicher.
18	F	The first company to develop hair coloring was (Clairol vs. L'Oreal)	Die erste Firma, die Haarfärbemittel entwickelte, war (Clairol vs. L'Oréal)	Beide Marken in DE bekannt – keine Anpassung nötig.
19	F	The TV show "Sex in the City" popularized which drink? (Cosmopolitan vs. Manhattan)	Welche Drink machte die Sendung "Sex and the City" populär? (Cosmopolitan vs. Manhattan)	Titel der Serie auf 'Sex and the City' korrigieren!

Item	Type	English (Original)	German Translation	Comments
20	F	Children typically start to teethe when they are (over vs. under) 1 year old?	Kinder bekommen typischerweise Zähne, wenn sie (über 1 Jahr alt vs. unter 1 Jahr alt) sind	Medizinisch korrekte und allgemein bekannte Information.
21	F	Toilet training should start around the age of (36 months vs. 12 months)	Die Sauberkeitserziehung sollte ungefähr mit (36 Monaten vs. 12 Monaten) beginnen	Begriff 'Sauberkeitserziehung' ist kulturell angepasst.
22	F	Children should not be given which medication? (ibuprofen vs. aspirin)	Kindern sollte welches Medikament nicht gegeben werden? (Ibuprofen vs. Aspirin)	Aspirin-Verbot bei Kindern auch in DE etabliert.
23	F	How many cups of water does it take to cook 1 cup of rice? (2 vs. 3)	Wie viele Tassen Wasser braucht man, um eine Tasse Reis zu kochen? (2 vs. 3 Tassen)	Tassen als Maßeinheit in Rezepten in DE bekannt.
24	F	Leftovers can be safely kept at room temperature for up to (4 hours vs. 2 hours)	Reste können bei Zimmertemperatur sicher bis zu (4 Stunden vs. 2 Stunden) aufbewahrt werden	Lebensmittelhygiene ähnlich geregelt wie in den USA.
25	F	If you don't have baking powder, you substitute baking soda plus (salt vs. cream of tartar)	Wenn kein Backpulver vorhanden ist, ersetzt man es durch Natron und (Salz vs. Weinstein)	'Weinstein' als Begriff in DE im Backbereich bekannt.
26	F	A roux is best described as a (sauce vs. cake)	Eine Mehlschwitze ist am besten beschrieben als (Soße vs. Kuchen)	'Mehlschwitze' ist die gängige Übersetzung für 'roux'.
27	F	Compared to men, women need more (iron vs. zinc)	Im Vergleich zu Männern benötigen Frauen mehr (Eisen vs. Zink)	Ernährungswissen, in DE auch medizinisch verbreitet.

Item	Type	English (Original)	German Translation	Comments
28	F	Which of these contains a natural mood enhancer? (chocolate vs. caviar)	Welches dieser Lebensmittel enthält einen natürlichen Stimmungsaufheller? (Schokolade vs. Kaviar)	'Schokolade als Seelentröster' – starkes kulturelles Stereotyp.
29	F	During pregnancy, morning sickness usually occurs in which trimester? (second vs. first)	Übelkeit in der Schwangerschaft tritt meistens im (zweiten vs. ersten) Trimester auf	Trimester-Wissen in Schwangerschaft universell.
30	F	Exercises that improve a woman's sex life are called (Kegel's vs. Pilates)	Übungen, die das Sexualleben einer Frau verbessern, heißen (Kegel-Übungen vs. Pilates)	'Kegel-Übungen' in DE bekannt; Kontext bleibt erhalten.

Note. M = Masculine-stereotyped knowledge; F = Feminine-stereotyped knowledge.

Table A7

Gender Knowledge Test Translation: Translator 2

Item	Type	English (Original)	German Translation	Comments
1	M	A motorcycle engine turning at 8000 rpms generates an exhaust sound at (4000 rpms vs. 8000 rpms)	Ein Motorradmotor, der mit 8000 U/min läuft, erzeugt einen Abgasgeräusch bei (4000 U/min vs. 8000 U/min)	'rpms' wird im Deutschen als 'U/min' (Umdrehungen pro Minute) übersetzt.
2	M	To help an engine produce more power you should (inject the fuel vs. reduce displacement)	Um einem Motor zu helfen, mehr Leistung zu erzeugen, sollten Sie (den Kraftstoff einspritzen vs. den Hubraum reduzieren)	'Displacement' wird im Deutschen als 'Hubraum' bezeichnet.

Item	Type	English (Original)	German Translation	Comments
3	M	In nature, the best analogy for a spark plug is (solar fire vs. lightning)	In der Natur ist das beste Analogon für eine Zündkerze (Sonnenfeuer vs. Blitz)	'Spark plug' wird im Deutschen als 'Zündkerze' bezeichnet.
4	M	Karate originated in martial arts developed in (Japan vs. China)	Karate entstand in den Kampfkünsten, die in (Japan vs. China) entwickelt wurden	Keine besonderen Anmerkungen.
5	M	Soldiers in WWII often used what type of guns? (Gatling vs. Tommy)	Welche Art von Gewehren wurde im Zweiten Weltkrieg oft von Soldaten verwendet? (Gatling vs. Tommy)	'Gatling' und 'Tommy' sind spezifische Begriffe, die im Deutschen gleich bleiben.
6	M	The groove inside the barrel of a revolver is (spiraled vs. smooth)	Die Rille im Inneren des Revolverlaufs ist (spiralförmig vs. glatt)	Keine besonderen Anmerkungen.
7	M	If you need to replace the tank ball in a toilet, ask for a (flapper vs. ball cock)	Wenn Sie die Tankkugel in einer Toilette ersetzen müssen, fragen Sie nach einem (Klappenventil vs. Kugelventil)	'Flapper' und 'ball cock' sind spezifische Begriffe, die im Deutschen als 'Klappenventil' und 'Kugelventil' bezeichnet werden.
8	M	The paste used for soldering joints is called (gel vs. flux)	Die Paste, die zum Löten von Verbindungen verwendet wird, heißt (Gel vs. Flussmittel)	'Flux' wird im Deutschen als 'Flussmittel' bezeichnet.
9	M	Hugh Hefner first published Playboy magazine in (1963 vs. 1953)	Hugh Hefner veröffentlichte das Playboy-Magazin erstmals in (1963 vs. 1953)	Keine besonderen Anmerkungen.
10	M	Arnold Schwarzenegger killed more people in	In welchem Film hat Arnold Schwarzenegger	Keine besonderen Anmerkungen.

