



You Don't Bend It Like Beckham if You're Female and Reminded of It: Stereotype Threat Among Female Football Players

Hilmar Grabow* and Melanie Kühl

Social and Political Psychology, Institute of Psychology, Kiel University, Kiel, Germany

Originally, the stereotype threat effect – poorer performance due to a fear of fulfilling a negative stereotype about one's group – was demonstrated for cognitive tasks (e.g. Steele and Aronson, 1995, or Steele, 1997). Drawing on the widespread stereotype of women being unable to play football we experimentally tested ($N = 80$) whether a respective threat affected female football players' goal scoring precision, i.e. a complex and demanding motor task. Those participants who were reminded of the stereotype scored significantly less hits than those not reminded. Additionally, deviations from the instruction during task execution (e.g. shooting from another distance than demanded or using the wrong foot) were recorded. Stereotype threat did not affect this comparatively more cognitive task of following instructions correctly. In order to explore underlying mechanisms of the observed stereotype effect, several potential mediators, e.g. measures of cognitive interference, or collective identification, were tested. None emerged as an unquestionable link between threat and motor performance. We discuss, however, why collective identification – in comparison to cognitive demand – appears to be the more promising explanatory concept.

Keywords: stereotype threat, football, soccer, motor task, precision, sport

OPEN ACCESS

Edited by:

Nathan Andrew Heflick,
University of Lincoln,
United Kingdom

Reviewed by:

Michael Häfner,
Berlin University of the Arts, Germany
Ana-Maria Bliuc,
Western Sydney University, Australia

*Correspondence:

Hilmar Grabow
h.grabow@psychologie.uni-kiel.de

Specialty section:

This article was submitted to
Personality and Social Psychology,
a section of the journal
Frontiers in Psychology

Received: 22 July 2018

Accepted: 09 August 2019

Published: 28 August 2019

Citation:

Grabow H and Kühl M (2019) You
Don't Bend It Like Beckham if You're
Female and Reminded of It:
Stereotype Threat Among Female
Football Players.
Front. Psychol. 10:1963.
doi: 10.3389/fpsyg.2019.01963

INTRODUCTION

Stereotypically, football¹, the most popular sport worldwide (e.g. Brüggemeier, 2006, or Kunz, 2007), is perceived as male. According to Csizma et al. (1988), it ranks amongst the masculine – as opposed to neutral (e.g. tennis) or feminine (e.g. figure skating) – sports. The comment of Pfister (2015, p. 565) that “Real” football is still men's football, concisely describes the status of women's football in the media and public. Even at highest administrative level, women's football is not taken seriously, as Sepp Blatter's suggestions concerning an increase of its popularity illustrate. During his reign as FIFA president, he proposed – as reported by Christenson and Kelso (2004): ‘Let the women play in more feminine clothes like they do in volleyball’, explicating ‘They could, for example, have tighter shorts’. Historically, women were kept from playing football over considerable periods of time. In Germany, for example, the football association Deutscher Fußball-Bund (DFB) even banned it between 1955 (Hoffmann and Nendza, 2007b) and 1970 (Hoffmann and Nendza, 2007a). Women were judged unfit for this sport. This categorical verdict may have changed over time. In 2017, 5,819 women's and 5,875 girls' teams were active under DFB auspices, with over

¹Following John Cleese's convincing argumentation (Vaske, 2006), we use the English term football for the sport that is in some areas of the world also known as associate football, or soccer.

one million female DFB members (Deutscher Fußball-Bund, n. d.a). Women's football, unfortunately, is still not taken seriously in relation to the men's game, with it on many occasions being impeded, depreciated, and ridiculed (e.g. Keller, 2018, or Schwaiger, 2011), and viewed as 'inappropriate for girls' (Scruton et al., 1999). Football pundit Max Merkel did not, in all likelihood, utter an exclusive opinion but echo the mood of at least a considerable part of the football-oriented public, when he bad-mouthed women's football as a 'distress for the eyes' (Merkel, 2003). Summing up, in Germany (and other countries), stereotypical ascriptions have developed over time. In the past, it was an established idea that women must not play football; nowadays, a widely shared assessment is that they are not (as) good (as men) at it.

Considering stereotypical characteristics of men and women, this predominant perception of football as male can hardly surprise: across many cultures, typically female characteristics comprise 'timid', 'mild', 'soft-hearted', 'submissive', or 'fearful' (Williams and Best, 1990). These appear to be incompatible with a physical competition like football. Typical male characteristics, however, are perfectly in line with it: 'active', 'ambitious', 'strong', 'robust', 'quick', or 'tough' (Williams and Best, 1990). Diekmann and Eagly (2000) argue that gender-related stereotypes change over time, following perceived changes in social roles. For the present, however, they report that males are ascribed masculine personality (e.g. competitive, daring, aggressive, or courageous) and physical attributes (e.g. rugged, muscular, or physically strong) to a higher degree than females. Again, these 'male' attributes are perfectly in line with the requirements of football. Lueptow et al. (1995) report similar results: their participants on average rated being athletic as well as competitive clearly more typical male than typical female. In contrast to Diekmann and Eagly (2000), Lueptow et al. (1995) consider gender stereotypes rather stable over time. According to Fiske et al. (2002), women are perceived to be warmer/more communal but less competent/agentive than men, i.e. viewed in the light of a paternalistic stereotype. Women are likely to suffer from a backlash effect if they do not fit into this stereotype (e.g. Rudman, 1998, or Rudman and Glick, 2001), which is highly likely for female football players who strive to excel in a typical male domain.

Football, indeed, requires athletic skills from the players. Therefore, the stereotype (the widely shared belief of women being bad at football) could be based on a better athletic performance of males in comparison to females (Knisel et al., 2009), mirroring average morphological and physiological differences between genders (Hottenrott and Neumann, 2012) such as muscle size (Högler et al., 2008). It is not reasonable to assume, however, that these differences directly impact other football-related competences like technical or tactical skills.

