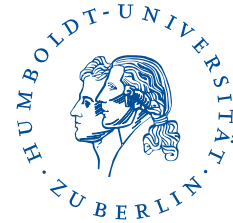


HUMBOLDT-UNIVERSITÄT ZU BERLIN



**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**

2026-01-03

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**

Anxiety, and aggression are common strains experienced by men when failing to conform to rigid masculinity norms (O'Neil, 2008). The precarious manhood thesis attributes this strain partially to the tenuous and elusive nature of manhood (Vandello et al., 2008; Vandello & Bosson, 2013). . A novel identity-based model of masculinity aims to explain the adverse effects associated with the precariousness of manhood (Stanaland et al., 2023). This model suggests that when a man's masculinity status is threatened, his compensatory response is influenced by differences in experienced self-discrepancy. The current thesis aims to operationalize and test this model within a German cultural context. Masculinity threat is induced through negative intra-group feedback on gender knowledge. Internalized and externalized threat responses are measured using a word fragment completion task. To assess different discrepancies in the masculine self, participants report both internal and external pressures to conform to masculinity norms. In line with Stanaland et al. (2023), it is hypothesized that external pressure to be masculine predicts externalized threat responses (aggression), while internal pressure predicts internalized threat responses (anxiousness).

*Keywords:* Fragile Masculinity, Identity Threat, Motivation

## Contents

1	Introduction .....	6
1.1	Masculinity Norms and Their Psychological Cost .....	6
1.2	Masculinity as a Precarious Social Status .....	7
1.3	A Motivational Framework for Masculinity Threats .....	8
1.4	The Present Research .....	10
1.4.1	Cross-Cultural Generalizability of Masculinity Threat Effects .....	10
1.4.1.1	Hypothesis 1a .....	11
1.4.1.2	Hypothesis 1b .....	11
1.4.2	Testing Motivational Predictors of Threat Response .....	12
1.4.2.1	Hypothesis 2a .....	12
1.4.2.2	Hypothesis 2b .....	13
1.4.3	Addressing Methodological Limitations .....	13
2	Methods .....	13
2.1	Ethical Considerations and Pre-Registration .....	13
2.2	Sample and Recruitment .....	13
2.3	Procedure .....	14
2.4	Materials .....	15
2.4.1	Motivation for Masculine Behavior .....	15
2.4.2	Gender Knowledge Test .....	15
2.4.3	Masculinity Threatening and Non-Threatening Feedback .....	16
2.4.4	Word Fragment Completion Task .....	16
2.4.5	Manipulation Check and Suspicion Probe .....	17
2.4.6	Demographic Questionnaire .....	17
2.5	Research Design .....	17

2.6 Data Analysis .....	17
3 Results .....	19
3.1 Participants .....	19
3.2 Motivation For Masculine Behaviour Validation .....	19
3.3 Masculinity Threat Effects .....	23
3.3.1 Masculinity Threat on Aggressive Cognition (Hypothesis 1a) .....	23
3.3.2 Masculinity Threat on Anxious Cognition (Hypothesis 1b) .....	23
3.4 Moderation of Masculinity Threat Effects .....	24
3.4.1 Pressured Motivation on Aggressive Cognition (Hypothesis 2a) .....	25
3.4.2 Autonomous Motivation on Anxious Cognition (Hypothesis 2b) .....	26
3.5 Experimental Manipulation Check .....	26
4 Discussion .....	27
4.1 Motivation for Masculine Behaviour in a German Sample .....	27
4.1.1 Factorial Structure of the MMB Scale .....	28
4.1.2 Sample considerations/translation .....	29
4.1.3 Scale construction .....	29
4.1.4 Conclusion on the MMB in a German sample .....	30
4.2 Masculinity Threat Effects .....	30
4.3 Measuring (Masculinity) Threat Response (or something like that) .....	32
4.4 Theory Discussion .....	34
4.5 Conclusion .....	35
References .....	36
Declaration of Independence .....	41
A Materials Supplement .....	44
Study Introduction .....	44

Motivation for Masculine Behavior .....	45
Documentation of the Translation of the Motivation for Masculine Behaviour Scale . .	46
Gender Knowledge Test .....	50
Documentation of the Translation of the Gender Knowledge Test .....	54
Masculinity Threatening and Non-Threatening Feedback .....	64
Threat Condition .....	64
No-Threat Condition .....	66
Word Fragment Completion Task .....	67
Manipulation Check and Suspicion Probe .....	68
Demographic Questionnaire .....	69
Debriefing .....	70
Consent After Debriefing .....	72
Sample Size Planning .....	73
Masculinity Threat Effects .....	73
Moderation by Motivation .....	73
Factorial Structure of Motivation for Masculine Behavior .....	74
Final Target Sample Size .....	74
B Results Supplement .....	76
Motivation for Masculine Behaviour .....	76
Word Fragment Completion Task: Detailed Psychometric Properties .....	76
Aggressive Word Fragment Items .....	77
Anxious Word Fragment Items .....	77
Sensitivity Power Analysis .....	78
C List of Appendices .....	81

## **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, ethnicity, 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 — offering stability amid the chaos and unpredictability of life. Yet, when these guardrails become too narrow, too rigid, they risk becoming barriers. Meant to keep us safe, they can instead steer us toward collision, limiting growth and distorting our path.

### **1.1 Masculinity Norms and Their Psychological Cost**

One particularly rigid and influential set of social norms are those associated with masculinity. From an early age, boys internalize anti-femininity norms, such as the belief that “boys shouldn’t 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). Further 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 its cultural background, they all shape male behavior and experience (Lease et al., 2013).

Conforming to rigid masculinity norms carries significant psychological and social costs. Attempting — and potentially failing — to meet these expectations 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, status-seeking and dominance-related pressures contribute to aggressive behaviors like sexual violence (Rando et al., 1998).

The psychological strain that occurs when a man’s personal needs, values, or emotions conflict with the narrow range of behaviors deemed socially acceptable for men is referred to

as **men's gender role conflict** (O'Neil et al., 1986). Although the negative consequences of this conflict are well documented (O'Neil, 2008), the processes through which this strain develops remain poorly understood. To understand why the costs of masculinity norms are so persistent, it is crucial to consider how masculinity is not just a set of traits, but a social status that must be earned and maintained.

## 1.2 Masculinity as a Precarious Social Status

“Men are made, not born”: the idea that masculinity functions as a socially conferred 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 and biological outcome — manhood is viewed as a status that must be earned, actively maintained, and socially validated (precarious manhood thesis; Vandello & Bosson, 2013). This structural fragility becomes especially apparent when men face threats to their masculinity, often eliciting *internalized* responses such as shame and anxiety (Vandello et al., 2008), or *externalized* responses such as aggression and dominance (Bosson et al., 2009). Men's aggressive threat response has led to the 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.

The precarious nature of masculinity bears significant implications not only for individuals but also for broader society. 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).

While the precarious manhood thesis offers a compelling account of why masculinity threats provoke defensive behaviors, it also reveals a divergence in the nature of these responses. Both anxious withdrawal and aggressive assertion appear to be plausible outcomes -

consistent with the psychological strain men experience in trying to meet rigid gender expectations and the need to repeatedly prove one's masculinity through dominant behavior. However, the factors that determine whether a threat leads to internalized or externalized reactions remain unclear. To address this gap, **it is critical to examine the underlying motivations for conforming to masculine norms**, which may help explain why men respond to threats in different ways.

### **1.3 A Motivational Framework for Masculinity Threats**

Although stereotypically masculine behaviors may appear uniform on the surface, the motivations behind them can differ significantly. For instance, one man might work out because he genuinely enjoys the activity and finds it personally meaningful, reflecting *autonomous motivation*. Another might exercise primarily because he feels it is expected of him and fears rejection, demonstrating *pressured motivation*. This distinction in motivations for masculine behavior has important implications for how men respond to threats to their masculine identity. Men with higher pressured motivation to conform to masculinity norms respond to such threats with increased aggression (Stanaland & Gaither, 2021). This interaction could not be explained via men's autonomous motivation or their collective self-esteem — that is, the degree to which they derive pride and self-worth from identifying as a man.

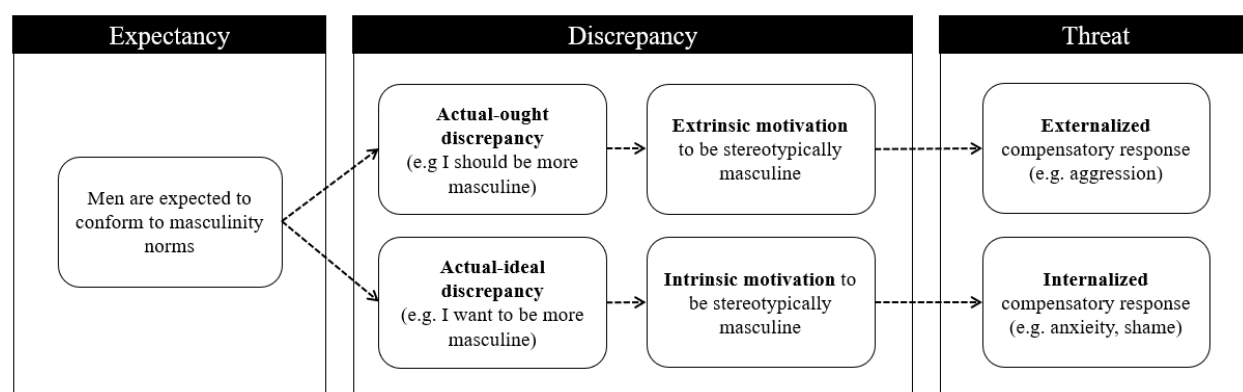
Building on this evidence, Stanaland et al. (2023) introduced the expectancy-discrepancy-threat (EDT) model of masculine identity (Figure 1). According to this model, boys and men are socialized into rigid masculinity norms and expectations. Drawing from self-discrepancy theory (Higgins, 1987) the EDT model posits that in response to this socialization process 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). In the case that men's actual-selves do not align with their ideal or ought selves — as in the context of a masculinity threat — they may experience two types of



identity discrepancies: one between who they are and who they believe they should be (actual-ought), and another between who they are and who they aspire to be (actual-ideal). Drawing from self-determination theory (Deci & Ryan, 1987) the EDT model suggests that an actual-ought discrepancy generates extrinsic motivation to reduce the tension which is associated with externalized responses to threat, such as aggression. On the other hand, an actual-ideal discrepancy produces intrinsic motivation that tends to result in internalized reactions, such as shame or anxiety. Importantly, men are unlikely to experience only one type of discrepancy; rather, they often experience both forms to different extents, depending on their underlying motivations to conform to masculinity norms.