Item	Type	English (Original)	German Translation	Comments
		which film? (True Lies vs. Total Recall)	mehr Menschen getötet? (True Lies vs. Total Recall)	
11	M	By Olympic rules, boxing gloves for all weight classes weigh (12 ounces vs. 10 ounces)	Nach den Olympischen Regeln wiegen Boxhandschuhe für alle Gewichtsklassen (12 Unzen vs. 10 Unzen)	‘Ounces’ wird im Deutschen als ‘Unzen’ bezeichnet.
12	M	When punching someone, you should aim your fist (a foot beyond target vs. directly at target)	Beim Schlagen sollte man seine Faust (etwa einen Fuß über das optimale Ziel hinaus vs. direkt auf das Ziel) richten	‘A foot beyond’ wird im Deutschen als ‘etwa einen Fuß über … hinaus’ übersetzt.
13	M	When punching someone, the majority of the force comes from (the speed of your fist vs. your upper arm and shoulder)	Beim Schlagen kommt der Großteil der Kraft aus (der Geschwindigkeit deiner Faust vs. deinem Oberarm und deiner Schulter)	Keine besonderen Anmerkungen.
14	M	What’s the best way to deflect a punch? (use the forearm vs. use hand)	Was ist die beste Methode, um einen Schlag abzuwehren? (den Unterarm verwenden, um ihn zu blockieren vs. die Hand verwenden, um ihn zu fangen)	Keine besonderen Anmerkungen.
15	M	When ramming a car to disable it, you should aim for the (rear passenger’s tire vs. front driver’s tire)	Wenn man ein Auto rammt, um es außer Gefecht zu setzen, sollte man auf den (hinteren Beifahrerreifen vs.	Keine besonderen Anmerkungen.

Item	Type	English (Original)	German Translation	Comments
			vorderen Fahrerreifen) zielen	
16	F	You wear Manolo Blahniks on your (head vs. feet)	Man trägt Manolo Blahniks an den (Kopf vs. Füßen)	‘Manolo Blahniks’ sind eine bekannte Schuhmarke, daher bleibt der Name im Deutschen gleich.
17	F	Botox temporarily erases wrinkles by (skin hydration vs. muscle paralysis)	Botox beseitigt Falten vorübergehend durch (Hautbefeuchtung vs. Muskelentspannung)	‘Muscle paralysis’ wird im Deutschen als ‘Muskelentspannung’ bezeichnet.
18	F	The first company to develop hair coloring was (Clairol vs. L’Oreal)	Das erste Unternehmen, das Haarfärbemittel entwickelte, war (Clairol vs. L’Oréal)	Keine besonderen Anmerkungen.
19	F	The TV show “Sex in the City” popularized which drink? (Cosmopolitan vs. Manhattan)	Welches Getränk wurde durch die TV-Show “Sex and the City” populär? (Cosmopolitan vs. Manhattan)	Keine besonderen Anmerkungen.
20	F	Children typically start to teethe when they are (over vs. under) 1 year old?	Kinder beginnen typischerweise mit dem Zahnen, wenn sie (über vs. unter) 1 Jahr alt sind?	Keine besonderen Anmerkungen.
21	F	Toilet training should start around the age of (36 months vs. 12 months)	Die Sauberkeitserziehung sollte etwa im Alter von (36 Monaten vs. 12 Monaten) beginnen	Keine besonderen Anmerkungen.
22	F	Children should not be given which medication? (ibuprofen vs. aspirin)	Welches Medikament sollte Kindern nicht gegeben werden? (Ibuprofen vs. Aspirin)	Keine besonderen Anmerkungen.

Item	Type	English (Original)	German Translation	Comments
23	F	How many cups of water does it take to cook 1 cup of rice? (2 vs. 3)	Wie viele Tassen Wasser braucht man, um 1 Tasse Reis zu kochen? (2 Tassen vs. 3 Tassen)	Keine besonderen Anmerkungen.
24	F	Leftovers can be safely kept at room temperature for up to (4 hours vs. 2 hours)	Reste können sicher bei Raumtemperatur für bis zu (4 Stunden vs. 2 Stunden) aufbewahrt werden	Keine besonderen Anmerkungen.
25	F	If you don't have baking powder, you substitute baking soda plus (salt vs. cream of tartar)	Wenn Sie kein Backpulver haben, können Sie Natron plus (Salz vs. Weinsteinbackpulver) ersetzen	'Cream of tartar' wird im Deutschen als 'Weinsteinbackpulver' bezeichnet.
26	F	A roux is best described as a (sauce vs. cake)	Eine Roux wird am besten als (Sauce vs. Kuchen) beschrieben	'Roux' bleibt im Deutschen gleich, da es ein Fachbegriff ist.
27	F	Compared to men, women need more (iron vs. zinc)	Im Vergleich zu Männern benötigen Frauen mehr (Eisen vs. Zink)	Keine besonderen Anmerkungen.
28	F	Which of these contains a natural mood enhancer? (chocolate vs. caviar)	Welches davon enthält einen natürlichen Stimmungsaufheller? (Schokolade vs. Kaviar)	Keine besonderen Anmerkungen.
29	F	During pregnancy, morning sickness usually occurs in which trimester? (second vs. first)	Während der Schwangerschaft tritt morgendliche Übelkeit normalerweise in welchem Trimester auf? (zweites vs. erstes)	Keine besonderen Anmerkungen.
30	F	Exercises that improve a woman's sex life are called (Kegel's vs. Pilates)	Übungen, die das Sexualleben einer Frau verbessern, heißen	'Kegel's' wird im Deutschen als 'Kegel-Übungen' bezeichnet.

Item	Type	English (Original)	German Translation	Comments
			(Kegel-Übungen vs. Pilates)	

Note. M = Masculine-stereotyped knowledge; F = Feminine-stereotyped knowledge.

Review & Adjudication. The review phase for the Gender Knowledge Test was more straightforward than for the MMB scale. The primary focus was on clarifying technical terminology where the two translations diverged, particularly for domain-specific terms in automotive, plumbing, and culinary contexts. Unlike the MMB scale, where conceptual nuances of masculinity required careful consideration, the Gender Knowledge Test items involved factual content with established German equivalents. Since the purpose of the Gender Knowledge Test was not to provide a psychometrically valid measure but rather to create a plausible and challenging test that could credibly threaten participants' gender identity, the review prioritized ensuring that all items appeared believable and appropriately difficult. The final translations were selected to maintain this sense of plausibility while using natural German phrasing.

Pre-Test. The same pre-test with nine male participants was used to evaluate the translated Gender Knowledge Test. No changes were made to the translations based on this pre-test. Participants reported that they enjoyed the questions, describing the experience as "game-like." This positive reception supported the intended design of the Gender Knowledge Test as an engaging measure that participants would take seriously and find credible.

Masculinity Threatening and Non-Threatening Feedback

The following feedback screens were displayed to participants after completing the Gender Knowledge Test. The threat condition received feedback indicating below-average performance relative to other men, while the control condition received feedback indicating average performance.

Threat Condition.