For the purposes of our study, the genesis of this specific negative stereotype is far less important than its potential effect on women's performance. Generally speaking, 'Stereotype threat is being at risk of confirming, as self-characteristic, a negative stereotype about one's group' (Steele and Aronson, 1995, p. 797). Traditionally, the stereotype threat literature has dealt with cognitive tasks (e.g. Steele, 1997, Aronson et al., 1999, Spencer et al., 1999, Cadinu et al., 2003). For a start, Steele and Aronson (1995)

reported worse results in a verbal test for African Americans compared to Whites if the racial stereotype about intellectual ability was salient. Since then, the stereotype threat effect was demonstrated across many samples comprising members of diverse (social) groups, e.g. women compared to men (Spencer et al., 1999), African vs. White Americans (Steele and Aronson, 1995), or White vs. Asian men (Aronson et al., 1999). Tempel and Neumann (2016) provide evidence for a moderation of stereotype activation by gender role orientation: the more feminine their participants rated themselves the less well they performed on mental rotation or mathematical tasks if gender stereotypes were activated. Thus, stereotype threat effects appear to be more severe the more relevant (in terms of self-concept) the respective identity is. Furthermore, the effect is not restricted to purely intellectual tasks; Stone et al. found it for motor tasks in golf amongst Black and White participants and stated that 'Blacks suffer when the stereotype concerning their supposed poor sports intelligence is made salient, whereas Whites suffer when the stereotype concerning their supposed natural athletic ability is made salient' (Stone et al., 1999, pp. 1225–1226). Based on the premise that 'Sport and athletics may represent one of the few domains in which Whites are stereotyped negatively and suffer psychologically as a result' (Stone et al., 1999, p. 1213), they chose a miniature golf task in order to test whether different stereotypes would deteriorate the performance of both members of minorities as well as majorities. Indeed, the participants' performance, measured by the number of strokes conducted in order to hit balls in a series of holes, suffered when participants were threatened by a fitting stereotype about their group. Not least because the participants were relative golf novices, the motor task in all likelihood was executed under articulate conscious control. In comparison to well-learned, automated movements, such explicit cognitive effort could be quite prone to cognitive interference – much like the cognitive tasks examined earlier in the stereotype threat literature. Chalabaev et al. (2008) avoided this potential problem by examining a dribbling task amongst experienced female football players, hence a well-learned routinely executed set of movements. The time for completion of a slalom course served as the dependent variable. Thus, results were not solely attributable to technical but also to athletic skills, hence, a dimension reflecting actual gender differences.

Summarising research findings, Keller (2008) extracted four preconditions whose presence increase the likelihood of a manifestation of the stereotype threat effect. First, a relevant stereotype has to be salient and directly applicable to the situation. Second, the respective person needs to highly identify with the domain (here: football). Third, the task has to be demanding, touching the person's performance limit. Fourth, the person needs to have the impression their performance is under scrutiny, i.e. the test has to be diagnostic.

In line with this research, we predicted and found that, compared to an unthreatened control group, women kick the ball less precisely if reminded of their gender. We chose this specific dependent variable because it is not plausible to expect an actual gender specific difference in shooting precision (or technical skills in general). Thus, the threat could not plausibly initiate remembrance of factual knowledge

about women's power and strength – unlike, e.g. in the study by Chalabaev et al. (2008).

In addition to a conceptual replication of the already well-established stereotype threat effect (employing a new dependent variable), we tested potential explanations for its occurrence. Apparently, there has not yet emerged an undisputed model of the mechanisms behind it. Several concepts/variables have been tested as mediators between stereotype threat and decreased performance, e.g. anxiety, self-efficacy, or evaluation apprehension (Spencer et al., 1999), or, for sensorimotor tasks, monitoring processes (Schmader et al., 2008). The last-mentioned authors argue that a deterioration of automated sensorimotor tasks can hardly be accounted for by working memory impairments – according to their integrated process model an important link between stereotype threat and worsening performance on cognitive tasks under controlled processing. Therefore, they propose that under threat, monitoring and controlling of a person's behaviour consume resources and impede performance (pp. 337–338).

In order to explain the expected effect of stereotype threat we draw on the social identity tradition, i.e. social identity theory (SIT, Tajfel and Turner, 1986), and self-categorisation theory (SCT, Turner et al., 1987). Tajfel and Turner present the concept of social categorisations – ‘cognitive tools that segment, classify, and order the social environment’ (Tajfel and Turner, 1986, p. 15), whereby individuals distinguish themselves (and others within the same category) from members of outgroups. Groups ensuing from these categorisations ‘provide their members with an identification of themselves in social terms’ (Tajfel and Turner, 1986, p. 16). Tajfel and Turner further assume that ‘Social groups or categories and the membership of them are associated with positive or negative value connotations. Hence, social identity may be positive or negative according to the evaluations’ (Tajfel and Turner, 1986, p. 16). Given that ‘Individuals strive to maintain or enhance their self-esteem: they strive for a positive self-concept’ (Tajfel and Turner, 1986, p. 16), being reminded of a decidedly negative evaluation of one's group needs to be considered a setback for this striving. Potential ways out of this aversive state of a threatened social identity are individual mobility (leaving or dissociating oneself from the respective group), social creativity (e.g. referring to a different dimension when comparing in- and outgroup), or social competition, i.e. direct competition with the outgroup (Tajfel and Turner, 1986). According to Turner et al. (1987), individuals feature a bundle of cognitive representations of the self which ‘take the form, amongst others, of self-categorisations’ (Turner et al., 1987, p. 44); the salience of a specific self-concept is dependent on ‘an interaction between the “relative accessibility” of that categorisation (...) and the “fit” between the stimulus input and category specifications’ (Turner et al., 1987, p. 54).

Concerning the participants and the setting of our study, both relative accessibility and fit for a self-categorisation as female football players were highly likely. Being reminded of the negative stereotype should constitute a threat to this collective identity and activate strategies in order to maintain the individual participants' self-esteem – e.g. dissociation from the group of female football players.

