An integrative study of facultative personality calibration

Christoph J. von Borell [a, b, c*], Tobias L. Kordsmeyer [a,b], Tanja M. Gerlach [a,b], Lars Penke [a,b]

[a] Department of Biological Personality Psychology, Georg Elias Müller Institute of Psychology, University of Goettingen, 37073 Göttingen, Germany

[b] Leibniz ScienceCampus Primate Cognition, 37077 Göttingen, Germany

[c] Research Training Group 2070 – Understanding Social Relationships, University of Goettingen, 37073 Göttingen, Germany

[*]Corresponding author:

Christoph J. von Borell

christoph.borell@psych.uni-goettingen.de

The theory of facultative calibration, which explains personality differences as responses to variation in other phenotypic traits of individuals, received mixed results throughout the last years. Whereas there is strong evidence that individual differences in human behavior are correlated with the self-perception of other traits, it still needs to be questioned whether they are also adjusted to objective differences in body condition (i.e. formidability). In two independent studies (N_1 = 119 men and 124 women, N_2= 165 men) we tested hypotheses of facultative personality calibration in an integrative way, assessing various outcomes of previous studies in the same samples (including Anger Proneness, Extraversion, Neuroticism, Narcissism, Shyness, Vengefulness, and Sociosexual Orientation). Formidability was derived from assessments of physical strength and various anthropometric measures from full-body 3D scans and paired with measures of self-perceived and other-rated physical attractiveness (based on rotating morphometric 3D body models and facial photographs). We could replicate positive correlations with self-perceived attractiveness across outcomes, though these were not corroborated by more objective assessments of attractiveness: an effect of other-rated attractiveness was clearly not supported in our results for either sex, regardless of the personality outcome. Anthropometric measures and physical strength were also largely unrelated to personality, with the exception of Extraversion, Utility of Personal Aggression, and Sociosexual Orientation. While the two samples differed in their results for domain-level Extraversion, at least the Extraversion facets Activity and Assertiveness were related to strength and masculinity in men. For Sociosexual Orientation the results of our two samples varied more substantially, a positive association was only present in Study 2. Future studies need to clarify whether formidability, potentially an indicator of genetic quality for males, enhances their orientation and success in short-term mating. Furthermore we propose longitudinal twin-difference studies as means to evaluate the theory of personality recalibration in a more controlled manner.

Keywords: facultative calibration, personality traits, reactive heritability, relative bargaining power, sociosexual orientation, state-behavior feedback, anthropometrics, physical attractiveness

1. Introduction

Throughout the last decades different theories emerged that explain personality variation among individuals from an evolutionary perspective (Penke et al., 2007; Buss & Hawley, 2010; Buss & Penke, 2015). One of them, the theory of facultative calibration, poses that individual differences in personality may not be actively selected for, but are calibrated to differences in other traits of individuals, such as physical strength, physical attractiveness, or intelligence (Haysom et al., 2015; Lukaszewski & Roney, 2011; Sell et al., 2009; Tooby & Cosmides, 1990). These traits are thought to enhance the formidability or expected relative bargaining power (RBP) of individuals, i.e. the ability to inflict costs on others or to extract benefits from them (Petersen et al., 2010). The heritability of personality should then be of reactive manner, linked to the heritability of these specific traits. Notably, such explanation of personality variation is not exclusive to humans. In animal behavioral ecology, a similar idea has been put forward as "state-behavior feedback loop", reflecting that behavior and behavioral repeatability (the stability of a behavioral trait throughout time) may be adaptively adjusted to slower-changing or fixed state variables such as size, energy reserves, or parasite infection (Sih et al., 2015; Wolf & Weissing, 2010). During the last years, various studies of humans and non-human animals empirically tested such links between behavioral and other phenotypic traits, however not necessarily under the same label. When we subsequently refer to the theory of facultative calibration in humans, we intend it to subsume synonymous terms as "recalibrational theory" or "condition-dependent calibration" that have been used in the literature before.