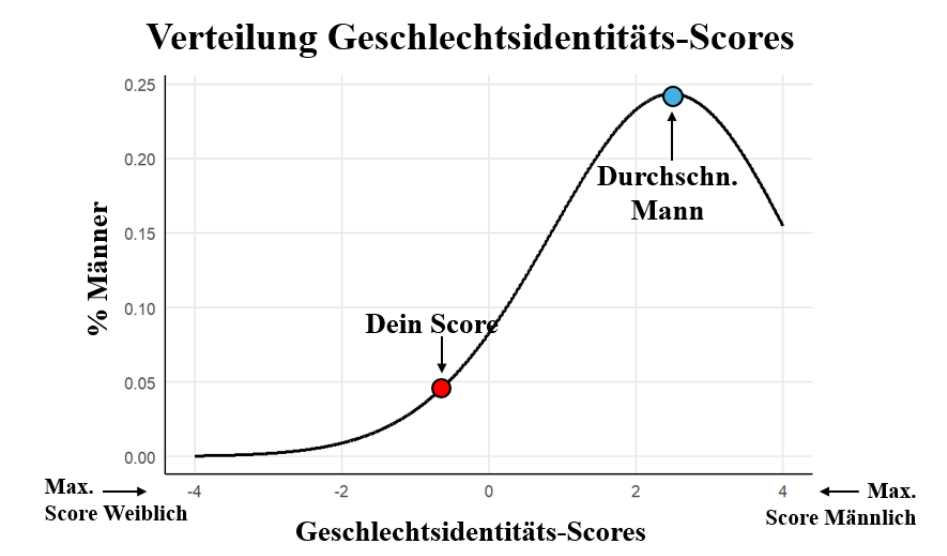
Glückwunsch – du hast den Geschlechtsspezifischen Wissenstest abgeschlossen!

Dieser Test wurde entwickelt, um Unterschiede im Wissen über geschlechtsspezifische Themen zu messen. Frühere Studien zeigen, dass bestimmte Wissensbereiche stärker mit traditionell männlich oder weiblich geprägten Rollen und Selbstbildern verbunden sind. Die Auswertung basiert auf einer Skala von −4 (ausgeprägt weiblich) bis +4 (ausgeprägt männlich).

Ihr Score beträgt: −1,83

Was bedeutet das?

Im Vergleich zu anderen männlichen Testteilnehmern, die im Durchschnitt einen Score von +2,24 erreichen, liegt ihr Ergebnis deutlich darunter. Dies weist darauf hin, dass ihre Antworten denen von Frauen ähnlicher sind als denen von Männern. In früheren Studien wurde ein solcher Score häufig mit einem weniger männlichen Selbstbild in Verbindung gebracht. Die folgende Grafik zeigt Ihren Wert im Vergleich zur typischen Verteilung unter männlichen Teilnehmern:



English translation: Congratulations – you have completed the Gender Knowledge Test! This test was developed to measure differences in knowledge about gender-specific topics. Scoring is based

on a scale from −4 (strongly feminine) to +4 (strongly masculine). Your score is: −1.83. Compared to other male participants, who on average achieve a score of +2.24, your result is significantly lower. This indicates that your answers are more similar to those of women than to those of men.

Control Condition.

> **Glückwunsch – du hast den Geschlechtsspezifischen Wissenstest abgeschlossen!**
> Dieser Test wurde entwickelt, um Unterschiede im Wissen über geschlechtsspezifische Themen zu messen. Frühere Studien zeigen, dass bestimmte Wissensbereiche stärker mit traditionell männlich oder weiblich geprägten Rollen und Selbstbildern verbunden sind. Die Auswertung basiert auf einer Skala von −4 (ausgeprägt weiblich) bis +4 (ausgeprägt männlich).
> **Ihr Score beträgt: +2,18**
> **Was bedeutet das?**
> Ihr Score liegt nahe beim Durchschnitt männlicher Teilnehmer, der bei +2,24 liegt. Das bedeutet, dass ihre Antworten typisch für männliche Testteilnehmer sind. In früheren Studien wurde ein solcher Score häufig mit einem ausgeglichenen und typischen Wissensprofil in geschlechtsspezifischen Themenbereichen in Verbindung gebracht. Die folgende Grafik zeigt Ihren Wert im Vergleich zur typischen Verteilung unter männlichen Teilnehmern:

English translation: Congratulations – you have completed the Gender Knowledge Test! Your score is: +2.18. Your score is close to the average of male participants, which is +2.24. This means that your answers are typical for male test-takers.

Word Fragment Completion Task

For the current study, a German version of the Word Fragment Completion Task was developed consisting of ten aggressive and ten anxious word fragments. These fragments were created by identifying German target words semantically related to aggression (e.g., "Wut" [anger]) and anxiety (e.g., "Bangen" [to fear]) using a word association thesaurus (https://wordassociations.net/de/). Potential fragments were generated by removing the first letter of each target word, ensuring each could be completed as either the target word or a neutral alternative (e.g., "_UT" becoming "WUT" [anger] or "HUT" [hat]).

Aggressive Word Fragments

Items. Table A8 presents the aggressive cognition items.

Table A8

Word Fragment Completion Task: Aggressive Cognition Items

Fragment	Aggressive	Neutral
_UT	Wut	Mut, Hut, Gut, Tut
_ORN	Zorn	Dorn, Korn, Horn, Vorn
_IER	Gier	Bier, Tier, Vier, Hier, Pier
_AGE	Rage	Vage, Tage, Sage, Lage, Page, Jage, Gage, Zage
_UNDE	Wunde	Runde, Hunde, Munde, Kunde, Bunde
_IEB	Dieb, Hieb	Sieb, Lieb, Rieb
_ASS	Hass	Nass, Fass, Bass, Dass, Pass, Lass, Mass
_OBEN	Toben	Loben, Hoben, Roben
_AGEN	Jagen, Wagen	Sagen, Tagen, Magen, Ragen, Nagen, Lagen, Zagen, Hagen
_AMPF	Kampf	Dampf, Mampf

Note. Fragment = word stem presented to participants. Aggressive = completions coded as aggressive cognition. Neutral = completions coded as neutral.

Table A9 presents response frequencies for the aggressive word fragment completion items.

Table A9

Response Frequencies for Aggressive Word Fragment Completion Items

Item	Fragment	Negative	Positive	% Positive	Missing
1	_UT	139	7	4.8	8
2	_ORN	130	20	13.3	4
3	_IER	138	11	7.4	5
4	_AGE	136	12	8.1	6
5	_UNDE	146	5	3.3	3
6	_IEB	74	78	51.3	2
7	_ASS	129	14	9.8	11
8	_OBEN	83	57	40.7	14
9	_AGEN	110	32	22.5	12
10	_AMPF	51	90	63.8	13

Note. Negative = neutral word completion; Positive = aggressive word completion. % Positive calculated excluding missing responses.

Item-level response rates revealed substantial variability in aggressive completion rates. Several items had very low positive rates, including Item 5 (3.3%) and Item 1 (4.8%), while others exceeded 50%, such as Item 10 (63.8%) and Item 6 (51.3%). Items 8, 10, 9, and 7 each had more than 10 missing values, suggesting these fragments may have been more difficult or ambiguous. This imbalance compromises the scale's ability to discriminate between individuals.