This reasoning led us to considering collective identification as a potential mediator of the hypothesised effect of stereotype threat on the motor task. In our analyses, we further examined the following potential mediators: cognitive interference, subjective vulnerability for sexist discrimination, and subjective task difficulty.

MATERIALS AND METHODS

Procedure

In order to recruit participants, managers of amateur women's football teams in northern Germany were approached. Data collection took place between 15 January and 8 April 2016 at times of regular training sessions of the co-operating clubs. After preparation of the test setting (see **Figure 1**; the blue square indicates the placement of footballs at the beginning of each session, the yellow circles represent the position of the video cameras used to film the test), the experimenter and her assistant – both female – welcomed the participants. In order to standardise the setting as far as possible, the team managers (all of them male) were excluded from the test situation. Thus, no males were immediately present who could involuntarily trigger the social category ‘gender’.

Participants took pens and clip-boards containing background information about the study, a consent form, and a questionnaire from a basket. In particular, they were informed about the test session being video-recorded (in order to ensure the impression of the test being diagnostic). From the outside, the clip-boards were undistinguishable, thus ensuring a random self-allocation to experimental conditions.

The first part of the questionnaire comprised questions referring to age, footballing experience, self-efficacy, and subjective importance of the sport. The following instruction contained the experimental manipulation which was imbedded into a bogus description of the study's scientific basis. The threatened participants read:²

Although men and women do not directly compete playing football, one can state on a scientific basis that men outperform women in motor tasks concerning force and velocity (Knisel, Opitz, Wossmann, and Ketelhut, 2009). Research supposes that there are hardly differences between men and women concerning the capability characteristics concentration, aplomb, and precision. During this study, the shooting precision of women shall be video-recorded and analysed in order to advance research.

The text for non-threatened participants differed:

In the realm of football one can state on a scientific basis that there are individual differences in motor performance concerning force and velocity (Knisel, Opitz, Wossmann, and Ketelhut, 2009). In how far there are individual differences concerning the capability characteristic shooting precision has not yet sufficiently been researched. During

²Differences between the conditions are italicised here. Of course, during the study, the respective texts were presented without highlighting.

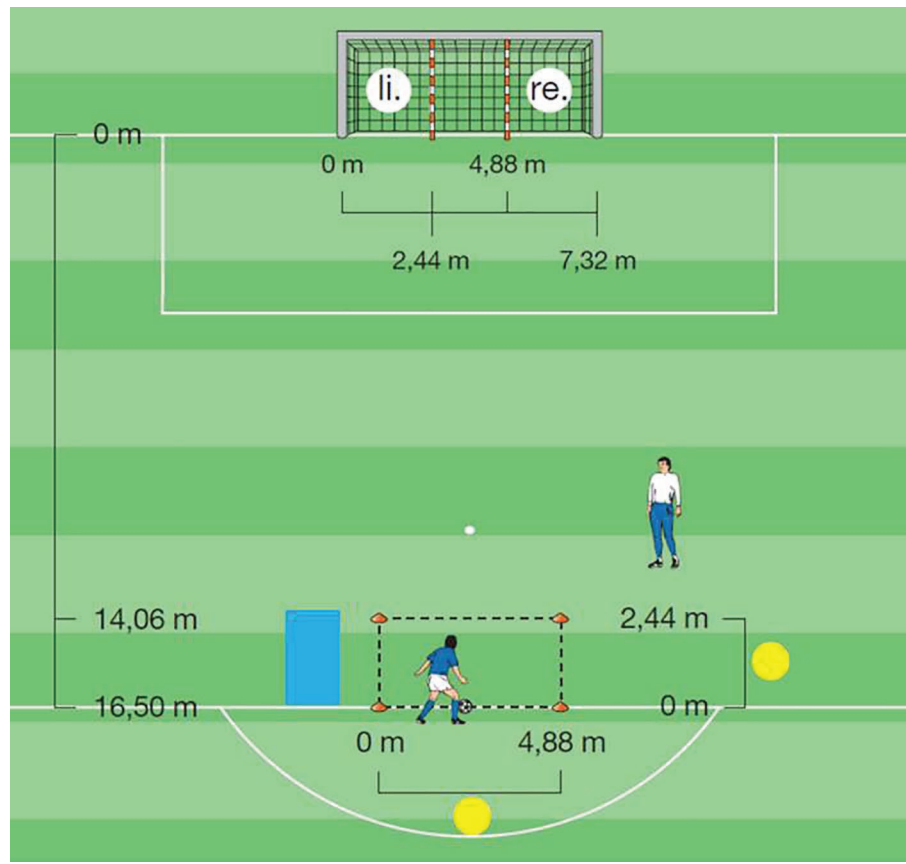


FIGURE 1 | Test setting [picture taken from Deutscher Fußball-Bund (n. d.b), and slightly amended].

this study, the shooting precision shall be video-recorded and analysed in order to advance research.

Hence, stereotype threat (the sole independent variable: threat vs. no threat) was manipulated between participants. The rest of the study was exactly the same for both threatened and non-threatened participants. The manipulation was followed by a written instruction for the actual motor task (taken from Deutscher Fußball-Bund, n. d.b): kicking the ball with maximum force from a designated square into either the left or right target field of the goal (see **Figure 1**). More specifically, the participants were supposed to take a ball from a pile, place it in front of the designated square, drive it into the square (first touch), and kick the rolling ball with either their left or right foot into a pre-specified area of the goal – either the left- or right-sided target field (second touch). Before each trial, the experimenter told the participants which foot to use and which target field to aim for. In order to enable familiarisation with their task, participants executed one test run followed by eight experimental trials.

After completion of the motor task, participants filled in the second part of the questionnaire, consisting of items concerning the manipulation check, subjective vulnerability for sexist discrimination, collective identification, cognitive interference, and subjective difficulty of the task. Upon completion, participants were thanked and debriefed.