**Figure 1**

*The Expectancy-Discrepancy-Threat Model of Masculine Identity*



*Note.* Men are socialized to conform to societal norms of masculinity. Different types of self-discrepancies—between the actual self (how one perceives oneself), the ideal self (who one wishes to be), and the ought self (who one believes they should be)—can give rise to distinct motivational processes. When masculinity is threatened, an actual-ought discrepancy is thought to elicit 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. anxiousness).

The EDT model of masculine identity offers a compelling framework for understanding the diverse emotional and behavioral responses men exhibit when their masculinity is challenged. However, despite its theoretical richness, empirical research testing the model remains limited, particularly in diverse cultural contexts. To address these gaps, the present study seeks to provide further empirical support for the EDT model by examining how men respond to masculinity threats and how these responses are shaped by underlying motivational factors.

## 1.4 The Present Research

By investigating masculinity threat responses in a novel cultural context using a theoretically grounded motivational framework, the present research makes several key contributions. First, it extends the EDT model by assessing its cross-cultural generalizability. Second, it addresses limited empirical evidence for the pathways proposed by the EDT model. Third, it directly tests the model's core predictions using implicit cognitive measures that capture both internalized and externalized threat responses. Through this multifaceted approach, the study aims to clarify the psychological processes that underlie fragile masculinity and deepen our understanding of how and why men respond differently when their gender identity is challenged.

### 1.4.1 Cross-Cultural Generalizability of Masculinity Threat Effects

Most research on fragile masculinity and the precariousness of manhood has been conducted in the United States. This raises concerns about the generalizability of these findings to other cultural contexts with differing gender norms and expectations. For example, the United States — particularly rural, Southern regions — is often described as an individualistic and “honor”-based culture that simultaneously emphasizes and challenges men's masculinity (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 (Stanaland, 2022).

fragile masculinity is only the externally motivated path of threat reaction

In support of the idea that masculine identity fragility also manifests outside the United States are widespread beliefs about the precariousness of manhood (i.e. the belief that manhood has to be earned and publicly demonstrated) including Germany and other European nations (Bosson et al., 2021). This is further supported by Polish and Norwegian men, who responded to masculinity threats with increased aggressive cognition (Valved et al., 2021) and reduced support for gender equality (Kosakowska-Berezecka et al., 2016). Notably, cross-cultural differences in threat responses appear to correspond with broader societal patterns: countries with higher gender inequality and lower human development scores tend to show stronger beliefs in the precariousness of manhood (Bosson et al., 2021). This is consistent with the stronger threat response observed in Polish men compared to Norwegian men (Valved et al., 2021).

Taken together, the prevalence of precarious manhood beliefs and empirical support from European populations make it plausible to expect compensatory responses to a masculinity threat in the present study. Although the present research is conducted in Germany – a country with relatively high gender equality (Rank 7/146; World Economic Forum, 2024) – it is expected that threat effects still emerge, albeit potentially with smaller effect sizes. This expectation is supported by prior findings showing that in Norway, a country with even higher levels of gender equality (Rank 3/146; World Economic Forum, 2024), men still exhibited increased aggressive cognition following a masculinity threat (Valved et al., 2021). Accordingly, it is hypothesized:

#### **Hypothesis 1a.**

Men exposed to a masculinity threat will exhibit higher levels of aggressive cognition than men receiving neutral feedback.

**Hypothesis 1b.** Men exposed to a masculinity threat will exhibit higher levels of anxious cognition than men receiving neutral feedback

a more  
meaningful  
heading?

### ***1.4.2 Testing Motivational Predictors of Threat Response***

To explain inter-individual differences in men's responses to masculinity threats, the present research aims to test the EDT model of masculine identity by applying its predictions in a novel cultural context. Although theoretically compelling, empirical support for the EDT model remains limited. Moreover, the model faces conceptual challenges from alternative frameworks that propose a sequential relationship between internalized and externalized threat responses - suggesting, that men may initially experience internal distress (e.g., shame or anxiety), which then escalates into outward aggression (Vescio et al., 2025). Here, externalized responses serve as a compensation to internal distress and is rooted in research that the expression of anger relieves feelings of discomfort (Jakupcak et al., 2005). These competing perspectives highlight the need for rigorous empirical testing of the EDT model's predictions.

According to the EDT model, pressured motivation - the external pressure to conform to masculine norms - is associated with externalized responses to masculinity threats, such as increased aggressive cognition. This pattern has been observed in both adult (Stanaland & Gaither, 2021) and adolescent (Stanaland et al., 2024) samples. Thus it is hypothesized

**Hypothesis 2a.** External pressure to conform to masculinity norms increases the positive effect of a masculinity threat (compared to a neutral feedback) on aggressive cognition.

Conversely, the EDT model posits that autonomous motivation is linked to internalized responses to masculinity threats, such as anxiety or shame. However, this prediction has not yet been empirically tested. Prior research has focused mainly on externalized responses, particularly aggression, leaving a gap in our understanding of how masculinity threats may also elicit internal distress. The present study addresses this gap by examining whether autonomous motivation predicts anxious cognition in response to a masculinity threat. Specifically it is hypothesized:

**Hypothesis 2b.** Internal pressure to conform to masculinity norms increases the positive effect of a masculinity threat (compared to a neutral feedback) on anxious cognition.

### ***1.4.3 Addressing Methodological Limitations***

Finally, methodological challenges persist in studying masculinity threat and the EDT model. While prior research has often operationalized masculinity threat responses through observable behaviors (e.g., aggression, dominance, social withdrawal), this study focuses on cognitive activation-specifically, aggressive and anxious cognition. This approach offers two key advantages. First, internalized emotional responses such as shame or anxiety are often incongruent with dominant masculinity norms and may be consciously downplayed or denied in self-report or behavioral tasks. Second, aggressive actions may sometimes reflect socially learned scripts rather than genuine emotional reactivity. Implicit cognitive tasks, such as the Word Fragment Completion Task (WFCT), are less susceptible to social desirability bias and provide a more direct measure of underlying threat-relevant mental processes. Thus, assessing cognitive activation allows for a more sensitive and nuanced examination of how masculinity threats are internally processed.

## **2 Methods**

### **2.1 Ethical Considerations and Pre-Registration**

Prior to data collection, ethical considerations were addressed 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 University of Berlin (reference number: 2025-41).

For transparency, the study's aims, hypotheses, methods, and data analytic plan 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 Recruitment**

A priori power analysis based on effect sizes from previous studies indicated a required sample size of 209 participants (see Appendix A, Sample Size Planning). The target sample size was set at 229 participants (10% higher to account for attrition). Eligibility criteria included self-identification as male, native-level proficiency in German, and a minimum age of 18 years. Participants were recruited via flyering in public spaces or an online study exchange portal (<https://surveycircle.com>). In exchange for participation, 1€ was donated to a public organization supporting men's counselling and public interests.

### **2.3 Procedure**

The procedure was adapted from Stanaland & Gaither (2021) into an online questionnaire and translated into German. Participants accessed the study via a link on the study exchange portal or a QR code on the flyers.

Deception regarding the study's true purpose was crucial to the experimental design. Participants were therefore informed that the study concerned personality, motivation, and performance ("Persönlichkeit, Motivation & Leistung") in men. They were reminded that eligibility required identifying as male, native-level proficiency in German, and a minimum age of 18 years. After providing informed consent, participants completed the nine randomized items of the MMB scale, presented as attitudes towards masculinity ("Einstellungen zum Thema Männlichkeit").

Next, participants were informed they would complete a test measuring knowledge in areas typically associated with men or women in society ("Geschlechtsspezifischer Wissenstest"). The items of the GKT were presented in a fixed randomized order.

Upon completion, participants received either (a) threatening feedback indicating worse performance than the average man, or (b) non-threatening feedback indicating performance comparable to the average man. This feedback was not based on participants' actual responses but was randomly assigned by experimental condition.

Following the feedback, participants were told they would take a test measuring their speed of thought (“Test zur Denkgeschwindigkeit”). In reality, they completed the 20 items of the WFCT. They were instructed to complete each word fragment by typing the complete word. Items were presented in randomized order, with a seven-second time limit per item to encourage spontaneous, intuitive responses. Participants then completed a demographic questionnaire, suspicion questions, and the manipulation check.

Upon completing all study components, participants received a full debriefing. This included an explanation of the study’s true purpose and a clarification that the feedback provided during the GKT was fabricated and randomly assigned. The rationale for the deception was explained, and participants were offered the option to withdraw their data.

## **2.4 Materials**

This section describes the measures and materials used in the study. The MMB scale and GKT were translated from English to German following the Translation, Review, Adjudication, Pretesting, and Documentation (TRAPD) protocol (Harkness et al., 2010). Complete materials including all item wordings in both languages and documentation of the translation process are provided in the Appendix.