The tetrachoric correlation (Figure 3) revealed weak and inconsistent correlations among items, including several negative correlations. If items measured the same latent construct, positive intercorrelations would be expected; this pattern suggests that the items do not form a coherent unidimensional construct. Item 7 was negatively correlated with the first principal component, indicating it may measure a different construct than the remaining items.

Anxious Word Fragments

Items. Table A10 presents the anxious cognition items.

Table A10

Word Fragment Completion Task: Anxious Cognition Items

Fragment	Anxious	Neutral
_ORGEN	Sorgen	Morgen, Borgen
_ANGEN	Bangen	Fangen, Wangen, Sangen, Rangen, Langen, Zangen
_EUE	Reue	Neue
_ROST	Trost	Prost, Frost
_AST	Last	Mast, Fast, Rast, Bast, Hast, Gast
_EERE	Leere	Beere, Heere, Teere
_ERN	Fern	Kern, Gern, Lern, Bern
_ESSEL	Fessel	Kessel, Sessel, Nessel
_ANISCH	Panisch	Manisch
_ANNE	Panne	Kanne, Tanne, Wanne

Note. Fragment = word stem presented to participants. Anxious = completions coded as anxious cognition. Neutral = completions coded as neutral.

Descriptive Statistics. Table A11 presents response frequencies for the anxious word fragment completion items.

Table A11

Response Frequencies for Anxious Word Fragment Completion Items

Item	Fragment	Negative	Positive	% Positive	Missing
1	_ORGEN	99	45	31.2	10
2	_ANGEN	140	5	3.4	9
3	_EUE	16	92	85.2	46
4	_ROST	107	38	26.2	9
5	_AST	94	46	32.9	14
6	_EERE	95	53	35.8	6
7	_ERN	121	23	16.0	10
8	_ESSEL	139	12	7.9	3
9	_ANISCH	15	90	85.7	49
10	_ANNE	120	24	16.7	10

Note. Negative = neutral word completion; Positive = anxious word completion. % Positive calculated excluding missing responses.

Item-level response rates revealed even greater variability than the aggressive items. Some items had very low positive rates, including Item 2 (3.4%) and Item 8 (7.9%), while others were completed with the anxious word by the vast majority of participants, such as Item 9 (85.7%) and Item 3 (85.2%). Notably, Items 3 and 9 also showed substantial missingness, with 46 and 49 missing responses respectively—approximately 30% of the sample. This pattern suggests these items may have been particularly difficult or ambiguous for participants.

Similar to the aggressive items, the tetrachoric correlation matrix (Figure 4) revealed weak and inconsistent correlations among items, including several negative correlations. Items 2, 7, and 8 were negatively correlated with the first principal component, suggesting they may not measure the same construct as the remaining items.

Manipulation Check and Suspicion Probe

Table A12 presents the items used to assess participants' recall of the experimental feedback and their suspicions about the study's purpose.

Table A12

Manipulation Check and Suspicion Probe Items and Response Options

Item	Question	Response Options
1	What feedback did you receive in this study regarding your knowledge of gender-specific topics?	1 (Typically feminine) to 10 (Typically masculine)
2	How would you rate your knowledge of gender-specific topics?	1 (Typically feminine) to 10 (Typically masculine)
3	During participation, did you have the impression that the study was about something other than what was stated?	Yes / No
4	What do you think this study was really about?	Open response
5	Do you have any thoughts about the study?	Open response

Note. German original questions and (response options):

1. Welche Rückmeldung haben Sie in dieser Studie zu Ihrem Wissen in geschlechtsspezifischen Themen erhalten? (1 = typisch weiblich bis 10 = typisch männlich).
2. Wie würden Sie Ihr Wissen in geschlechtsspezifischen Themen einschätzen? (1 = typisch weiblich bis 10 = typisch männlich).
3. Hatten Sie während der Teilnahme den Eindruck, dass es in der Studie um etwas anderes ging als angegeben? (Ja / Nein).

4. Was denken Sie, worum es in dieser Studie wirklich ging? (Offene Antwort).
5. Haben Sie Gedanken zur Studie? (Offene Antwort).

Item 4 was only displayed if participants responded "Yes" to Item 3.

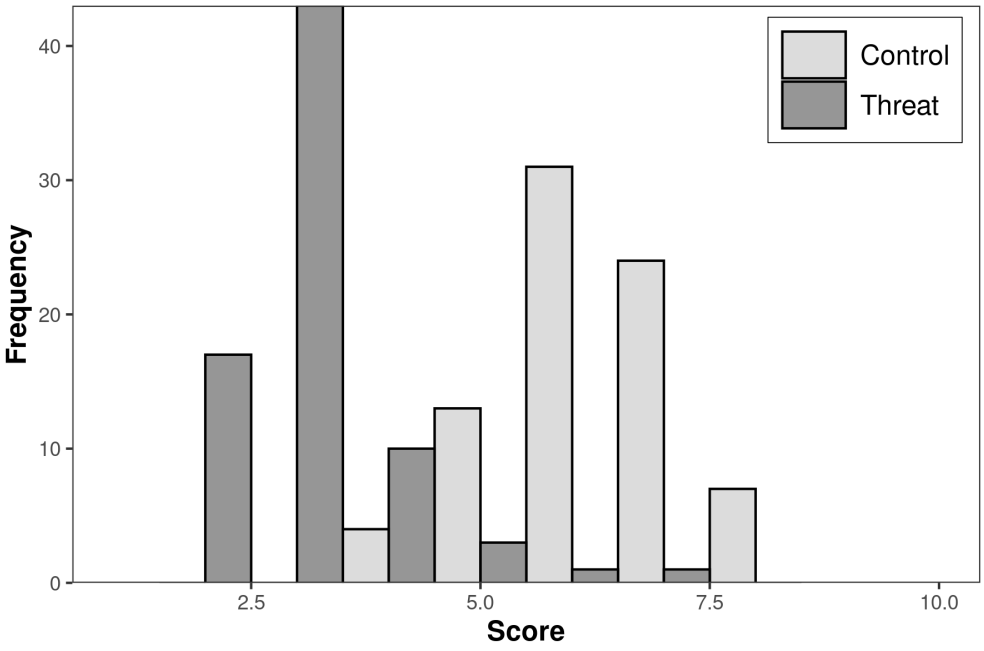
Success of Manipulation

To ensure the experimental manipulation was successful, participants were asked to recall the nature of the feedback they received on the Gender Knowledge Test (Table A12, Item 1). Responses were recorded on a scale ranging from 1 (*typically feminine*) to 10 (*typically masculine*). The results indicated that participants accurately recalled the feedback consistent with their assigned condition. Those in the threat condition correctly identified receiving more feminine feedback (M = 3.1, SD = 0.9), whereas participants in the control condition reported significantly higher scores, reflecting the more masculine feedback they received (M = 6.2, SD = 1).