Measures

Dependent Variables

The participants' task was to kick a football into specified areas of the goal. They had eight tries, each. Thus, the first DV, the number of hits, could range between 0 and 8. Additionally, the experimenter assessed the correct (or incorrect) execution of the task per try. If the player used the incorrect leg for kicking, and/or shot from outside the designated square the try was noted as a mistake. Thus, the second DV, the number of mistakes, could also range from 0 to 8.

Additional Variables

Some data were collected before the experimental manipulation. Concerning demographics only age was measured. As a potential control variable, participants indicated the number of years of experience as active football players. Moreover, we asked them to rate the importance they assign to football ('Football is very important to me.' and 'My sporting performance within the field of football is very important to me.'). (1) 'does not apply'; (2) 'does rather not apply'; (3) 'rather does apply'; (4) 'does apply'; Cronbach's $\alpha = 0.788$), and task-related self-efficacy ('I always know what to do when the task is scoring a goal', 'I always achieve the task of scoring a goal when I try to', 'I cope well with surprising goal-scoring tasks', 'I am looking forward to goal-scoring tasks calmly because I can trust my

abilities,' 'I never know how to deal with new goal-scoring tasks' [reverse coded], 'I am going to master the goal-scoring task because I can rely on my skills'; (1) 'does not apply'; (2) 'does rather not apply'; (3) 'rather does apply'; (4) 'does apply'; Cronbach's $\alpha = 0.824$).

After completion of the task, the participants answered some additional questions – each on a 4-point-scale ranging from 1 to 4. The first of them – 'During the goal-scoring task I thought about the cliché "Women are worse football players than men"' – served as manipulation check (plus potential mediator). The second one – 'The cliché "Women are worse football players than men" annoys me.' – is an indicator of subjective vulnerability for sexist discrimination. The third and fourth items measured collective identification with women, respectively, female football players whilst completion of the task: 'Whilst executing the goal scoring task I identified with the group of (women/female football players)'. Three more items were included in order to capture cognitive interference during task completion: 'I thought about other things during the goal-scoring task', 'Whilst completing the goal-scoring task I had difficulties to concentrate,' and 'My attention was completely focused on the goal-scoring task' (reverse coded); Cronbach's $\alpha = 0.741$. Finally, the participants assessed the subjective difficulty of the goal-scoring task: 'I perceived the goal scoring task as difficult', with higher values indicating larger difficulty.

Sample

Eighty³ female football players took part in the study. They were randomly assigned to the threat ($N = 41$) or no-threat ($N = 39$) condition. Their age ranged from 16 to 53, with a mean of 23.59 years ($SD = 7.634$). Their mean footballing experience was 12.09 years ($SD = 6.042$), ranging from 0 to 28 years. No participant quit during data collection or withdrew her consent afterwards. No cases were excluded from analysis.

³An *a priori* power analysis employing G*Power 3.1.9.2 (Faul et al., 2009) yielded a target sample size of 74, given an effect size $f^2 = 0.15$ and power = 0.95 for a one-tailed, or of 89 for a two-tailed test. *Post hoc*, calculated with the actual sample size of $N = 80$, and an effect size of $f^2 = 0.058$ [calculated as proposed by Selya et al. (2012)], the power equals 0.691 (one-tailed), or 0.569 (two-tailed). Data collection stopped when the sample size was considered sufficiently large – before any analyses were conducted.

RESULTS

Neither the items capturing importance, nor those referring to task-related self-efficacy – all measured before the experimental manipulation – showed a statistically significant difference between threat and no-threat condition; $|ts| < 0.982$, $ps > 0.329$ (concerning means and standard deviations of the scales or single-item measures mentioned above and below, see **Table 1**).

The same applies to some of the items measured after task completion – namely those measuring cognitive interference ($|ts| < 1.05$, $ps > 0.298$), subjective difficulty of the goal-scoring task [$t(78) = -0.289$, $p = 0.773$], or identification with women [$t(78) = -0.740$, $p = 0.462$]. The participants within the threat condition ($M = 2.73$, $SD = 1.001$) did, however, identify somewhat less with female football players than those not threatened ($M = 3.15$, $SD = 0.933$): $t(78) = 1.949$, $p = 0.055$. Participants' identification as female football players ($M = 2.94$, $SD = 0.985$) was stronger than that as women ($M = 2.25$, $SD = 1.131$), $t(79) = -5.488$, $p < 0.001$.

Whereas the subjective vulnerability for sexist discrimination did not significantly differ between conditions ($M_{\text{threat}} = 2.63$, $SD_{\text{threat}} = 1.067$; $M_{\text{no threat}} = 2.51$, $SD_{\text{no threat}} = 1.211$; $t(78) = -0.476$, $p = 0.635$), the manipulation check indicates that the intended stereotype threat worked: $M_{\text{threat}} = 1.83$, $SD_{\text{threat}} = 1.202$; $M_{\text{no threat}} = 1.05$, $SD_{\text{no threat}} = 0.320$; $t(78) = -3.911$, $p < 0.001$. It should be noted, however, that the majority of threatened participants (26, i.e. 63.4% versus 38, i.e. 97.4% of the non-threatened) chose the answer option 'does not apply'.

In accordance with our hypothesis, threatened participants scored less hits ($M = 1.44$, $SD = 0.923$) than those not threatened ($M = 2.00$, $SD = 1.556$). In order to statistically test⁴ this difference, we conducted a multiple regression analysis including two control variables, namely years of footballing experience and task-related self-efficacy. The results, shown in **Table 2**, indicate the significant role of stereotype threat on shooting precision⁵.

⁴Due to conventions, we conducted a two-tailed test in spite of our directional hypothesis.

⁵Without consideration of the control variables, the effect of stereotype threat on hits reaches a level of significance of $p = 0.052$ ($t(78) = 1.973$), indicating a suppressor effect.

TABLE 1 | Means and standard deviations.