### ***2.4.1 Motivation for Masculine Behavior***

The MMB scale was taken from Stanaland et al. (2023) Study 2. It consists of nine items measuring participants’ motivation for enacting masculine behavior. Responses were recorded on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The scale comprises two subscales: Pressured Motivation (5 items; e.g., “In general, I’m masculine because I want others’ acceptance and approval”) and Autonomous Motivation (4 items; e.g., “I enjoy being masculine”). See Table A1 for all items with German translations.

### ***2.4.2 Gender Knowledge Test***

The GKT, originally developed by Rudman & Fairchild (2004), was used in a version culturally adapted for European participants by Valved et al. (2021). Thirty items covering stereotypically masculine (e.g., sports, combat, home repair) and feminine (e.g., cooking, childcare, fashion) knowledge domains were included. Items were designed with moderate to high difficulty to support the credibility of the subsequent bogus feedback. Each item presented 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”). See Table A4 for all items with German translations.

#### ***2.4.3 Masculinity Threatening and Non-Threatening Feedback***

To administer the masculinity threat, participants were shown a diagram depicting a fictitious distribution of scores on the GKT. In the threat condition, participants were told they scored  $-1.83$ , substantially below the purported male average of  $+2.24$ . This feedback was accompanied by a statement that their answers were “more similar to women than to men.” In the no-threat condition, participants were told they scored  $+2.18$ , close to the male average, indicating “typical” performance for male participants. Both conditions included a visual distribution graph highlighting the participant’s position relative to other men (see Appendix for screenshots).

#### ***2.4.4 Word Fragment Completion Task***

The WFCT was used to assess aggressive and anxious cognition. Word fragments were developed by first identifying German words semantically related to aggression (“Wut” [anger], “Ärger” [frustration]) and anxiety (“Angst” [fear], “Furcht” [fear]) using the a word association thesaurus (<https://wordassociations.net/de/>). For each target word, potential word fragments were generated by removing the first letter. Each fragment could be completed as either a target word (aggressive or anxious) or a neutral word. For example, “\_UT” could become “WUT” (anger; aggressive) or “HUT” (hat; neutral). Similarly, “\_ANGEN” could become



“BANGEN” (to fear; anxious) or “WANGEN” (cheeks; neutral). See Table A7 and Table A8 for all word fragments and their possible completions.

#### **2.4.5 Manipulation Check and Suspicion Probe**

Two manipulation check items verified that participants correctly perceived the experimental manipulation. First, participants recalled the type of feedback they received on the GKT on a scale from 1 (*typically feminine*) to 10 (*typically masculine*). Second, they provided a self-evaluation of their own gender knowledge on the same scale. 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. Participants were also given the opportunity to share any other thoughts about the study in an open-text field. See Table A9 for all items with German translations.

#### **2.4.6 Demographic Questionnaire**

Participants reported their age, gender, German language proficiency, highest educational degree, and highest occupational degree. See Table A10 for all items with translations.

### **2.5 Research Design**

The study employed a between-subjects experimental design with random assignment to one of two conditions: masculinity threat (threatening feedback) or no-threat (non-threatening feedback). Two dependent variables were assessed: aggressive cognition and anxious cognition, each operationalized as the proportion of target word completions on the respective WFCT subscales. MMB was measured prior to the experimental manipulation and examined as a moderator. Specifically, pressured motivation was hypothesized to moderate the effect of threat on aggressive cognition, whereas autonomous motivation was hypothesized to moderate the effect of threat on anxious cognition.

### **2.6 Data Analysis**

All analyses were conducted using R version 4.5.2 (R Core Team, 2023). To validate the MMB scale, confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) were used to assess item structure. Model fit was evaluated against the following criteria: comparative fit index (CFI) > .95, Tucker-Lewis Index (TLI) > .95, root mean square error of approximation (RMSEA) < .06, and standardized root mean square residual (SRMR) < .08.

Hypothesis testing was conducted on aggressive and anxious word fragment completion scores 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 violations of the homoscedasticity assumption on transformed scores. To address this, the analyses employed a generalized linear model (GLM) with a binomial family 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. Prior to the main analyses, Levene's tests assessed homogeneity of variance across conditions. Model diagnostics were evaluated using simulated residuals via the DHARMA package (Hartig, 2024). Influential observations may distort the validity of results (Aguinis et al., 2013). To address this, a sensitivity analysis identified cases exceeding a Cook's distance threshold of  $\frac{4}{n-k}$ , removed them, and refitted the model to confirm the stability of findings.

A significance level of  $\alpha = .05$  was applied throughout. To evaluate the detectability of effects, sensitivity analyses were conducted using G\*Power 3 (Cunningham & McCrum-Gardner, 2007). Sensitivity analysis determines the minimum effect size detectable with 80% power at  $\alpha = .05$ , avoiding the limitations of post-hoc observed power (Hoenig & Heisey, 2001). If the observed effect is substantially smaller than the minimum detectable effect, this suggests absence of a meaningful effect rather than insufficient statistical power.

### 3 Results

#### 3.1 Participants

A final sample of 196 participants was collected over the course of two months. This missed the goal of 229 participants from power analysis.

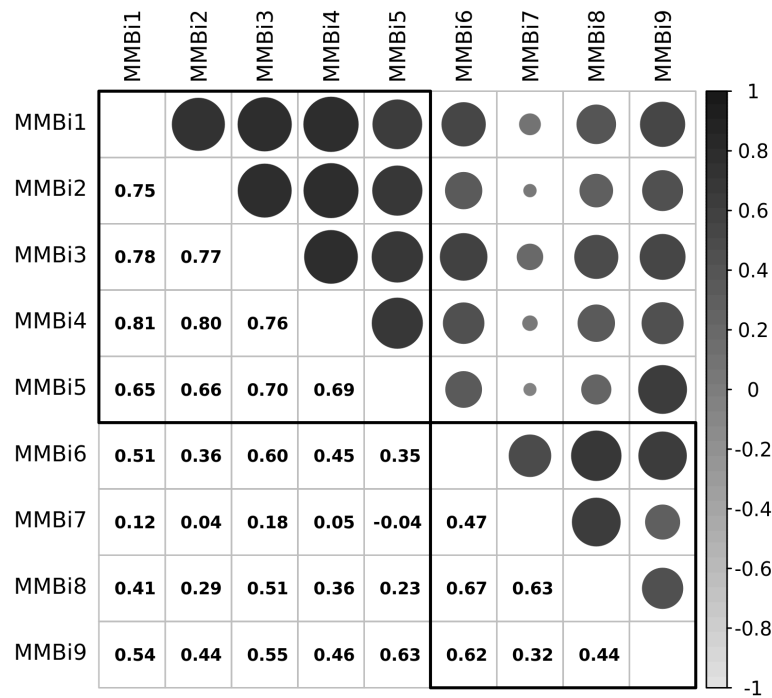
Participants were excluded based on pre-registered criteria. First, 8 participants withdrew consent after debriefing. Of the remaining sample, 13 participants did not meet eligibility requirements (9 reported a gender other than male, 3 were under 18 years of age, and 1 reported insufficient German proficiency). Furthermore, 9 participants were excluded for expressing strong suspicion about the study's true purpose (open-text responses were independently coded, and participants who correctly identified the masculinity threat manipulation were excluded), 11 for completing fewer than 50% real words on the WFCT, and 1 for missing all items on the MMB scale. After applying these criteria, 42 participants were excluded in total, resulting in a final sample of  $N = 154$  ( $n_{\text{threat}} = 79$ ,  $n_{\text{no-threat}} = 75$ ).

Participants ranged in age from 18 to 69 years ( $M = 30.6$ ,  $SD = 10.9$ ). The sample was predominantly highly educated: 80.5% held at least a general university entrance qualification (Abitur), and 60.4% had completed a university degree.

#### 3.2 Motivation For Masculine Behaviour Validation

##### Figure 2

*Polychoric Correlation Matrix for MMB Items*



*Note.* Items 1–5 = Pressured Motivation subscale; Items 6–9 = Autonomous Motivation subscale. Rectangles indicate the hypothesized factor structure. Darker shading indicates stronger positive correlations.

Prior to the moderation analysis, the validity and reliability of the MMB scale were assessed (see Figure 2 for correlations and Table B1 for item descriptives). The Kaiser-Meyer-Olkin criterion yielded an overall Measure of Sampling Adequacy (MSA) of .85, indicating the data were suitable for factor analysis. The hypothesized two-factor model by Stanaland & Gaither (2021) specified uncorrelated latent factors for Pressured Motivation (Items 1–5) and Autonomous Motivation (Items 6–9). A CFA was conducted to validate this structure; however, the fit indices did not satisfy the predefined criteria (CFI = .85, TLI = .8, RMSEA = .18, SRMR = .25).

To better understand the data structure, an EFA with oblimin rotation was conducted. A parallel analysis and scree plot supported a two-factor solution. The resulting factor loadings (Table 1) showed strong coherence for Items 1–5 on the first factor (Pressured Motivation) and

moderate coherence for Items 6–9 on the second factor (Autonomous Motivation). However, Item 9 (“It is important to me not to be feminine”) displayed a cross-loading on both factors, suggesting poor discriminant validity.

**Table 1**

*EFA Factor Loadings for 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	
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.* Factor loadings below .30 are suppressed for clarity. Extraction method: Maximum Likelihood with oblimin rotation.