One of the first studies that explicitly tested facultative calibration in humans predicted a relationship between RBP and anger proneness (Sell et al., 2009). It was reasoned that differences in RBP would lead to differences in the perception of a personal welfare trade-off ratio WTR (i.e., how much an actor A would expect an actor B to value his welfare compared to B's own; Tooby et al., 2008), which in turn should affect the propensity to feel and express

anger (see Sell et al., 2009). Furthermore, the authors predicted sex differences in which variables would factor into an individual`s RBP. Since men are usually stronger and tend to monopolize the use of force in social negotiations, strength should factor strongly into men's RBP (by the means of increased physical formidability; see Lukaszewski, 2013). In women, RBP should be more closely linked to physical attractiveness, a main indicator of fertility and reproductive potential. Given that access of males to female sexuality is more limited than the other way round, attractiveness should thus provide a powerful leverage to women. Indeed, Sell et al. (2009) found upper body strength to be positively correlated with a variety of anger-relevant measures in men, but not in women. Self-perceived attractiveness, in turn, showed positive relationships to anger in females, but only a few such associations were found in men.

Price et al. (2012) extended these results by linking anger proneness to a greater variety of anthropometric measures (e.g. chest circumference, bicep circumference, waist circumference, and overall body shape masculinity). They found measures of body shape and upper body masculinity in particular to be positively related to proneness to anger in men. However, this pattern was merely evident in a subsample of younger participants aged 18 to 23. In contrast, neither physical masculinity, nor anthropometric measures of attractiveness (such as waist-hip ratio or BMI) were related to proneness to anger in female participants, with the exception of leg-body ratio (LBR). Nevertheless they could replicate the results of Sell et al. (2009) showing a positive relationship between different measures of self-perceived attractiveness and anger for females (but not for males). Recent findings from a large sample of Swiss adolescents (N = 1447; Sell, Eisner, & Ribeaud, 2016) found no correlation between height and aggressiveness, but a small relationship of aggressiveness with weight and BMI. However, in multiple regressions also including overall fighting ability (a composite measure based on self-report items and flexed biceps circumference), these effects disappeared. Also, a single effect of biceps circumference was no longer significant when controlling for the self-

report measure of fighting ability. Hence the authors conclude that actually fighting ability predicts aggressive bargaining rather than individual anthropometric measures per se. Archer and Thanzami (2007) suggested a more fine-grained differentiation in anger and aggression related outcomes. While they found a relationship between trait measures of direct (physical) aggressiveness with height, weight, and strength in a non-western sample of young Indian men, they did not find a correlation between physical formidability and proneness to anger. In fact, based on the concept of Resource Holding Power (i.e., the ability to win a fight; Parker, 1974; Stulp et al., 2012) they explicitly hypothesized size and strength to be unrelated to indirect measures of aggressiveness such as anger or hostility. Furthermore, it has been argued that vengefulness, the propensity to harm others or withhold benefits *in response* to a previous cost-inflicting or benefit withholding event, could be adjusted based on mechanism of facultative calibration as well (McCullough et al., 2011).

Lukazsewski and Roney (2011) hypothesized facultative calibration of extraversion, arguing that since extraverted individuals, as compared to more introverted individuals, are more likely to proactively seek social status, influence, and relationships, they will be exposed more often to conflicts of interest with others. The entailed cost-benefit ratio should thus be more favorable to stronger individuals and also to more attractive individuals, since they are known to be preferred in relationships and cooperative exchanges. However, they did not predict an isolated effect of facultative calibration but an integrative model wherein facultative calibration and a pleiotropic genetic effect acting on both extraversion and somatic features influence individual variability in extraversion together. In fact, they found extraversion to be related to other-rated- and self-perceived attractiveness in both sexes, to physical strength in men and independently to a polymorphism in the androgen receptor gene in men (which had been previously linked to strength and extraversion, though overall results are mixed). Their results therefore imply that the heritability of extraverted personality is comprised of both the

influence of genetic polymorphisms acting (rather) directly on behavioral trait regulation and of reactive heritability reflected in facultative calibration. However, in a subsequent study, Lukaszewski (2013) could not replicate a relationship between other-rated attractiveness and various personality traits (e.g. extraversion, emotionality, and fear of rejection) in either sex. Physical strength was again related to extraversion, this time in both sexes. Similar results have been found in the Tsimane, a group of forager-horticulturalists living in the Bolivian amazon (von Rueden et al., 2015). In this study, physical strength in both sexes was related to the personality dimension of Prosocial Leadership Orientation, which is specific to the Tsimane and represents a mixture of high Extraversion, high Agreeableness, high Openness to Experience, and low Neuroticism (based on items of the Big Five Inventory). Physical strength explained about 15% of the additive heritability of Prosocial Leadership Orientation (von Rueden et al., 2015).