Variables	Overall	No threat	Threat
	M (SD)	M (SD)	M (SD)
Importance	3.56 (0.509)	3.58 (0.466)	3.54 (0.552)
Task-related self-efficacy	2.91 (0.510)	2.90 (0.505)	2.91 (0.521)
Manipulation check	1.45 (0.967)	1.05 (0.320)	1.83 (1.202)
Subjective vulnerability for sexist discrimination	2.58 (1.134)	2.51 (1.211)	2.63 (1.067)
Collective identification with women	2.25 (1.131)	2.15 (1.159)	2.34 (1.109)
Collective identification with female football players	2.94 (0.985)	3.15 (0.933)	2.73 (1.001)
Cognitive interference	2.02 (0.750)	1.93 (0.742)	2.10 (0.757)
Subjective difficulty	2.41 (0.837)	2.38 (0.877)	2.44 (0.808)
Hits	1.71 (1.295)	2.00 (1.556)	1.44 (0.923)
Mistakes	0.71 (1.150)	0.64 (1.181)	0.78 (1.129)

TABLE 2 | Regression analysis with scored hits as criterion.

Predictor and control variables	<i>B</i>	<i>SE B</i>	β	<i>t</i>	<i>p</i> ^a	Uniquely explained variance (%)
				<i>t</i> (76)		
Stereotype threat (IV) ^b	-0.578	0.273	-0.224	-2.114	0.038	5.03
Task-related self-efficacy	0.754	0.276	0.297	2.728	0.008	8.38
Footballing experience	0.010	0.023	0.047	0.436	0.664	0.21

$F^2 = 0.144$; $F_{adj}^2 = 0.110$; $F(3, 76) = 4.269$, $p = 0.008$; $VIFs < 1.06$.

^aTwo-tailed.

^bNo threat = 0; threat = 1.

The number of mistakes, however ($M_{threat} = 0.78$, $SD_{threat} = 1.129$; $M_{no\ threat} = 0.64$, $SD_{no\ threat} = 1.181$) did not significantly differ. Again, a multiple regression analysis, with the same control variables, was conducted. The results are shown in **Table 3**.

Concerning hits as well as mistakes, the role of task-related self-efficacy is almost identical – it explains 8.38% of the variance in the former and 8.22% in the latter. Footballing experience did not significantly contribute to variance explanation in either analysis. Thus, the pattern between the prediction of hits and mistakes deviates only for stereotype threat.

In order to test potential indirect effects of stereotype threat on hits and mistakes, we conducted mediation analyses (see **Figure 2** for a depiction of the model) as proposed by Hayes (2013).

The first six analyses contained stereotype threat as predictor, hits as criterion, task-related self-efficacy and experience as controls and one of the following potential mediators per calculation: thinking about the negative stereotype (our manipulation check), subjective vulnerability for sexist discrimination, collective identification with women, collective identification with female football players, cognitive interference, or subjective difficulty of the task. Each of the 95% confidence intervals for the indirect effect contained the 0 (10,000 bootstrap samples), indicating no statistically significant mediation effect. The second set of six analyses differed only concerning the criterion – now mistakes. Again, no statistically significant indirect effect was found.

Path a of the above mentioned analyses – between predictor and mediator – reached respectively approached statistical significance for two predictors: stereotype threat significantly affected thinking about the negative stereotype ($b = 0.784$, $\beta = 0.408$, $p < 0.001$) and collective identification with female football players ($b = -0.418$, $\beta = -0.214$, $p = 0.055$).

Additionally, we tested whether stereotype threat exerted an indirect effect on hits *via* mistakes. This could be another indicator of (conscious) cognitive interference as a mediator. Again, task-related self-efficacy and experience were included as control variables, and again, the 95% confidence interval included 0.

Before discussing implications of the study's results, it should be considered whether a stereotype effect was – according to Keller's (2008) criteria – likely to emerge. Apparently, the stereotype of women being unable to play football proper does fit excellently into the (experimental) situation and is directly applicable to the task. Results indicate that football (the relevant domain) is important to the participants: means for the respective items (≥ 3.45) are close the scale's maximum (4). The demanding character of the task is indicated by different measures: the relatively low

number of hits (overall mean = 1.71), the overall mean (2.41, i.e. above the scale midpoint) for the subjective difficulty measure, and some participants' oral feedback after test sessions. The impression of one's performance being under scrutiny is highly likely to have arisen amongst participants due to the test setting and procedure. Again, participant feedback confirms this assumption. Summing up, it appears reasonable to consider the four criteria met and expect a stereotype effect likely to emerge.

DISCUSSION

In line with our hypothesis, threatened participants scored less hits. They did not, however, make significantly more mistakes during task execution. Thus, the rather difficult motor task – tapping in well-learned, automated motion sequences – was impeded when the negative stereotype was salient; the mainly cognitive task of following instructions correctly, was not. This is surprising insofar as the stereotype threat effect was introduced into the psychological literature as one affecting cognitive/intellectual tasks (Steele and Aronson, 1995). The same applies to our finding that neither one of the three single indicators of cognitive interference nor the respective scale were affected by threat. Still, it is not impossible that the stereotype threat affected working memory *via* verbal ruminations or worries (Rydell et al., 2009); both concepts are not (well) captured by our measures of cognitive interference. In this case, however, one would expect both fewer hits and more mistakes within the threat condition. It should be noted that the average number of mistakes is rather low (0.71 overall), which might be indicating a comparably low difficulty of the task of following the instructions correctly – making the manifestation of a stereotype effect on this particular measure less likely. Employing a more sensitive measure of mistakes and/or more demanding instructions seem appropriate strategies for potential replications of our study.

The fact that stereotype threat hampered goal scoring complies with the reasoning of Schmader et al. (2008) insofar as their integrated process model predicts an effect on sensorimotor tasks. Indeed, it seems plausible that the threat 'in combination with the motivation to disconfirm the stereotype, translates into a strong motivation to avoid failure' which in turn lets the participants 'focus attention on themselves and their performance, becoming more vigilant to detect signs of failure' (Schmader et al., 2008, p. 343). Keller and Dauenheimer (2003) hypothesised a mediation of stereotype threat effects on

TABLE 3 | Regression analysis with mistakes as criterion.