Based on the EFA results, Item 9 was removed due to its cross-loading. Examination of modification indices further suggested a residual covariance between Items 7 and 8. This modification is conceptually supported, as both items express positive affect toward masculine identity (“enjoy,” “happy”). A final CFA on the 8-item scale yielded an almost acceptable fit (CFI = .97, TLI = .95, RMSEA = .1, SRMR = .07). The CFI and SRMR satisfied the target values of .95 and .08, respectively. While the TLI fell slightly below the .95 threshold and the RMSEA remained elevated, both showed substantial improvement. This pattern may be partially attributable to the limited sample size ( $N = 154$ ). No further modifications were made, as

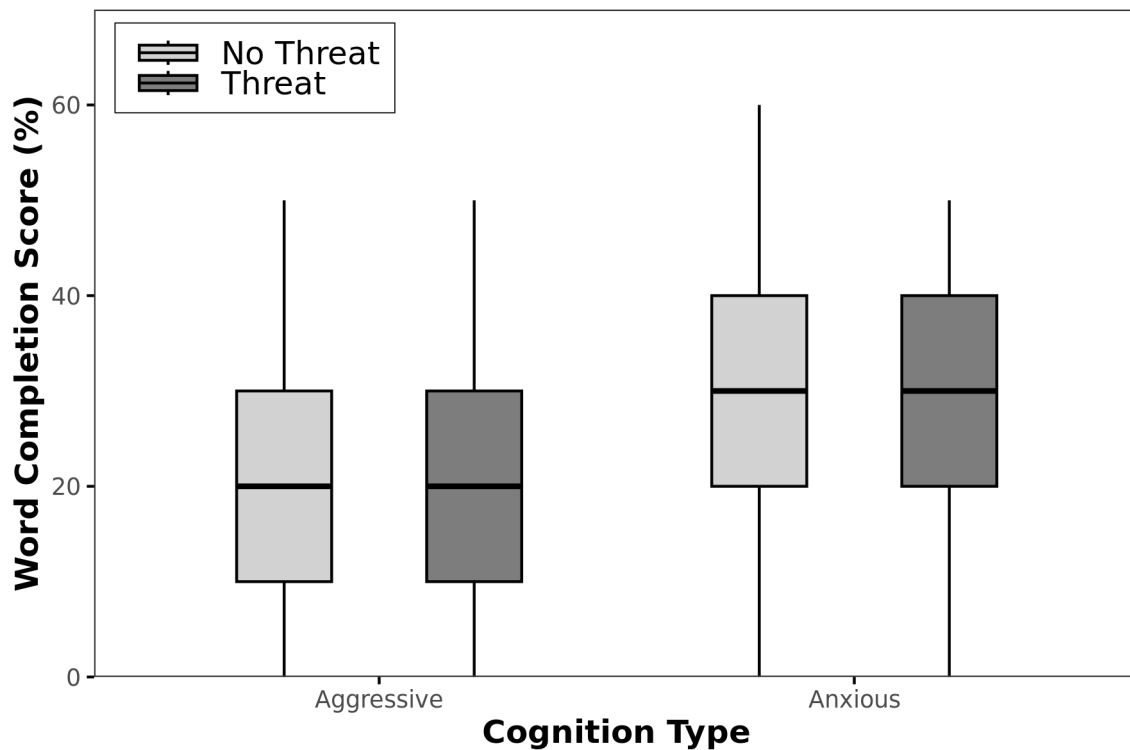
additional changes could not be theoretically justified and would risk overfitting the model to the data.

Finally, reliability was assessed using McDonald's omega ( $\omega$ ), which does not assume tau-equivalence. The Pressured Motivation subscale showed excellent reliability ( $\omega = .93$ ), while the Autonomous Motivation subscale showed acceptable reliability ( $\omega = .70$ ).

The two subscales showed a moderate positive correlation ( $r = .39$ ). For the subsequent moderation analyses, subscale scores were computed as row means, with Item 9 excluded from the Autonomous Motivation scale.

**Figure 3**

*Aggressive and Anxious Word Completion Scores by Threat Condition*



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

### 3.3 Masculinity Threat Effects

Figure 3 displays the distribution of aggressive and anxious word completion scores by threat condition.

#### 3.3.1 Masculinity Threat on Aggressive Cognition (Hypothesis 1a)

Hypothesis 1a predicted higher aggressive cognition in the threat condition compared to the no-threat condition. Contrary to this prediction, aggressive word completion scores were lower in the threat condition ( $M = 19.5$ ,  $SD = 12.3$ ) than in the no-threat condition ( $M = 22.8$ ,  $SD = 12.3$ ).

Assumption checks were satisfactory: Levene's test indicated homogeneous variances, and DHARMA residual diagnostics revealed no significant deviations from the expected distribution. A binomial GLM revealed no statistically significant difference between conditions ( $z = -1.59$ ,  $p = .111$ ). Given the directional hypothesis (predicted  $\beta > 0$ ) and the observed effect in the opposite direction, the one-tailed  $p$ -value was  $p = .944$ .

A sensitivity analysis was performed by identifying and excluding five influential observations. The GLM refitted on the reduced sample continued to show no significant effect ( $z = -1.48$ ,  $p = .139$  one-tailed  $p = .93$ ).

A sensitivity power analysis (see Figure B1) indicated that, given the sample size ( $n_{\text{threat}} = 79$ ,  $n_{\text{no-threat}} = 75$ ), the study had 80% power to detect effects of  $d \geq 0.40$  at  $\alpha = .05$ . The observed effect ( $d = -0.27$ ) was in the opposite direction and smaller in magnitude than this threshold, suggesting the absence of the hypothesized effect rather than insufficient statistical power.

#### 3.3.2 Masculinity Threat on Anxious Cognition (Hypothesis 1b)

Hypothesis 1b predicted higher anxious cognition in the threat condition compared to the no-threat condition. Anxious word completion scores were virtually identical across conditions: threat condition ( $M = 28$ ,  $SD = 12.8$ ) and no-threat condition ( $M = 27.6$ ,  $SD = 14.5$ ).

Assumption checks were satisfactory: Levene's test indicated homogeneous variances, and DHARMA residual diagnostics revealed no significant deviations from the expected distribution. A binomial GLM revealed no statistically significant difference between conditions ( $z = 0.18, p = .859$ ). Given the directional hypothesis (predicted  $\beta > 0$ ), the one-tailed  $p$ -value was  $p = .430$ .

A sensitivity analysis was performed by identifying and excluding five influential observations. The GLM refitted on the reduced sample continued to show no significant effect ( $z = 1.04, p = .3$  one-tailed  $p = .15$ ).

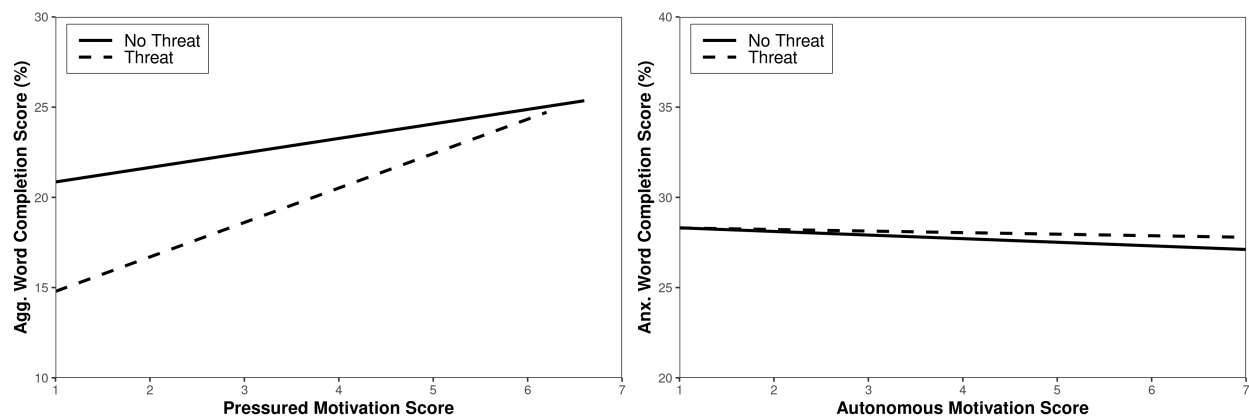
A sensitivity power analysis (see Figure B1) indicated that the study had 80% power to detect effects of  $d \geq 0.40$  at  $\alpha = .05$ . The observed effect ( $d = 0.03$ ) was substantially smaller than this threshold, suggesting the absence of a meaningful effect rather than insufficient statistical power.

### 3.4 Moderation of Masculinity Threat Effects

Having found no main effects of threat condition on either aggressive or anxious cognition, the moderation hypotheses were tested to examine whether individual differences in motivation for masculine behavior influenced threat responses.

**Figure 4**

*Moderation of Word Completion Scores by Motivation for Masculine Behavior*





*Note.* Left panel: Aggressive word completion scores (%) by threat condition and pressured motivation. Right panel: Anxious word completion scores (%) by threat condition and autonomous motivation.

### ***3.4.1 Pressured Motivation on Aggressive Cognition (Hypothesis 2a)***

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. As illustrated in Figure 4, descriptive patterns showed a steeper positive slope between pressured motivation scores and aggressive word completion scores in the threat condition compared to the no-threat condition, consistent with the predicted interaction.

Assumption checks were satisfactory: Levene's test indicated homogeneous variances, and DHARMA residual diagnostics revealed no significant deviations from the expected distribution. A binomial GLM with 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 ( $z = 0.92, p = .357$ ). Given the directional hypothesis (predicted  $\beta > 0$  for the interaction), the one-tailed  $p$ -value was  $p = .178$ , leading to a failure to reject the null hypothesis.

A sensitivity analysis was performed by identifying and excluding four influential observations. The GLM refitted on the reduced sample continued to show no significant interaction effect ( $z = 1.45, p = .146$  one-tailed  $p = .073$ ), though the effect approached marginal significance.

A sensitivity power analysis (see Figure B2) indicated that, with  $N = 154$  and 3 predictors (threat condition, pressured motivation, and their interaction), the study had 80%

power to detect interaction effects of  $f^2 \geq 0.052$  at  $\alpha = .05$ . This corresponds to a small-to-medium effect size according to Cohen (2013) conventions.