As visualized in the Figure A1, the distribution of scores reveals two distinct, non-overlapping peaks. The threat condition is characterized by a high-frequency cluster at the lower end of the scale (around a score of 3), while the control condition shows a broader but clearly shifted distribution toward the higher end of the scale. This distinct bimodal pattern across groups confirms that the experimental manipulation was perceived as intended.

Figure A1

Histogram of Recalled Gender Knowledge Feedback



Note. Participants (N = 154) were asked to recall the feedback they received on the Gender Knowledge Test. Responses ranged from 1 (*typically feminine*) to 10 (*typically masculine*). Participants in the threat condition reported a more feminine feedback (M = 3.1, SD = 0.9) than participants in the control condition (M = 6.2, SD = 1).

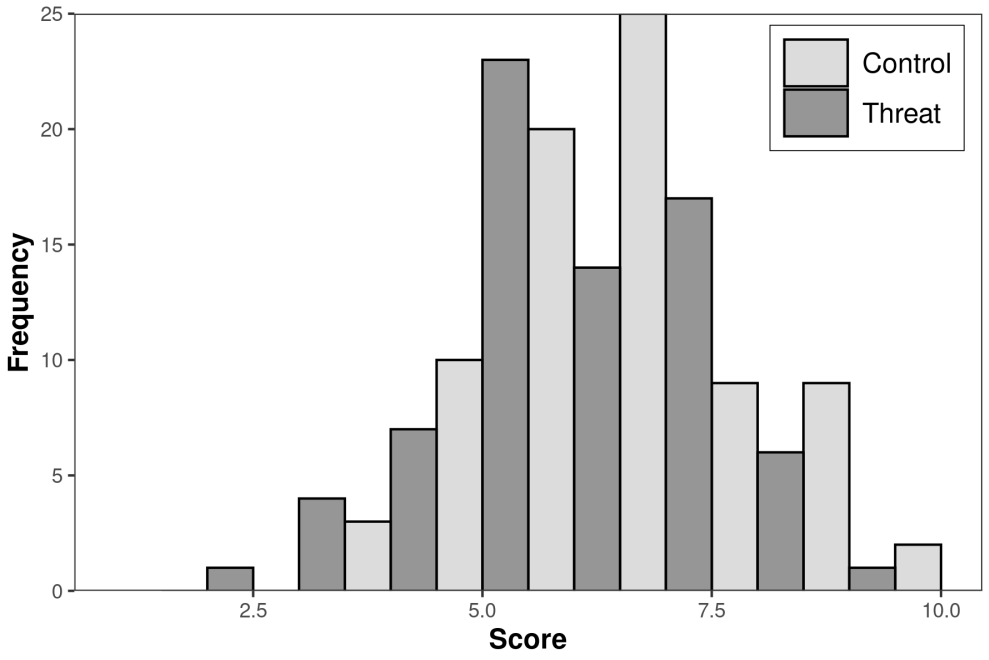
In addition to feedback recall, participants provided a self-evaluation of their own gender knowledge (Table A12, Item 2) using the same 10-point scale (1 = typically feminine, 10 = typically masculine). This measure served to assess the credibility of the experimental manipulation, determining whether the feedback actually shifted participants' internal perceptions of their gender-related expertise. Consistent with the intended effect, the masculinity feedback was perceived as credible. Participants in the threat condition reported lower self-evaluations (M = 5.7, SD = 1.4) compared to those in the control condition, who rated their knowledge as more masculine (M = 6.8, SD = 1.4).

The histogram for self-evaluated knowledge (Figure A2) reveals a more nuanced distribution than the recall measure. While both groups show a central tendency toward the middle of

the scale, there is a visible distributional shift between conditions. The threat condition's scores are clustered more densely between the 4 and 6 range, whereas the control condition shows a clear rightward shift, with a higher frequency of scores between 6 and 8. This overlap is expected for a self-evaluation measure, yet the distinct separation of the peaks confirms that the feedback successfully influenced participants' subjective appraisals of their gender knowledge.

Figure A2

Histogram of Self-Evaluated Gender Knowledge



Note. Participants (N = 154) were asked to self-evaluate their gender knowledge. Responses ranged from 1 (*typically feminine*) to 10 (*typically masculine*). Participants in the threat condition evaluted their gender knowledge as more feminine (M = 5.7, SD = 1.4) than participants in the control condition (M = 6.8, SD = 1.4).

Demographic Questionnaire

Table A13 presents the demographic items administered at the end of the study.

Table A13

Demographic Questionnaire Items and Response Options

Item	Question	Response Options
1	How old are you?	Open numeric entry
2	What is your gender?	Male, Female, Diverse, Prefer not to say
3	How proficient is your German?	None, Basic, Good, Very good, Native speaker
4	What is your highest educational degree?	No degree yet, No degree (left school), Lower secondary, Intermediate school, Technical college, University entrance (Abitur)
5	What is your highest occupational degree?	None yet, Apprenticeship, Vocational school, Technical school, University of applied sciences, University, Other

Note. German original questions and (response options):

1. Wie alt sind Sie?
2. Welches Geschlecht haben Sie? (Männlich, Weiblich, Divers, Keine Angabe).
3. Wie gut sind Ihre Deutschkenntnisse? (Keine Deutschkenntnisse, Grundkenntnisse, Gute Kenntnisse, Sehr gute Kenntnisse, Deutsch ist meine Muttersprache).
4. Was ist Ihr höchster Schulabschluss? (Noch kein Abschluss, Schule beendet ohne Abschluss, Haupt-/Volksschule, Mittlerer Schulabschluss, Fachhochschulreife, Abitur).
5. Welchen höchsten beruflichen Abschluss haben Sie? (Noch keinen, Lehre, Berufsfachschule, Fachschule, Fachhochschule, Universität, Anderer).

Debriefing

The following debriefing text was displayed to participants after completing all study components.

> **Aufklärung zum Studienzweck**
>
> Vielen Dank für Ihre Teilnahme!

Bevor Sie die Studie abschließen, möchten wir Sie über den tatsächlichen Hintergrund des Experiments informieren:

Wichtiger Hinweis: Das Feedback, das Sie im Rahmen des „geschlechtsspezifischen Wissenstests“ erhalten haben, war nicht echt. Es wurde unabhängig von Ihrer tatsächlichen Leistung zufällig generiert und ist nicht aussagekräftig für Ihre kognitive Leistungsfähigkeit oder Ihr Wissen.

Warum wurde das gemacht? Die Studie untersucht ein psychologisches Phänomen, das als „fragile Männlichkeit“ (engl. fragile masculinity) bekannt ist. Dabei geht es um die Frage, wie Männer auf Situationen reagieren, in denen ihre Geschlechtsidentität – also das Selbstbild als Mann – subtil infrage gestellt oder als „bedroht” erlebt wird, z. B. durch vermeintlich negatives Feedback in geschlechtsbezogenen Leistungsbereichen.