Predictor and control variables	B	SE B	β	t	p ^a	Uniquely explained variance (%)
				t(76)		
Stereotype threat (IV) ^b	0.154	0.249	0.067	0.617	0.539	0.45
Task-related self-efficacy	-0.663	0.252	-0.294	-2.631	0.010	8.22
Footballing experience	-0.008	0.021	-0.040	-0.360	0.720	0.15

$R^2 = 0.097$; $R^2_{adj} = 0.062$; $F(3, 76) = 2.726$, $p = 0.050$; $VIFs < 1.06$.

^aTwo-tailed.

^bNo threat = 0; threat = 1.

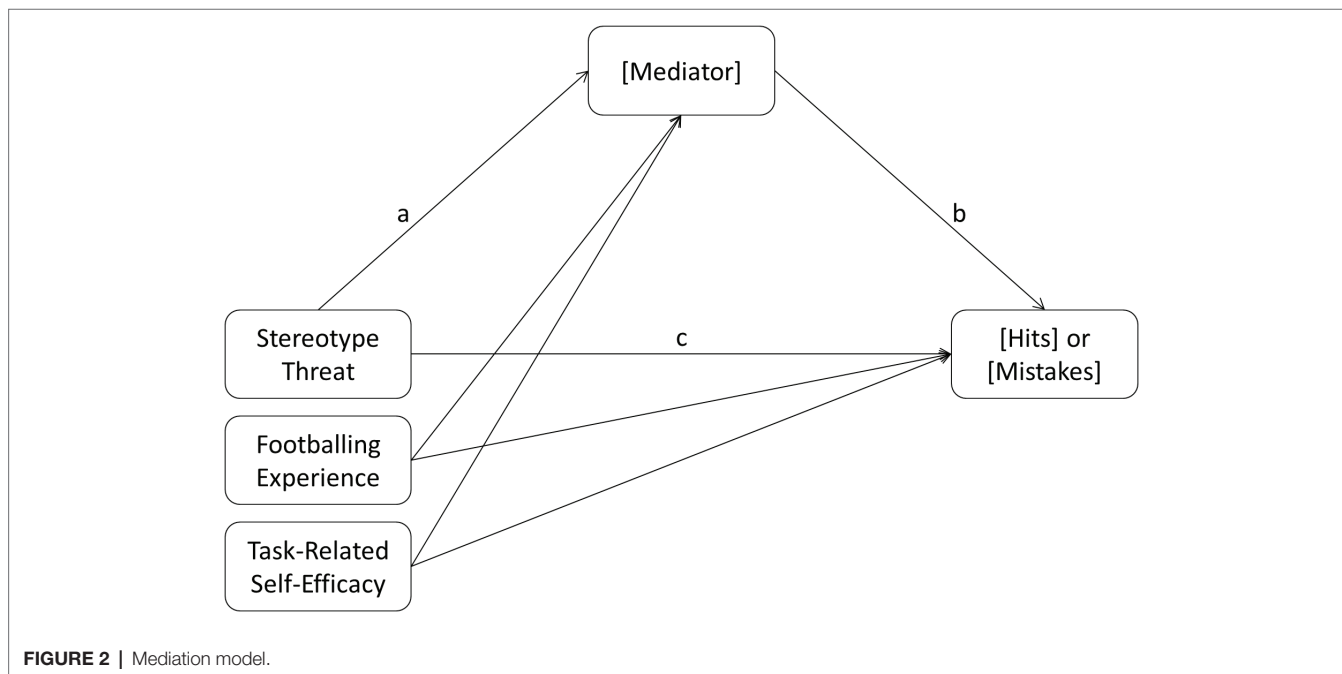


FIGURE 2 | Mediation model.

performance *via* regulatory focus and tested it with specific emotions typical of one or the other regulatory focus; they found empirical support for at least one kind of emotions: dejection significantly mediated the effect of gender-related stereotype threat on women's performance in a mathematics test. Seibt and Förster (2004) argue similarly when they introduce regulatory focus as mediator between stereotype threat and performance: being confronted with a negative stereotype would induce a prevention focus, whereas positive stereotypes foster a promotion focus. The former is associated with 'safety and nonlosses', the latter with 'advancement and gains' (Seibt and Förster, 2004, p. 39). Applied to our study, with accuracy at the core of the task, however, both regulatory foci should motivate the participants to score as many hits as possible, either in order to maximise a positive outcome (i.e., hits) or to minimise a negative outcome (i.e., misses). Concerning this measure, thus, our results do not support this line of reasoning. The same applies to the number of mistakes: 'promotion focus emphasis on strategic eagerness should lead to a riskier processing style' (Seibt and Förster, 2004, p. 39) presumably resulting in more mistakes, 'whereas prevention focus emphasis on strategic

vigilance should lead to a more careful processing style' (Seibt and Förster, 2004, p. 40), presumably resulting in fewer mistakes. The number of mistakes, however, did not differ significantly. Because we did not assess the participants' regulatory focus, an evaluation of our results from this particular perspective lacks empirical data. Therefore, the above reasoning needs to be considered cautiously. With regard to Schmader et al. (2008), our measures of cognitive interference could be interpreted as indicators of ongoing monitoring processes. In this case, threatened participants should score significantly higher on the respective items, which is not the case. Nevertheless, we consider it plausible that the pursuit of a certain regulatory focus could be a strategy in order to maintain high self-esteem when an individual's collective identity is threatened. Further research, integrating social identity and regulatory focus approaches, could help to understand the respective constructs' relations better.