### **3.4.2 Autonomous Motivation on Anxious Cognition (Hypothesis 2b)**

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. As illustrated in Figure 4, descriptive patterns did not support this prediction: anxious word completion scores showed no meaningful change with increasing autonomous motivation scores in either condition.

Assumption checks were satisfactory: Levene's test indicated homogeneous variances, and DHARMA residual diagnostics revealed no significant deviations from the expected distribution. A binomial GLM with 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 ( $z = 0.07$ ,  $p = .946$ ). Given the directional hypothesis (predicted  $\beta > 0$  for the interaction), the one-tailed  $p$ -value was  $p = .473$ , leading to a failure to reject the null hypothesis.

A sensitivity analysis was performed by identifying and excluding influential observations. The GLM refitted on the reduced sample continued to show no significant interaction effect ( $z = 0.41$ ,  $p = .678$  one-tailed  $p = .339$ ).

A sensitivity power analysis (see Figure B2) indicated that the study had 80% power to detect interaction effects of  $f^2 \geq 0.052$  at  $\alpha = .05$ . This corresponds to a small-to-medium effect size according to Cohen (2013) conventions.

## **3.5 Experimental Manipulation Check**

Given the null findings across all four hypotheses, it is essential to verify that the experimental manipulation successfully induced a perceived threat to masculine identity. To assess whether the manipulation was effective, two measures were examined.

First, participants were asked at the end of the questionnaire to indicate which feedback they had received, on a scale from 1 (stereotypically feminine) to 10 (stereotypically masculine). Descriptive statistics confirmed a clear distinction between conditions: the threat condition ( $M = 3.1$ ,  $SD = 0.9$ ) reported substantially lower scores than the no-threat condition ( $M = 6.2$ ,  $SD = 1$ ).

Second, participants were asked to self-evaluate their gender knowledge on the same scale. Descriptive statistics showed lower self-evaluations in the threat condition ( $M = 5.7$ ,  $SD = 1.4$ ) than in the no-threat condition ( $M = 6.8$ ,  $SD = 1.4$ ). A  $t$ -test confirmed that participants in the threat condition rated their gender knowledge significantly lower than those in the no-threat condition ( $t(154) = 4.74$ ,  $p = < .001$ ).

## 4 Discussion

The current studies aim was to replicate masculinity threat effects and their motivational moderators in a German cultural sample. Based on the proportion of aggressively and anxiously completed items in the WFCT, the results of this study did not reveal the hypothesized increase of aggressively and anxiously coded items in the threat condition. Furthermore, this effect of the threat motivation was not moderated by the pressured and autonomous motivation to be masculine for aggressive and anxious trials respectively. Nonetheless, also the failure to replicate these effects highlights prolific insights into the assessment of masculinity threat effects and the motivation for gendered behaviour, that will be beneficial to future work.

### 4.1 Motivation for Masculine Behaviour in a German Sample

A key objective of the current study was to establish a translation of the MMB in a German cultural sample and test whether it reveals the same two-factorial structure distinguishing autonomous and pressured motivation to be masculine.

#### **4.1.1 Factorial Structure of the MMB Scale**

- The CFA showed a sufficiently good fit for a two-factor model on the scale with a few constraints:
- Firstly Stanaland & Gaither (2021) original model showed the two factors to be uncorrelated. The current study did not show this, as a modeling the factors as uncorrelated produced a bad model fit. From a theoretical perspective, it makes sense to consider these two factors as correlated, since: 1. The underlying constructs are actually related (which is plausible, considering that my own ideals and the expectations of others influence each other). 2. It may be hard for participants to distinguish between their internal and external motivations in these ‘simple’ one-line sentence. Additionally, the original studies by Good & Sanchez (2010) that investigated the distinction between external and internal pressured motivation to conform to a gender stereotype, also reported a moderate correlation between their two factors ( $r = \dots$ ).
- Secondly, one item of the autonomous motivation scale (“Es ist mir wichtig, mich nicht weiblich zu verhalten”, original: [It is important to me not to be feminine]) was excluded from the scale, because it showed equally high loadings (.39) on the pressured and autonomous factor. Interestingly, a very similar pattern emerged in the original validation of the scale (Study 2; Stanaland et al., 2023), where the item loaded only marginally higher on autonomous than pressured motivation (pressured: .36, autonomous: .43). The fact that this pattern seems to emerge consistently is interesting and may stem from an inherent property of the item: It may be, that the rejection of femininity may inherently be something that is associated with an external pressure,

rather than an intrinsic motivation. In the current study, we chose to exclude the item, because it showed a poor fit into the hypothesized structure and removing it significantly improved the model fit.

- Thirdly, a residual covariate was allowed for items seven and eight of the autonomous motivation factor. This was justified, because both items express positive affect toward masculine identity (“enjoy,” “happy”) and allowing for the correlation significantly improved the model fit.

In summary ...

#### ***4.1.2 Sample considerations/translation***

Although participants largely come from an academic background, this did not yield very low MMB agreements (previous consideration that a largely university sample may be biased).

A lot of people in the sample however reported suspicion and were able to identify the true purpose of the study (i.e. experimental manipulation of feedback). This provides a significant challenge when conducting this paradigm in an academic sample, since the participants are more sceptical/ more used to experimental manipulations.

#### ***4.1.3 Scale construction***

- Different methods in assessing the MMB scale: (Study 1; Stanaland et al., 2023) let participants describe their idea of an ideal man and then asked participants with regard to the ideal they described (e.g. In general, I present myself like the man I described because I want others’ acceptance and approval.). In contrast (Study 2; Stanaland et al., 2023) that was also the base for this study, omitted the description of an ideal man and asked more generally “In general, I’m masculine because I want others’ acceptance and

approval.”. Notably in their first study, Stanaland & Gaither (2021) observed a much smaller correlation between the two factors ( $r = .15$ ) than in the second study  $r = .5$ . As this study again produced a high correlation between the two factors, one may assume, that formulating of an ideal man helps participants distinguish better autonomous and pressured components of their masculinity.

Difficulties translating the scale (männlich sein, vs. männlich verhalten) discussed in the Appendix ...

-> current results show that the items were comprehensible and none of the items produces severe bottom or ceiling effects.

-> No factor emerged in the FA that would suggest a factor for phrasing acting vs. state

#### **4.1.4 Conclusion on the MMB in a German sample**

- Successful translation of the scale and showed a valid and reliable measure for pressured and autonomous motivation for masculine behaviour that replicated previous studies (Stanaland et al., 2023; Stanaland & Gaither, 2021).

Future studies that want to investigate pressured and autonomous motivation for masculine behaviour may regard the “state vs. agentiveness” of translating traits related to gender (being vs. doing gender), as well as consider the benefit of an introductory text.

## **4.2 Masculinity Threat Effects**

- The study failed to show masculinity threat effects: Both the hypothesized higher aggressive and anxious word completions (Hypothesis 1a & 1b), as well as the moderation of this effect via autonomous and pressured motivation (Hypothesis 2a & 2b) remained non-significant.
- Three possible reasons for the current sample not showing masculinity threat effects: 1. Failure in operationalization of the threat (i.e. participants did not perceive the stimuli

as a threat). 2. There is an issue with the WFCT as a dependent measure for aggressive and anxious cognition. 3. German participants truly do not react with aggressive or anxious responses to a threat to their masculinity (this would challenge the theoretical framework in a German context).

- The first reason for this could be that the study failed to successfully induce a masculinity threat to the participants.
- However, with regard to this, the study contained two checks of experimental manipulation:
  - Firstly, participants were asked to recall the feedback that they got from the GKT. Participants in the threat condition reported that their gender knowledge was rated much more feminine feedback than participants in the no threat condition. This shows that participants paid attention to the feedback, as well as the feedback was comprehensible.
  - Secondly, participants were asked to evaluate their own Gender Knowledge from typically feminine to typically masculine. Participants in the threat condition reported a significantly less masculine gender knowledge, than people in the no threat condition. This shows that the feedback the participants received was also credible.
- Summing up, these results indicate, that the operationalization of masculinity threat was successful.
- The second reason for failing to measure a masculinity threat response, could be an issue with the WFCT as a dependent measure for aggressive and anxious cognition.

- The current study used this measure, as previous studies were able to successfully use this measure in an English version
- In favor of this argument is a lack of construct validity, assessed via internal consistency of the aggressive and anxious word completion scales: Both scales showed practically no internal consistency of the items, meaning the items of the scales covaried randomly.
- The last possibility, could be that the study did not show a masculinity threat effects, could be that the participants really did not experience aggressive or anxious emotion after a threatening feedback to their masculinity.
- This is a serious point to discuss, as it would have a serious implication on the theoretical background (precarious manhood, fragile masculinity).
- Although a masculinity threat effect was found with the lower self evaluation of Gender Knowledge, this must not be linked to an aggressive or anxious response.
- A limitation that has to be regarded that the studied population was largely academic.

In favor of the failure of dependent measures is the study by (TODO), that tried to establish a German WFCT with validated measures, but failed to do so.

#### **4.3 Measuring (Masculinity) Threat Response (or something like that)**

- As previously highlighted, the lack of validity of the WFCT is likely to have caused the non-observation of masculinity threat effects.

Discuss the benefits, as well as disadvantages of dependent measures for aggressive and anxious cognition

##### **1. WFCT**



- Advantages: Implicit, which is particularly relevant with a dependent measure that is subject to social desirability (anxious response is incongruent to a masculine stereotype and thus may be underreported). Also participants may not be self-aware to their emotions and thus not be able to report them.
- Disadvantages: Very high measurement error -> there are so many reasons why a subject may complete a word fragment in one way or another -> it would need a lot of items and participants to counter that. An unpublished study by (TODO) tried to cross-validate a word fragment completion task with established measures of aggressive responses, but failed to do so.