Wir interessieren uns dafür, wie sich solche Situationen auf Denken, Emotionen und Verhalten auswirken. Um dies zu untersuchen, haben wir den Teilnehmenden zufällig entweder neutrales oder negatives Feedback zum Wissenstest gegeben. Ihre emotionale Reaktion wurde anschließend mithilfe der Wortvervollständigungsaufgabe im „Test zur Denkgeschwindigkeit“ gemessen.

Die Annahme ist, dass sich emotionale Zustände (z. B. Ärger oder Verunsicherung) darin zeigen können, wie Personen unvollständige Wörter ergänzen.

Damit Ihre Reaktion möglichst authentisch und unverfälscht bleibt, war es notwendig, den wahren Zweck der Studie nicht vorab offenzulegen.

Wichtig: Sie sind nicht „hereingefallen“ Das verwendete Feedback wurde bewusst sehr glaubwürdig gestaltet – viele Teilnehmende hielten es ebenfalls für real. Ihre Reaktion ist verständlich und vollkommen normal. Die Täuschung diente ausschließlich

> wissenschaftlichen Zwecken und basiert auf etablierten Methoden der psychologischen Forschung.
>
> Auch wenn Ihnen nun bewusst ist, dass das Feedback nicht echt war, kann es dennoch nachwirken. In der Psychologie ist bekannt, dass selbst widerlegte Informationen unser Denken und Fühlen weiterhin beeinflussen können. Diese Reaktion ist normal und gut erforscht – bitte seien Sie daher nachsichtig mit sich, falls das Feedback bei Ihnen nachhallt.
>
> Wenn Sie Fragen, Bedenken oder Feedback zur Studie haben, können Sie sich jederzeit an die Studienleitung wenden:
>
> Studienleitung: Linus Widmer Institut für Psychologie, Humboldt-Universität zu Berlin widmerli@student.hu-berlin.de
>
> Wichtig: Bevor Sie das Fenster schließen lesen Sie bitte den unten stehenden Hinweis zur Freiwilligkeit und Einwilligungserklärung

English translation: Thank you for your participation! The feedback you received on the "Gender Knowledge Test" was not real—it was randomly generated regardless of your actual performance. The study investigates "fragile masculinity"—how men respond when their gender identity is subtly challenged. We randomly assigned either neutral or negative feedback and measured emotional reactions through the word completion task. You did not "fall for" anything; the feedback was designed to be believable, and your reaction is completely normal. Even though you now know the feedback was fake, it may still linger—this is a well-documented psychological phenomenon.

Consent After Debriefing

Following the debriefing, participants were given the option to withdraw their data.

> **Freiwilligkeit & Einwilligungserklärung**
>
> Hinweis zur Freiwilligkeit und Einwilligungserklärung:

> Da wir Sie zu Beginn der Studie nicht über alle Aspekte informieren konnten, möchten wir Ihnen nun die Möglichkeit geben, der Verwendung Ihrer Daten zu widersprechen. Ihre Entscheidung hat keinen Einfluss auf die zugesagte Spende an das Bundesforum Männer e.V. Wenn Sie widersprechen, werden keine Ihrer Antworten aus diesem Fragebogen gespeichert – mit Ausnahme der folgenden Informationen:
>
> - Ihre Entscheidung über die Datenspeicherung (Ja/Nein)
> - Welches Feedback Sie im Wissenstest erhalten haben (neutral oder negativ)

English translation: Since we could not inform you about all aspects of the study at the beginning, we would now like to give you the opportunity to object to the use of your data. Your decision will not affect the promised donation. If you object, none of your responses will be saved except your decision about data storage and which feedback condition you received.

Appendix B

B Sample Size Planning

To determine the appropriate sample size, four a priori power analyses were conducted using G*Power 3.1 (Faul et al., 2007). Additionally, the statistical complexity of the measurement model for the Motivation for Masculine Behavior scale was explored. Table B1 summarizes effect sizes from prior studies on masculinity threat.

Table B1

Effect Sizes for Masculinity Threat and Motivation of Masculine Behavior

Study	Dependent Variable	Effect Size	Note
Vandello et al. (2008), Study 4	anxious cognition	d = 0.55	-
Stanaland & Gaither (2021)	aggressive cognition	d = 0.41	-
Vandello et al. (2008), Study 5	aggressive cognition	d = 0.91	-
Stanaland et al. (2024)	aggressive cognition	d = 0.35	adolescents
Stanaland & Gaither (2021)	pressured motivation	f = 0.17	adolescents

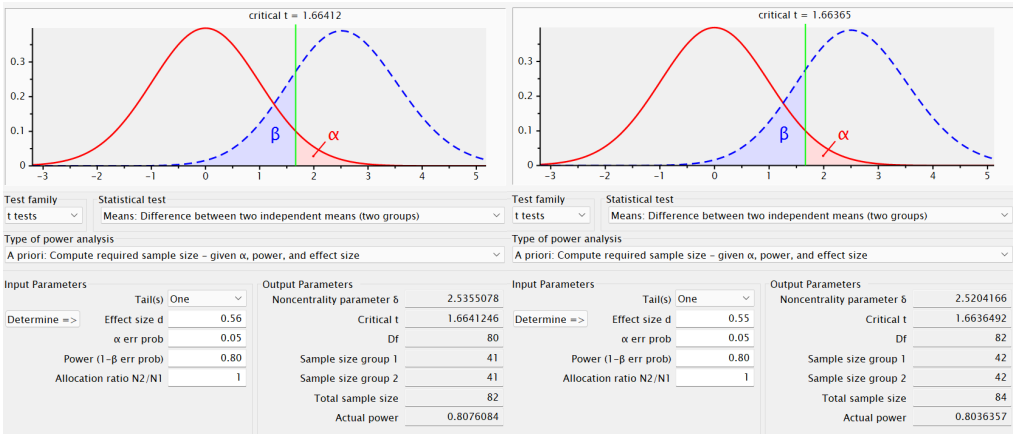
Note. Effect sizes for aggressive and anxious cognition induced by a masculinity threat. All studies measured aggressive and anxious cognition using a Word Fragment Completion Task and operationalized masculinity threat via negative feedback on the Gender Knowledge Test.

Masculinity Threat Effects

Previous studies have reported medium-sized effects of masculinity threat on aggressive cognition (d = 0.56) and anxious cognition (d = 0.55), both measured via the Word Fragment Completion Task. Based on these estimates, a priori power analyses were conducted using one-tailed independent samples t-tests (α = .05, power = .80). The analyses indicated required sample sizes of 82 participants for detecting effects on aggressive cognition and 84 participants for anxious cognition (see Figure B1).