Moreover, the lack of influence of threat on (our indicators for) cognitive performance goes hand in hand with a lack of evidence for an assumed mediating effect on motor tasks: our findings do not corroborate the idea that stereotype threat strains sparse cognitive resources. An explanation considering threat's

impact on collective identification appears more promising. Admittedly, the indirect path of threat *via* identification as female football player on hits was not statistically significant. However, at least the path from threat to identification approached statistical significance ($p = 0.055$). An interesting aspect of this finding is the contrast to the second measure of collective identification – that with women *per se* – where threat made no difference. According to the SCT, ‘given two equally “accessible” categories, the one that better “fits” the perceptual data will become salient’ (Turner et al., 1987, p. 55). Here, in the context of a football-related task, the narrower, more specific social group of female football players can reasonably be expected to fit better than the broader, rather unspecific group of females *per se* (both categories can be expected to be very well accessible to our participants). The overall means of identification reflect this assumption as the participants identified more strongly as female football players than as women. In the light of the SCT, the negative impact of our manipulation on the identification as female football players, but not females *per se*, makes perfect sense: apparently, they employed an individual mobility strategy in order to maintain self-esteem. Because identification was measured after the goal scoring task, it is also feasible that poor performance mediated the effect. However, this was not found to be the case.

Although we did not find compelling evidence for an influence of identification as female football player on the task, such a relationship between collective identification and performance should not be categorically ruled out. In fact, the respective link is well established, e.g. in the context of organisational outcomes (Haslam et al., 2003), or *via* motivation (Van Knippenberg, 2000); moreover, social identity is perceived as major factor in leadership – hence, again related to performance (Haslam et al., 2011). Further research should employ different operationalisations of a wider range of collective identities and consider potential moderators or mediators. Moreover, future samples should be larger in order to secure high statistical power. Although we found a significant effect of threat on hits, a lack of power may have contributed to our not being able to detect a full mediation, or a more than only marginally significant effect of threat on identification as female football player.

Another fruitful idea for future research could be an investigation into task-related self-efficacy as a potential mediator.

REFERENCES

- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., and Brown, J. (1999). When white men can't do math: necessary and sufficient factors in stereotype threat. *J. Exp. Soc. Psychol.* 35, 29–46. doi: 10.1006/jesp.1998.1371
- Brüggemeier, F.-J. (2006). Vom Randphänomen zum Massensport. *Inf. polit. Bild.* Available at: <https://www.bpb.de/izpb/8749/vom-randphaenomen-zum-massensport> (Accessed May 26, 2018).
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., and Latinotti, S. (2003). Stereotype threat: the effect of expectancy on performance. *Eur. J. Soc. Psychol.* 33, 267–285. doi: 10.1002/ejsp.145
- Chalabaev, A., Sarrazin, P., Stone, J., and Cury, F. (2008). Do achievement goals mediate stereotype threat? An investigation on females' soccer performance. *J. Sport Exerc. Psychol.* 30, 143–158. doi: 10.1123/jsep.30.2.143

Given its reliable and comparably strong influence on hits as well as mistakes, it appears worthwhile to test whether a stereotype threat reduces task-related self-efficacy.

Summing up, we could demonstrate the stereotype effect for a complex and demanding motor task. In the light of the ongoing replicability crisis (e.g. Świątkowski and Dompnier, 2017), our findings can therefore strengthen the confidence in the existence of the stereotype effect. Hypothesised mediational paths could, however, not be corroborated. Especially interesting is the absence of evidence for an influence of threat *via* cognition on motor performance. Therefore, it might be wise to take a step back and reconsider possible pathways between stereotype threat and motor tasks, then try to replicate the study and bolster it with measures relating to this re-developed theory.

ETHICS STATEMENT

This study was carried out in accordance with the ‘Berufsethische Richtlinien des Berufsverbandes Deutscher Psychologinnen und Psychologen e.V. und der Deutschen Gesellschaft für Psychologie e.V.’ (professional ethics guidelines of the occupational union of German psychologists and the German society of psychology). All subjects gave written informed consent in accordance with the Declaration of Helsinki. An ethics approval was not required as per the university's guidelines or national regulations. The research presented no discernible risk to the participants.

AUTHOR CONTRIBUTIONS

HG wrote the article and supervised data collection. MK contributed the initial research idea, collected the data, and contributed information during writing. HG and MK developed the research idea and hypotheses, designed the study, and analysed the data.

FUNDING

We acknowledge financial support by Land Schleswig-Holstein within the funding programme Open Access Publikationsfonds.

- Christenson, M., and Kelso, P. (2004). Soccer chief's plan to boost women's game? Hotpants. *The Guardian*. Available at: <https://www.theguardian.com/uk/2004/jan/16/football.gender> (Accessed June 07, 2019).
- Csizma, K. A., Wittig, A. F., and Schurr, K. T. (1988). Sport stereotypes and gender. *J. Sport Exerc. Psychol.* 10, 62–74. doi: 10.1123/jsep.10.1.62
- Deutscher Fußball-Bund (n. d.a). Mitglieder-Statistik 2017. Available at: https://www.dfb.de/fileadmin/_dfbdam/141295-Mitglieder-Statistik_2017_final.pdf (Accessed December 16, 2018).
- Deutscher Fußball-Bund (n. d.b). *Testmanual für die technomotorische Leistungsdiagnostik*.
- Diekmann, A. B., and Eagly, A. H. (2000). Stereotypes as dynamic constructs: women and men of the past, present, and future. *Personal. Soc. Psychol. Bull.* 26, 1171–1188. doi: 10.1177/0146167200262001
- Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