Furthermore, narcissism, a complex psychological trait typically correlated to extraversion which includes feelings of superiority, entitlement, and power (Paulhus & Williams, 2002; Wetzel et al., 2016), has been shown to form an instantiation of a personality-formidability correlation, as it is associated with physical attractiveness (Holtzman & Strube, 2010). On the one hand, a correlation between narcissism and physical attractiveness could be explained in evolutionary terms by a convergence of selection pressures on narcissistic traits and physical attractiveness in short-term mating contexts (Holtzman & Strube, 2010). On the other hand, the proposed association also dovetails with the notion of a relationship between anger proneness and formidability / RBP, since the sense of entitlement is (a) involved in setting a WTR, influencing subsequent expressions of anger (Sell et al., 2009), and (b) constitutes a core component of the narcissistic personality disposition (Ackerman et al., 2011; Campbell et al., 2010). In a meta-analysis, Holtzman and Strube (2010) calculated a mean correlation of 0.14 between measures of narcissism and other-rated physical attractiveness.

Finally, facultative calibration has been proposed as a mechanism of individual differences in sociosexual orientation. Lukaszewski et al. (2014) found positive correlations of composite scores of physical strength and attractiveness (comprising both self- and other-rated measures) with an uncommitted (short-term) mating orientation in men. These relationships were predicted to emerge because strength and attractiveness were assumed to serve as cues of genetic quality and were thus likely to be preferred by ancestral women in uncommitted mating. Additionally, physical strength was most likely an asset in intrasexual contests arising in this setting. Such relationships were neither predicted nor found for women (Lukaszewski et al., 2014).

Notably, several studies have challenged the theory of facultative personality calibration in recent years. For instance, a longitudinal study on the ontogeny of aggressiveness in children showed that boys with greater aggressive and antisocial tendencies at age 11 had greater increases in physical strength during the following six years of puberty, while not being consistently stronger than their peers at age 11 (Isen et al., 2015). These findings are not consistent with the theory of facultative calibration. Thus, Isen et al. (2015) proposed a joint hormonal mediation of behavioral and physiological traits as an alternative explanation. Haysom et al. (2015) found no correlation between extraversion and height or BMI in men or women of a large twin sample (N=1659). In addition, low but significant phenotypic correlations between extraversion and facial attractiveness were not genetically mediated in this study and could also be explained by general learning processes (Haysom et al., 2015). Overall, the theoretical foundations of facultative calibration have been discussed in a critical light by Zietsch (2016), since for example the large mutational target size of complex behavioral traits likely affects personality variation, or strategies counter to previously proposed mechanisms could be at work (e.g. physically unattractive men may try to attract females with extraverted behavior). However, in some points, fairness must be maintained with previous studies who empirically tested possible explanations of the proposed optimal strategy (e.g. by showing a

correlation between self-perceived bargaining power and fear of rejection; Lukaszewski, 2013) or did simply not claim that facultative calibration is the only mechanism driving variation in personality related traits (Lukaszewski & Roney, 2011).

In sum the literature shows convincing evidence for a relationship between self-rated trait measures (such as self-perceived physical attractiveness) and personality, but only mixed results or missing evidence for relationships between personality and other-rated or direct anthropometric measures of attractiveness or formidability (see Table 1 for an overview). Thus it could be possible that the heritability of personality variation is, contrary to the theory of facultative calibration, not reactive to heritability of other phenotypic traits, but solely related to how people *perceive* themselves in these traits. In the latter case the causal effect could go in the other direction, with individuals with certain personalities having a tendency to perceive themselves as more attractive or formidable, independent of their objective physique.

The aim of the present study was to probe the relationship between personality traits, formidability and physical attractiveness in an integrated way. In order to do so, we collected a broad range of formidability indicators and personality traits in two relatively large community samples from two countries, the UK and Germany.

Table 1

An Overview of main results in previous studies of facultative calibration

Women	Self-perceived attractiveness	Other-rated attractiveness	Physical strength	Anthropometric Measurements	References

Anger	Positive association	Not tested	No association	No association	Price et al.; 2012, Sell et al., 2009
Extraversion	Positive association	Mixed Results	Mixed results	No association	Fink et al., 2016; Haysom, 2015; Lukaszewski, 2013; Lukaszewski & Roney, 2011, von Rueden et al., 2015
Narcissism	Positive association	Positive association	Not tested	Not tested	Gabriel et al., 1994; Holtzman & Strube, 2010
Sociosexual orientation	Unclear*	Unclear*	Unclear*	Not tested	Lukaszewski et al., 2014