Figure B1

Power Analysis for Masculinity Threat on Aggressive and Anxious Cognition



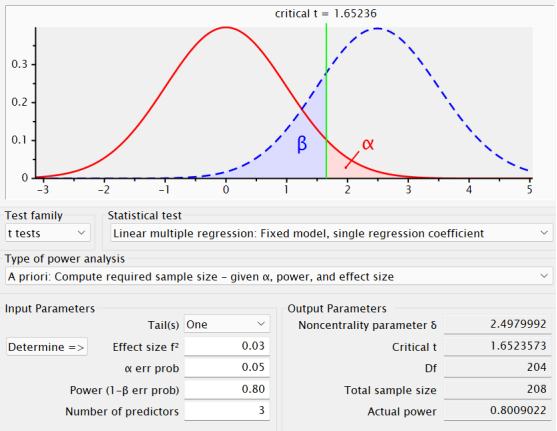
Note. Left = Masculinity threat on aggressive cognition; Right = Masculinity Threat on Anxious Cognition

Moderation by Motivation

For the moderation of pressured motivation on aggressive cognition, a small effect size was assumed (f = 0.17), based on prior findings on motivational moderators (Stanaland & Gaither, 2021). Due to the absence of previous studies on the moderation of autonomous motivation on anxious cognition, the same conservative small effect size (f = 0.17) was used. Based on a linear multiple regression model including three predictors (threat condition, autonomous or pressured motivation, and their two-way interaction), G*Power indicated a required sample size of 208 participants (α = .05, power = .80; see Figure B2).

Figure B2

Power Analysis for Masculinity Threat on Aggressive & Anxious Cognition



Factorial Structure of Motivation for Masculine Behavior

To evaluate the hypothesized two-factor structure of the Motivation for Masculine Behavior scale via Confirmatory Factor Analysis, sample size requirements were determined based on common recommendations for CFA model stability. Following conservative guidelines, a minimum of 200 participants to ensure stable estimation and generalizability of the factor structure, particularly when factors are correlated and loadings are moderate (Hoogland & Boomsma, 1998).

Final Target Sample Size

Based on the largest required sample size (moderation analysis; 208 participants) and accounting for an estimated 10% exclusion rate, the final target sample size was set at 229 participants.

Appendix C

C Results Supplement

The analysis code can be viewed publicly at (https://github.com/Linuswidmer/master-thesis). An anonymized version of the data is supplied in the same repository. Anonymization involved coding all free-text responses for study suspicion and then removing the free-text content, as it may contain information that could identify participants. The following sections provide additional statistic details about the sample and post-hoc power analysis.

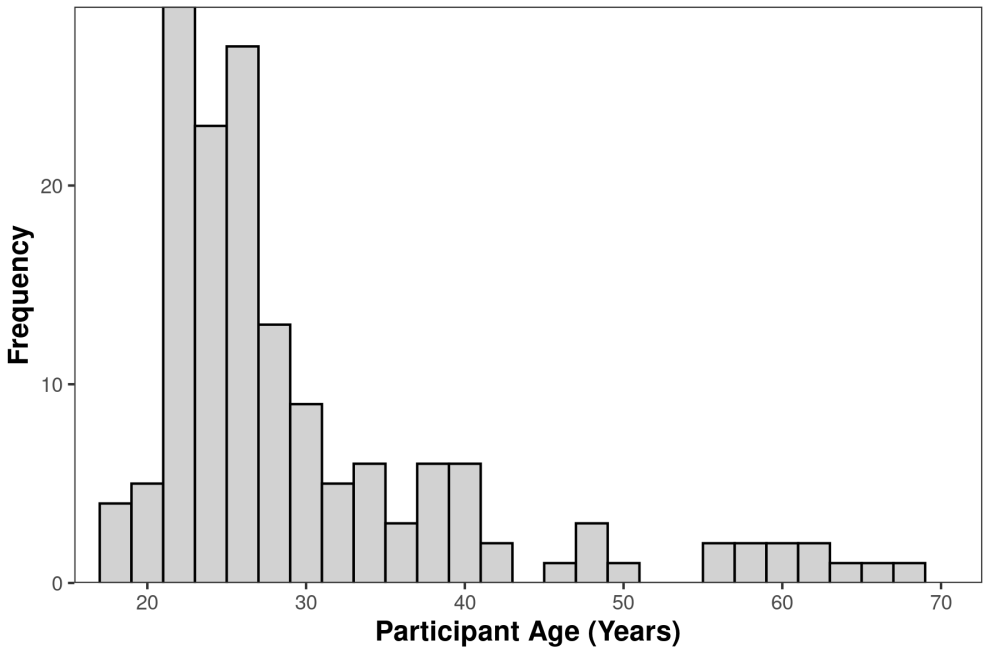
Sample Charateristics

Age

The final sample ranged in age from 18 to 69 years (M = 30.6, SD = 10.9). As illustrated in Figure C1, the distribution is positively skewed, with the majority of participants concentrated in the early-to-mid twenties. Frequency counts diminish steadily for older age brackets, though the sample remains representative of a broad adult developmental span.

Figure C1

Histogram of Participant Age



Note. Participant (N = 154) age in years (M = 30.6, SD = 10.9).

Educational Background

The educational background of the sample is summarized in Table C1. The participants were predominantly highly educated, with the vast majority holding a university entrance qualification (80.5%). An additional 14.3% had obtained a technical college entrance qualification, while the remaining 5.1% of the sample reported intermediate, lower, or no current secondary school degrees. This distribution indicates a sample that is significantly more academic than the general population, a factor considered in the interpretation of the results regarding psychological savvy and response patterns.

Table C1

Educational Background of the Sample

Education Level	Count (n)	Percent (%)
University Entrance Qualification	124	80.5
Technical College Entrance Qualification	22	14.3
Intermediate School Leaving Certificate	6	3.9
Lower Secondary School Certificate	1	0.6
No Degree Yet / Student	1	0.6

Note. N = 154. German original categories were: University Entrance Qualification = Abitur, allgemeine oder fachgebundene Hochschulreife; Technical College Entrance Qualification = Fachhochschulreife / Fachoberschule; Intermediate School Leaving Certificate = Mittlerer Schulabschluss / Realschule; Lower Secondary School Certificate = Haupt- / Volksschule; No Degree Yet = Noch kein Abschluss / Schüler.

Occupational Background

The occupational background of the participants is detailed in Table C2. Consistent with the educational data, the sample was characterized by a high level of academic achievement, with

60.4% of participants holding a university degree. A significant portion of the sample reported having no occupational degree yet (20.8%), likely reflecting the high proportion of students within the participant pool. Vocational and technical qualifications, including apprenticeships and technical academy degrees, accounted for the remaining 18.8% of the sample.

Table C2

Occupational Background of the Sample

Occupational Degree	Count (n)	Percent (%)
University Degree	93	60.4
No Occupational Degree Yet	32	20.8
Vocational or Commercial School	13	8.4
Technical Academy (Master/Technician)	7	4.5
Apprenticeship (Dual System)	5	3.2
University of Applied Sciences	4	2.6

Note. N = 154. German original categories were: University Degree = Universität / Hochschule; No Occupational Degree Yet = Noch keinen Berufsabschluss; Vocational or Commercial School = Berufsfachschule / Handelsschule; Technical Academy = Fachschule (Meister-, Technikerschule, Berufs- oder Fachakademie); Apprenticeship = Lehre (beruflich-betrieblich); University of Applied Sciences = Fachhochschule / Ingenieurschule.