## 2. Self-reporting:

- Advantage: High validity
- Disadvantage: Subject to social desirability & lack of self-awareness

## 3. More implicit measures

- e.g. assessing aggressive cognition via agreement to discriminatory statements (e.g. TODO).
- Advantage: Good validity, has been successfully used in many studies (TODO).
- Disadvantage: More difficult in studentic/academic samples (as the agreement to discriminatory statements may be generally very low (TODO)).

## 4. Behavioral measures

- e.g. aggressive behaviour (TODO), stress levels (TODO) ...

Advantages: High validity, less subject to social desirability or self-awareness problems.

Disadvantages: High cost for setup.

Summary: Future studies for masculinity threat effects

Should:

- Implement measures to verify operationalization of masculinity threats (e.g. like in the current study)
- Include suspicion checks (like open text responses)
- Use validated measures for measuring threat response
- pay attention to social desirability and self-awareness problems
- pay attention to the interference between their dependent measures and population being studies (e.g. agreement to statements)

Could:

- Use distractor items, to hide the purpose of the study
- Use open text fields to ask about the purpose of the study
- use implicit measures, like the agreement to discriminatory statements

#### 4.4 Theory Discussion

Discuss different theoretical approaches to why masculinity threat responses:

##### 1. The EDT Model

- the current study did not provide evidence in favor of the EDT Model, since it did not show a moderation of masculinity threat effects by aggressive and anxious MMB.
- yet this is likely to a failed measurement of aggressive and anxious behaviour, as the study failed to reproduce simple masculinity threat responses.
- Conclusion: It remains challenging to empirically test the proposition of the EDT Model.

First research is pointing into that direction (cite studies, stanaland etc.) but more research is needed. Especially in the domain of autonomous motivation leading to increased anxious responses, no evidence was yet found for that. Thus Stanaland & Gaither (2021) even hypothesized that “Therefore, manhood may not be precarious in it’s entirety but only fragile”, and only an externalized pathway for the moderation of masculinity threat effects exists. Yet Vandello et al (2004), Study 4, (and others ??? look

again into literature) found an increase of anxious cognition due to masculinity threat and further research is needed to explain the underlying mechanisms for this response.

Another account for the affective response after a masculinity threat could be the control beliefs. This account has not been discussed in the literature so far, but it could provide important insights.

- Briefly explain control beliefs and why they may have an impact on the response to a masculinity threat.

... future studies should take this into account and may switch to a motivational x control belief model of masculinity threats.

Outside of the context of identifying and studying mechanisms that underly masculinity threat responses, recent studies highlight a challenge to the idea of precarious manhood altogether: By studying female participants responses to a subjective femini threat (negative evaluation of physical femininity), Wittlin et al. (2024) female participants showed an increase in anxiety and reduced self-esteem. This challenges the original precarious manhood thesis, since it explained that manhood was something that must be earned, while womanhood is viewed as something inherently given. This naturalistic (is this the right word ???) view of manhood and womanhood has received criticism ever since (Addis & Schwab, 2013; Chrisler, 2013). It may be beneficial, not just methodologically, but for egalitarianism in scientific research to decouple our traditional understandings in the binarity of gender and perceive masculinity and femininity on a more broad view (... something like that)

## 4.5 Conclusion

...write a sentence to each of the arguments made in the discussion.

## References

- Addis, M. E., & Schwab, J. R. (2013). Theory and research on gender are always precarious. *Psychology of Men & Masculinity*, 14(2), 114.
- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16, 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, 411–419.
- 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.
- 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.
- 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.
- 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.
- Chrisler, J. C. (2013). *Womanhood is not as easy as it seems: Femininity requires both achievement and restraint.*

- 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.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge.
- 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://go.gale.com/>
- Dahl, J., Vescio, T., & Weaver, K. (2015). How threats to masculinity sequentially cause public discomfort, anger, and ideological dominance over women. *Social Psychology*.
- 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.
- 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.
- 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.
- Gilmore, D. D. (1990). *Manhood in the making: Cultural concepts of masculinity*. Yale University Press.
- 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.
- Harkness, J. A., Villar, A., & Edwards, B. (2010). Translation, adaptation, and design. *Survey Methods in Multinational, Multiregional, And Multicultural Contexts*, 115–140.
- Hartig, F. (2024). *DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. <https://doi.org/10.32614/CRAN.package.DHARMA>

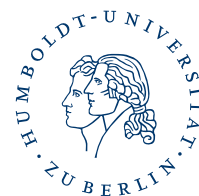
- Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological Review*, 94(3), 319.
- Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. *The American Statistician*, 55(1), 19–24.
- Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, 26(3), 329–367.
- 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.
- 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.
- 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.
- 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*.
- 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.
- 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, 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.
- 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.
- 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.
- 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.
- Stanaland, A. W. (2022). *Fragile Masculinity: Operationalizing and Testing a Novel Model of Identity Fragility* [Doctoral dissertation].
- 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.
- 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.
- 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), e13544.

- 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.
- 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.
- 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.
- 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.
- Weaver, K. S., & Vescio, T. K. (2015). The justification of social inequality in response to masculinity threats. *Sex Roles*, 72, 521–535.
- Winer, B. J., Brown, D. R., Michels, K. M., & others. (1971). *Statistical principles in experimental design* (Vol. 2). McGraw-hill New York.
- 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.
- 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.

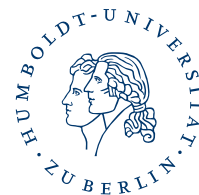
Zusätzlich versichere ich, dass ich beim Einsatz von IT-/KI-gestützten Werkzeugen diese Werkzeuge in der unten genannten „Übersicht verwendeter Hilfsmittel“, mit ihrem Produktnamen und der Versionsnummer, meiner Bezugsquelle (z.B. URL) und Angaben zur Nutzung vollständig aufgeführt sowie die Checkliste 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, 03.01.2026

---

Ort, Datum, Unterschrift



### Übersicht verwendeter Hilfsmittel

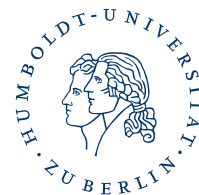
**Ausfüllhinweis:** Bitte geben Sie an, welche der folgenden Programme Sie im Rahmen Ihrer Arbeit verwendet haben. Bitte tragen Sie dazu bei jedem Programm ein, wie Sie dieses genutzt haben. Tragen Sie dazu die jeweilige Ziffer in die Spalte „Nutzung“ ein. Es können mehrere Ziffern je Programm eingetragen 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:**

<b>Programm</b> (Bitte geben Sie auch weitere Programme an, die Sie verwendet haben, um Ihre Arbeit zu schreiben.)	<b>falls relevant: betroffene Abschnitte</b> (Bitte geben Sie Seitenzahlen an, sofern es nicht auf das gesamte Dokument zutrifft.)	<b>Nutzung</b> (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	

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 Supplement

#### 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 SurveySwap oder 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.
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 et al. (2023) Study 2.

### ***Documentation of the Translation of the Motivation for Masculine Behaviour Scale***

The translation of the MMB scale was conducted following the TRAPD protocol (Harkness et al., 2010), using two independent native German, fluent English translators. In the first step, translators were given the original scale items and were instructed to find the most appropriate translation focusing on natural and semantically correct phrasing. Table A2 and Table A3 present the independent translations and translator comments from this initial translation phase.

**Table A2***MMB 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.
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	„Ich glaube“ wurde eingefügt, um die subjektive Wahrnehmung

Item	English (Original)	German Translation	Comments
		Leute dann nicht mögen würden.	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.
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 A3**

*MMB 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 denken, dass ich so sein sollte.	'Around other people' wird im Deutschen als 'in Gegenwart anderer' übersetzt.
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.

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 stative and agentic phrasings was adopted in the final version. The purely stative 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.

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.

### **Gender Knowledge Test**

Table A4 presents the 30 items of the Gender Knowledge Test with their original English wording and German translations.

**Table A4**

*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)

Item	Type	English (Original)	German (Translation)
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)
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)

Item	Type	English (Original)	German (Translation)
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)
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)

Item	Type	English (Original)	German (Translation)
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.

### ***Documentation of the Translation of the Gender Knowledge Test***

The translation of the Gender Knowledge Test followed the same TRAPD protocol as the MMB scale, using two independent native German, fluent English translators. Table A5 and Table A6 present the independent translations and translator comments.

**Table A5**

#### *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

Item	Type	English (Original)	German Translation	Comments
				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. Thompson-Maschinenpistole)	‘Tommy Gun’ als ‘Thompson’ bekannt; ‘Gatling’ weniger gebräuchlich.
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.

Item	Type	English (Original)	German Translation	Comments
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.
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.



Item	Type	English (Original)	German Translation	Comments
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!
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	Tassen als Maßeinheit in Rezepten in DE bekannt.

Item	Type	English (Original)	German Translation	Comments
			Tasse Reis zu kochen? (2 vs. 3 Tassen)	
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.
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' in DE bekannt; Kontext bleibt erhalten.

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

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

**Table A6**

*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.
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.

Item	Type	English (Original)	German Translation	Comments
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 which film? (True Lies vs. Total Recall)	In welchem Film hat Arnold Schwarzenegger mehr Menschen getötet? (True Lies vs. Total Recall)	Keine besonderen Anmerkungen.
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	Beim Schlagen sollte man seine Faust (etwa einen Fuß über das optimale Ziel hinaus vs.	‘A foot beyond’ wird im Deutschen als ‘etwa einen Fuß über ... hinaus’ übersetzt.

Item	Type	English (Original)	German Translation	Comments
13	M	beyond target vs. directly at target)	direkt auf das Ziel) richten	Keine besonderen Anmerkungen.
		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)	
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. vorderen Fahrerreifen) zielen	Keine besonderen Anmerkungen.
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	Botox beseitigt Falten vorübergehend durch	'Muscle paralysis' wird im Deutschen als

Item	Type	English (Original)	German Translation	Comments
18	F	hydration vs. muscle paralysis) The first company to develop hair coloring was (Clairol vs. L'Oreal)	(Hautbefeuchtung vs. Muskelentspannung) Das erste Unternehmen, das Haarfärbemittel entwickelte, war (Clairol vs. L'Oréal)	'Muskelentspannung' bezeichnet. 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.
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	Keine besonderen Anmerkungen.