- Fiske, S. T., Cuddy, A. J. C., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* 82, 878–902. doi: 10.1037/0022-3514.82.6.878
- Haslam, S. A., Egghins, R. A., and Reynolds, K. J. (2003). The ASPIRe model: actualizing social and personal identity resources to enhance organizational outcomes. *J. Occup. Organ. Psychol.* 76, 83–113. doi: 10.1348/096317903321208907
- Haslam, S. A., Reicher, S. D., and Platow, M. J. (2011). *The new psychology of leadership: Identity, influence and power*. Hove: Psychology Press.
- Hayes, A. F. (2013). "Introduction to mediation, moderation, and conditional process analysis: a regression-based approach" in *Methodology in the social sciences*. New York, NY: Guilford Press.
- Hoffmann, E., and Nendza, J. (2007a). Der Deutsche Fußball-Bund hebt das Frauenfußballverbot auf. Available at: <http://www.bpb.de/gesellschaft/gender/graue-spielzeit/65068/dfb-hebt-verbot-auf> (Accessed December 16, 2018).
- Hoffmann, E., and Nendza, J. (2007b). DFB verbietet seinen Vereinen Damenfußball. Available at: <http://www.bpb.de/gesellschaft/gender/graue-spielzeit/65063/das-dfb-verbot?p=all> (Accessed December 16, 2018).
- Högler, W., Blimkie, C. J., Cowell, C. T., Inglis, D., Rauch, F., Kemp, A. F., et al. (2008). Sex-specific developmental changes in muscle size and bone geometry at the femoral shaft. *Bone* 42, 982–989. doi: 10.1016/j.bone.2008.01.008
- Hottenrott, K., and Neumann, G. (2012). Geschlechtsspezifische Formel für optimale Trainingsherzfrequenzen. *Schweiz. Z. Sportmed. Sporttraumatol.* 60, 102–105.
- Keller, J. (2008). "Stereotype als Bedrohung" in *Stereotype, Vorurteile und soziale Diskriminierung*, eds. L.-E. Petersen and B. Six (Weinheim: Beltz), 88–92.
- Keller, J., and Dauenheimer, D. (2003). Stereotype threat in the classroom: dejection mediates the disrupting threat effect on women's math performance. *Personal. Soc. Psychol. Bull.* 29, 371–381. doi: 10.1177/0146167202250218
- Keller, T. (2018). Bei diesem Frauenfußball-Bild scrollt ihr gleich weiter? Das nervt! Available at: <https://www.bento.de/sport/frauenfussball-frauen-werden-nicht-ernstgenommen-maenner-schon-warum-a-00000000-0003-0001-0000-000002357342> (Accessed December 16, 2018).
- Knisel, E., Opitz, S., Wossmann, M., and Ketelhut, K. (2009). Sport motivation and physical activity of students in three European schools. *Int. J. Phys. Educ.* 46, 40–53.
- Kunz, M. (2007). 265 million playing football. Available at: https://www.fifa.com/mm/document/fifafacts/bcoffsurv/emaga_9384_10704.pdf (Accessed May 26, 2018).
- Lueptow, L. B., Garovich, L., and Lueptow, M. B. (1995). The persistence of gender stereotypes in the face of changing sex roles: evidence contrary to the sociocultural model. *Ethol. Sociobiol.* 16, 509–530. doi: 10.1016/0162-3095(95)00072-0
- Merkel, M. (2003). "Ein Elend für die Augen". Available at: <http://www.taz.de/!703861/> (Accessed December 15, 2018).
- Pfister, G. (2015). Assessing the sociology of sport: on women and football. *Int. Rev. Sociol. Sport* 50, 563–569. doi: 10.1177/1012690214566646
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management. *J. Pers. Soc. Psychol.* 74, 629–645. doi: 10.1037/0022-3514.74.3.629
- Rudman, L. A., and Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *J. Soc. Issues* 57, 743–762. doi: 10.1111/0022-4537.00239
- Rydell, R. J., McConnell, A. R., and Beilock, S. L. (2009). Multiple social identities and stereotype threat: imbalance, accessibility, and working memory. *J. Pers. Soc. Psychol.* 96, 949–966. doi: 10.1037/a0014846
- Schmader, T., Johns, M., and Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychol. Rev.* 115, 336–356. doi: 10.1037/0033-295X.115.2.336
- Schwaiger, R. (2011). Der unbegründete Hype um die Frauenfußball-WM. Available at: <https://www.profil.at/home/frauenfussball-der-hype-frauenfussball-wm-301630> (Accessed December 16, 2018).
- Scruton, S., Fasting, K., Pfister, G., and Bunuel, A. (1999). It's still a man's game? The experiences of top-level European women footballers. *Int. Rev. Sociol. Sport* 34, 99–111. doi: 10.1177/101269099034002001
- Seibt, B., and Förster, J. (2004). Stereotype threat and performance: how self-stereotypes influence processing by inducing regulatory foci. *J. Pers. Soc. Psychol.* 87, 38–56. doi: 10.1037/0022-3514.87.1.38
- Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., and Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2 , a measure of local effect size, from PROC MIXED. *Front. Psychol.* 3:111. doi: 10.3389/fpsyg.2012.00111
- Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28. doi: 10.1006/jesp.1998.1373
- Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *Am. Psychol.* 52, 613–629. doi: 10.1037/0003-066X.52.6.613
- Steele, C. M., and Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69, 797–811. doi: 10.1037/0022-3514.69.5.797
- Stone, J., Lynch, C. I., Sjomeling, M., and Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *J. Pers. Soc. Psychol.* 77, 1213–1227. doi: 10.1037/0022-3514.77.6.1213
- Świątkowski, W., and Dompnier, B. (2017). Replicability crisis in social psychology: looking at the past to find new pathways for the future. *Int. Rev. Soc. Psychol.* 30, 111–124. doi: 10.5334/irsp.66
- Tajfel, H., and Turner, J. C. (1986). "The social identity theory of intergroup behavior" in *Psychology of intergroup relations*. 2nd Edn. eds. S. Worchel and W. G. Austin (Chicago: Nelson-Hall), 7–24.
- Tempel, T., and Neumann, R. (2016). Gender role orientation moderates effects of stereotype activation on test performances. *Soc. Psychol.* 47, 63–73. doi: 10.1027/1864-9335/a000259
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., and Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Basil Blackwell.
- Van Knippenberg, D. (2000). Work motivation and performance: a social identity perspective. *Appl. Psychol.* 49, 357–371. doi: 10.1111/1464-0597.00020
- Vaske, H. (Director) (2006). *The art of football von A bis Z* [DVD].
- Williams, J. E., and Best, D. L. (1990). *Sex and psyche: Gender and self viewed cross-culturally (1. pr)*. Cross-cultural research and methodology series. Vol. 13, Newbury Park, CA: Sage Publications.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Grabow and Kühl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.