Post-hoc Power Analysis

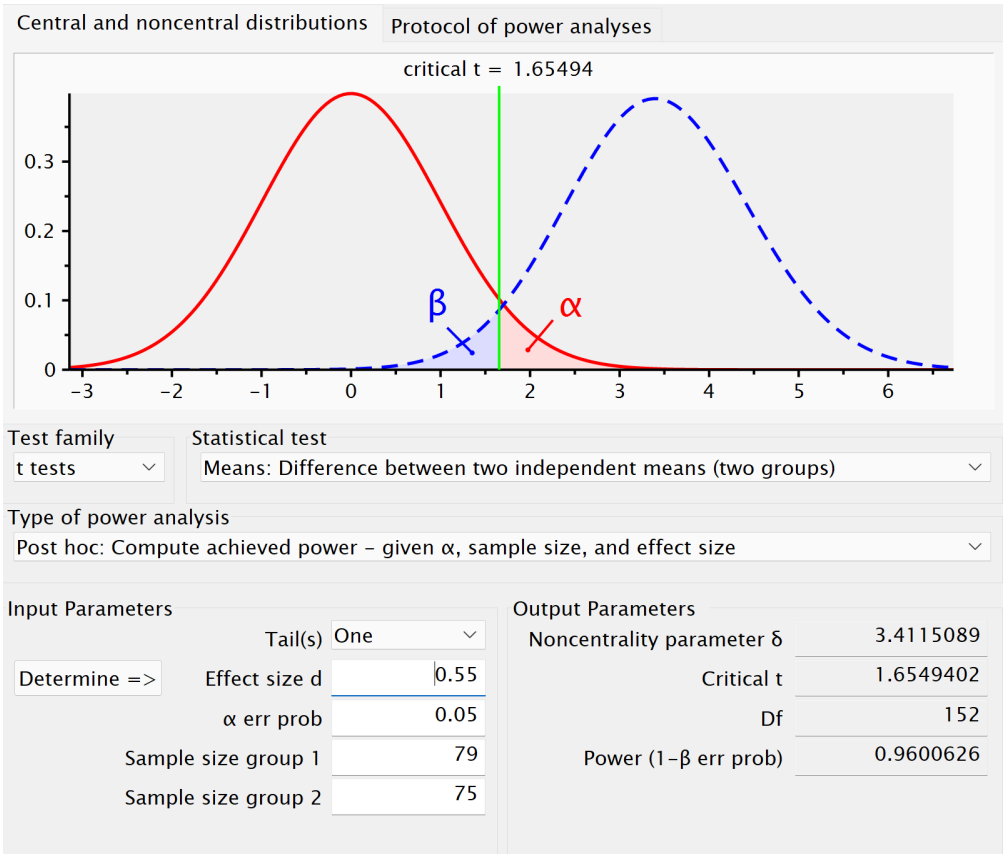
Post-hoc power analysis were conducted using G*Power 3.1 (Cunningham & McCrum-Gardner, 2007) to determine the achieved test power.

Masculinity Threat Effects

A post-hoc power analysis indicated that the study achieved a power of $1 - \beta = .96$ to detect the expected medium effect ($d = 0.55$) of the masculinity threat (Figure C2).

Figure C2

Post-hoc Power Analysis for Masculinity Threat Effects



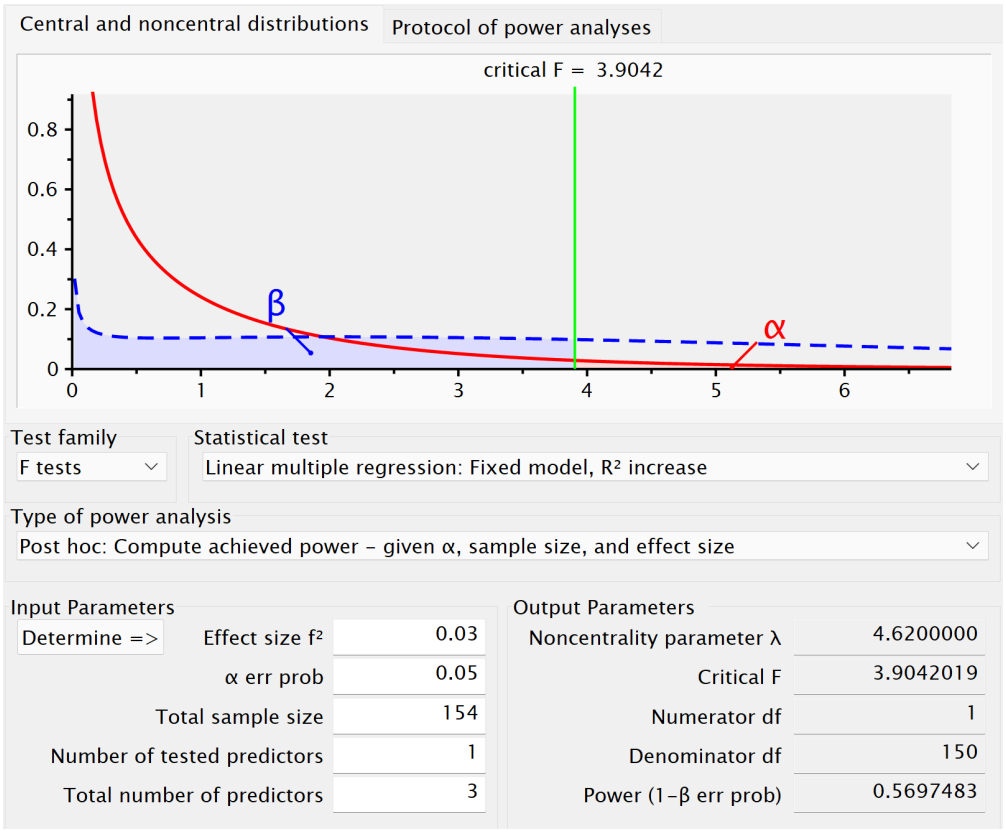
Note. G*Power output for post-hoc power analysis. Test family: t tests; Statistical test: Means: Difference between two independent means (two groups). With $n_1 = 79$, $n_2 = 75$, $\alpha = .05$, assumed effect size $d = 0.55$. Post-hoc achieved power was $1 - \beta = .96$.

Moderation of Masculinity Threat Effects

A post-hoc power analysis indicated that the study achieved a power of $1 - \beta = .56$ to detect the expected smalle effect ($f = 0.17$) of the masculinity threat (Figure C3).

Figure C3

Post-hoc Power Analysis for Moderation of Masculinity Threat Effects



Note. G*Power output for post-hoc power analysis. Test family: F tests; Statistical test: Linear Multiple Regression: Fixed Model, R^2 increase. With $n = 154$, $\alpha = .05$, number of tested predictors = 1, number of total predictors = 3 assumed effect size $f = 0.17$. Post-hoc achieved power was $1 - \beta = .56$.

Appendix D

D List of Appendices