Item	Type	English (Original)	German Translation	Comments
			Stunden) aufbewahrt werden	
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-Übungen vs. Pilates)	'Kegel's' wird im Deutschen als 'Kegel-Übungen' bezeichnet.

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

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 GKT 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.

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 GKT 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 no-threat 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

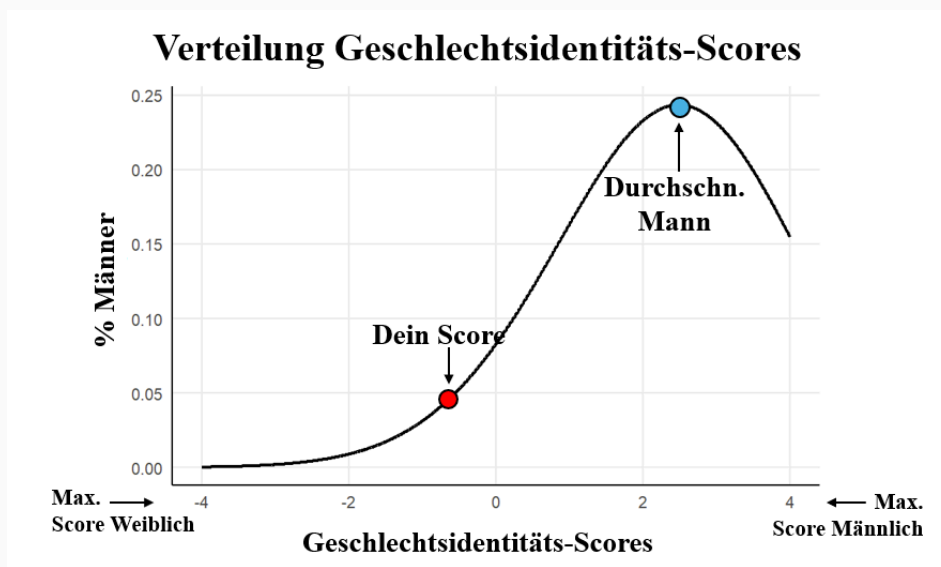


traditionell männlich oder weiblich geprägten Rollen und Selbstbildern verbunden sind. Die Auswertung basiert auf einer Skala von  $-4$  (ausgeprägt weiblich) bis  $+4$  (ausgeprägt männlich).

**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.

### ***No-Threat Condition***

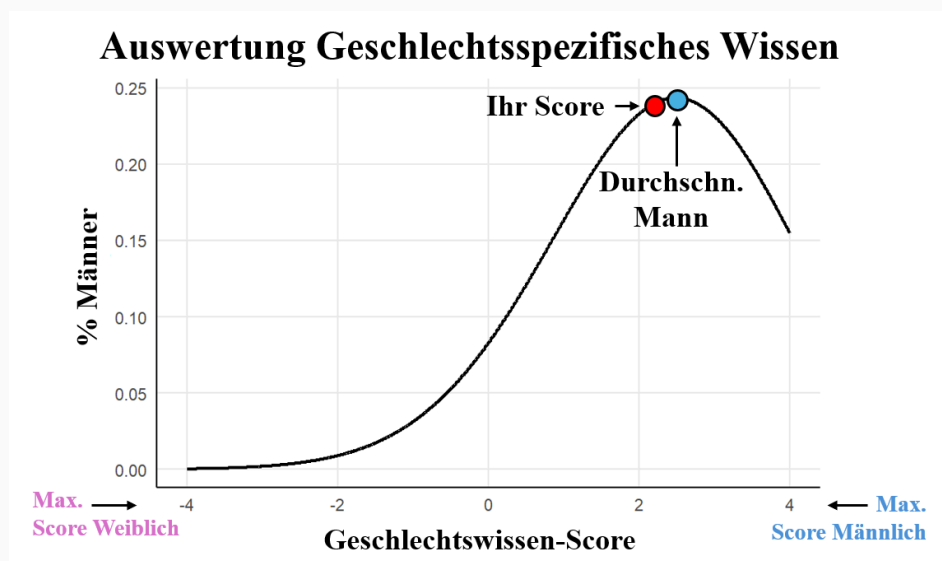
#### **Glückwunsch – du hast den Geschlechtsspezifischen Wissenstest abgeschlossen!**

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: +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

Participants completed word fragments by typing the missing first letter. Each fragment could be completed as either a target word (aggressive or anxious) or a neutral word. A seven-second time limit was imposed per item to encourage spontaneous responses. Table A7 presents the aggressive cognition items, and Table A8 presents the anxious cognition items.

**Table A7**

*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

Fragment	Aggressive	Neutral
_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 A8**

*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.

### **Manipulation Check and Suspicion Probe**

Table A9 presents the items used to assess participants' recall of the experimental feedback and their suspicions about the study's purpose.

**Table A9**

*Manipulation Check and Suspicion Probe Items*

Item	German	English
1	Welche Rückmeldung haben Sie in dieser Studie zu Ihrem Wissen in geschlechtsspezifischen Themen erhalten? <i>Skala: 1 (typisch weiblich) – 10 (typisch männlich)</i>	What feedback did you receive in this study regarding your knowledge of gender-specific topics? <i>Scale: 1 (typically feminine) – 10 (typically masculine)</i>
2	Wie würden Sie Ihr Wissen in geschlechtsspezifischen Themen einschätzen? <i>Skala: 1 (typisch weiblich) – 10 (typisch männlich)</i>	How would you rate your knowledge of gender-specific topics? <i>Scale: 1 (typically feminine) – 10 (typically masculine)</i>
3	Hatten Sie während der Teilnahme den Eindruck, dass es in der Studie um etwas anderes ging als angegeben? <i>Ja / Nein</i>	During participation, did you have the impression that the study was about something other than what was stated? <i>Yes / No</i>
4	Was denken Sie, worum es in dieser Studie wirklich ging? <i>Offene Antwort</i>	What do you think this study was really about? <i>Open response</i>
5	Haben Sie Gedanken zur Studie? <i>Offene Antwort</i>	Do you have any thoughts about the study? <i>Open response</i>

*Note.* Items 1 and 2 served as manipulation checks. Items 3–5 assessed participant suspicion.

Item 4 was only shown when participants responded ‘yes’ to item 3.

### Demographic Questionnaire

Table A10 presents the demographic items administered at the end of the study.

**Table A10**

#### *Demographic Questionnaire Items*

Item	German	English
1	Wie alt sind Sie?	How old are you?

Item	German	English
2	Welches Geschlecht haben Sie? <i>Männlich, Weiblich, Divers, Keine Angabe</i>	What is your gender? <i>Male, Female, Diverse, Prefer not to say</i>
3	Wie gut sind Ihre Deutschkenntnisse? <i>Keine Deutschkenntnisse, Grundkenntnisse, Gute Kenntnisse, Sehr gute Kenntnisse, Deutsch ist meine Muttersprache</i>	How proficient is your German? <i>No German, Basic, Good, Very good, German is my native language</i>
4	Was ist Ihr höchster Schulabschluss? <i>Noch kein Abschluss, Schule beendet ohne Abschluss, Haupt-/Volksschule, Mittlerer Schulabschluss, Fachhochschulreife, Abitur</i>	What is your highest educational degree? <i>No degree yet, Left school without degree, Secondary school, Middle school, Technical college entrance, University entrance (Abitur)</i>
5	Welchen höchsten beruflichen Abschluss haben Sie? <i>Noch keinen, Lehre, Berufsfachschule, Fachschule, Fachhochschule, Universität, Anderer</i>	What is your highest occupational degree? <i>None yet, Apprenticeship, Vocational school, Technical school, University of applied sciences, University, Other</i>

*Note.* Response options are abbreviated. Full German response options were presented to participants.

### 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 diene 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.

### **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 was explored. Table A11 summarizes effect sizes from prior studies on masculinity threat.

### ***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 WFCT. Based on these estimates, a priori power analyses were conducted using one-tailed independent samples t-tests ( $\alpha = .05$ , 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 A1).

### ***Moderation by Motivation***

For the moderation of pressured motivation on aggressive cognition, a small effect size was assumed ( $f = 0.17$ ), consistent with prior findings on motivational moderators. 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 ( $\alpha = .05$ , power = .80; see Figure A2).

### ***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 (CFA), 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***

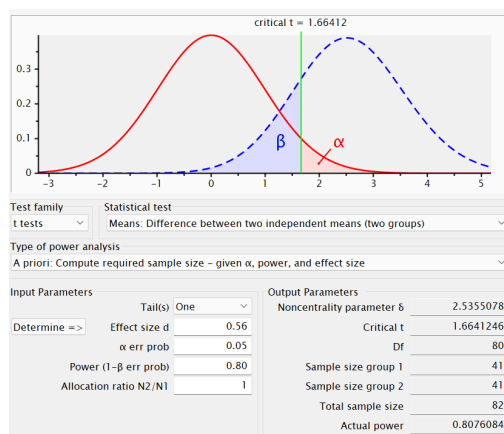
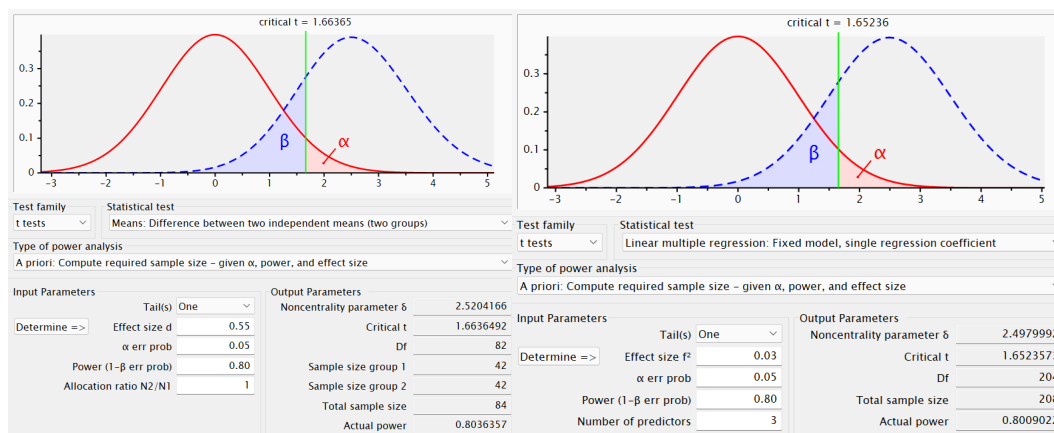
Based on the largest required sample size (moderation analysis; 208 participants) and accounting for an estimated 10% exclusion rate, the final target sample size was set at 229 participants.

**Table A11**

*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 et al. (2024)	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 WFCT and operationalized masculinity threat via negative feedback on the Gender Knowledge Test.

**Figure A1***Power Analysis for Masculinity Threat on Aggressive Cognition***Figure A2***Power Analysis for Masculinity Threat on Aggressive & Anxious Cognition*

## Appendix B

### B Results Supplement

The analysis code can be viewed publicly at

[add link](#)

. 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.

Table B1 presents descriptive statistics for the Motivation for Masculine Behavior items.

#### Motivation for Masculine Behaviour

**Table B1**

*Descriptive Statistics for Motivation for Masculine Behavior Items*

Item	German Item	<i>M</i>	<i>SD</i>
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
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).

#### Word Fragment Completion Task: Detailed Psychometric Properties

Table B2 and Table B3 present response frequencies for the aggressive and anxious word fragment completion items, respectively.

### ***Aggressive Word Fragment Items***

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 matrix 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 Fragment Items***

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 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.

## **Table B2**

### ***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.

**Table B3**

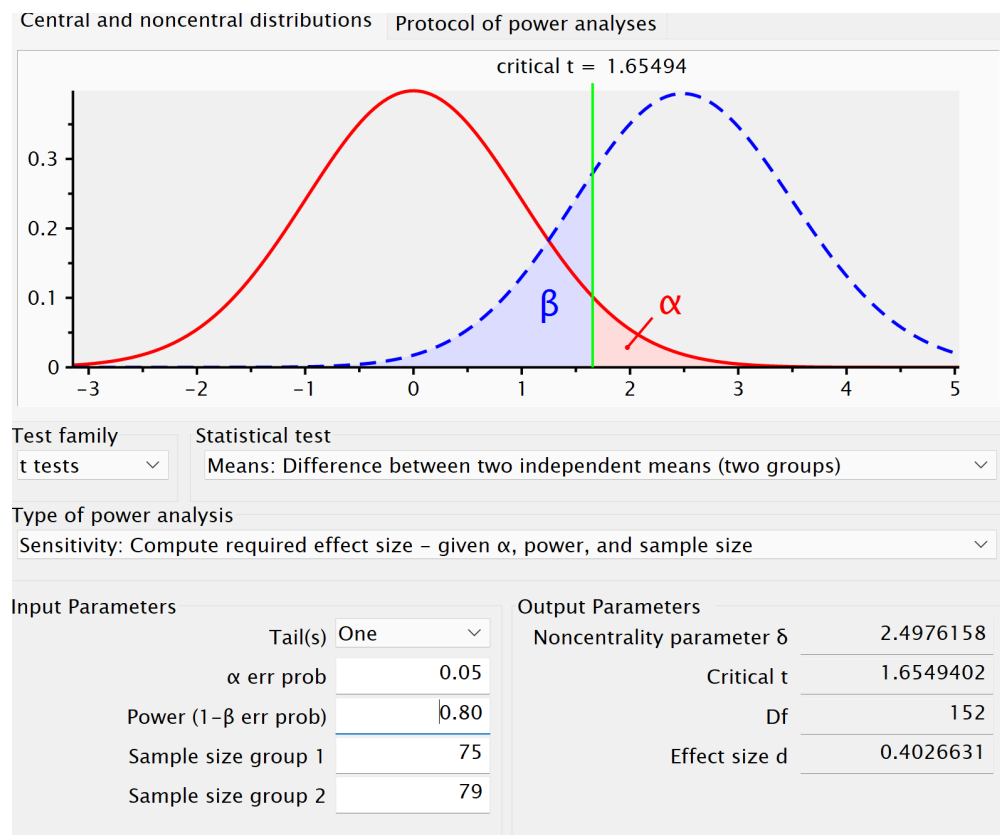
*Response Frequencies for Anxious Word Fragment Completion Items*

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.

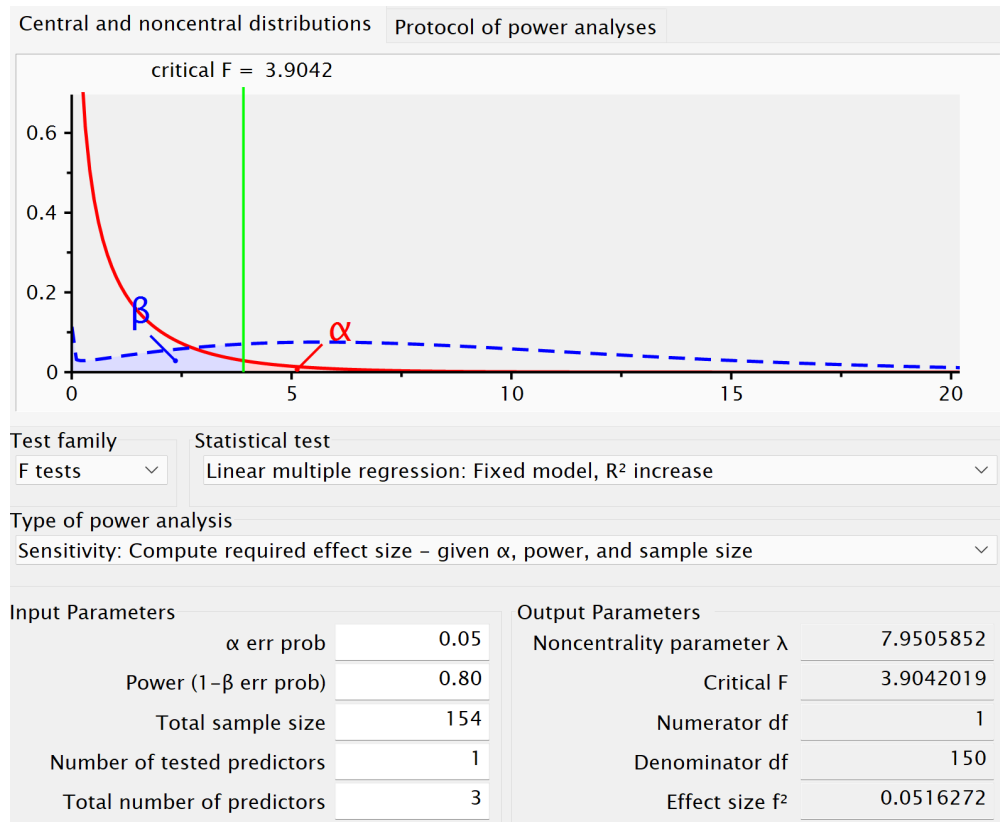
### **Sensitivity Power Analysis**

Sensitivity analyses were conducted using G\*Power 3.1 to determine the minimum effect sizes detectable with 80% power at  $\alpha = .05$  given the achieved sample size.

**Figure B1***Sensitivity Analysis for Hypotheses 1a and 1b*

*Note.* G\*Power output for sensitivity analysis. Test family: t tests; Statistical test: Means: Difference between two independent means (two groups). With  $n_1 = 79$ ,  $n_2 = 75$ ,  $\alpha = .05$ , and power = .80, the minimum detectable effect size is  $d = 0.40$ .

**Figure B2***Sensitivity Analysis for Hypotheses 2a and 2b*



*Note.* G\*Power output for sensitivity analysis. Test family: F tests; Statistical test: Linear multiple regression: Fixed model,  $R^2$  increase. With  $N = 154$ , 3 predictors, 1 tested predictor (interaction),  $\alpha = .05$ , and power = .80, the minimum detectable effect size is  $f^2 = 0.052$ .



## Appendix C

### C List of Appendices

A Materials Supplement .....	44
Study Introduction .....	44
Motivation for Masculine Behavior .....	45
Documentation of the Translation of the Motivation for Masculine Behaviour Scale . . .	46
Gender Knowledge Test .....	50
Documentation of the Translation of the Gender Knowledge Test .....	54
Masculinity Threatening and Non-Threatening Feedback .....	64
Threat Condition .....	64
No-Threat Condition .....	66
Word Fragment Completion Task .....	67
Manipulation Check and Suspicion Probe .....	68
Demographic Questionnaire .....	69
Debriefing .....	70
Consent After Debriefing .....	72
Sample Size Planning .....	73
Masculinity Threat Effects .....	73
Moderation by Motivation .....	73
Factorial Structure of Motivation for Masculine Behavior .....	74
Final Target Sample Size .....	74
B Results Supplement .....	76
Motivation for Masculine Behaviour .....	76
Word Fragment Completion Task: Detailed Psychometric Properties .....	76
Aggressive Word Fragment Items .....	77

Anxious Word Fragment Items .....	77
Sensitivity Power Analysis .....	78
C List of Appendices .....	81