Men	Self-perceived attractiveness	Other-rated attractiveness	Physical strength	Anthropometric Measurements	References
Anger proneness	Mixed Results	Not tested	Mixed results	Mixed results	Archer & Thanzami, 2007; Price et al., 2012; Sell et al., 200; Sell et al., 2016
Extraversion	Positive association	Mixed results	Positive association	No association	Fink et al., 2016; Haysom, 2015; Lukaszewski, 2013; Lukaszewski & Roney, 2011, von Rueden et al., 2015
Narcissism	Positive association	Positive association	Not tested	Not tested	Gabriel et al., 1994; Holtzman & Strube, 2010

Sociosexual orientation	Unclear*	Unclear*	Unclear*	Not tested	Lukaszewski et al., 2014

*No separate tests of significance for self-rated and objective trait measures

2. Study 1

2.1 Methods

2.1.1 Participants and Procedure

A total of 119 men and 124 women from Edinburgh (Scotland, UK) participated in the study (age 18-29 years, M = 21.5, SD = 2.1; years of completed education M = 15.6, SD = 3.5). Participants were either undergraduate students recruited through the student subject pool of the University of Edinburgh and received course credit, or came from the local community, recruited via social networks and small advertisements, and received a compensation of £10. A total of 83.1% of the sample indicated their ethnicity as white, while the rest self-identified as Indian (2.9%), Chinese (5.3%), mixed (4.1%), or other (4.5%).

Laboratory assessments were conducted by same-sex experimenters. After signing an informed consent form, participants were seated upright in front of a 3DMD facial camera system with glasses and facial jewelry removed, asked to maintain a neutral expression, and a 3D picture of their faces was taken. Then they were asked to change into provided, tight fitting standardized underwear and scanned three times standing upright with a TC^2 NX-16 3D body scanner (Cary, NC, USA), following the procedure described in Price et al. (2012). Afterwards height and weight was measured with a statiometer and a digital scale, respectively. Flexed biceps circumference was measured with an anthropometric tape measure. Hand grip and upper body strength was measured with a dynamometer, and lung function with a spirometer. Finally, participants filled out a computerized battery of questionnaires (described below) in private.

The study was approved by the University of Edinburgh Psychology Research Ethics Committee (application numbers 25-1112, 299-1112, 40-1213/2).

2.1.2 Measures

2.1.2.1 Personality Measures

The *anger* measurements were adopted from Sell et al. (2009), measured on a 5-point Likert scale from 1 = "strongly disagree" to 5 = "strongly agree", and included *Proneness to Anger (*11 items, Cronbach's α=0.77), *Utility of Personal Aggression (*16 items, Cronbach's α=0.77), *Success in Conflict (*7 items, Cronbach's α=0.81), and *History of Fighting* (5 items, Cronbach's α=0.73). Typical items for each scale were as follows: ''It is harder to get me angry than other people'' (*Proneness to Anger*, reverse coded), ''If I don't respond to provocations and do something to make the wrong-doers pay, they'll just do more to hurt me in the future'' (*Utility of Personal Aggression*), ''When there's a dispute, I usually get my way'' (*Success in Conflict*), ''I have physically intimidated someone who had it coming'' (*History of Fighting*).

Vengefulness (4 items, 5-point Likert scale, 1 = "strongly disagree" to 5 = "strongly agree", Cronbach's α=0.72) was measured using the Revenge subscale of the Transgression-Related Interpersonal Motivations (TRIM) inventory (McCullough et al., 1998). A typical item was: "When someone angers me or hurts my feelings, I usually find a way to make this person regret it."

Extraversion (48 items, 5-point Likert scale, 1 = "strongly disagree" to 5 = "strongly agree", Cronbach's α=0.91) was measured using the NEO-PI-R Extraversion scale (Costa & McCrae, 1992), including the six facets Warmth (Cronbach's α=0.81), Gregariousness (Cronbach's α=0.81), Assertiveness (Cronbach's α=0.81), Activity (Cronbach's α=0.72), Excitement

Seeking (Cronbach's α=0.65), and Positive Emotion (Cronbach's α=0.81). All facets consisted of 8 items. Typical items for each facet were: "I really enjoy talking to people." (Warmth), "I like to have a lot of people around me." (Gregariousness), "I am dominant, forceful, and assertive." (Assertiveness), "I often feel as if I'm bursting with energy." (Activity), "I like to be where the action is." (Excitement Seeking), "I am a cheerful, high-spirited person." (Positive Emotion).

Dominance (11 items, 5-point Likert scale, 1 = "strongly disagree" to 5 = "strongly agree", Cronbach's α=0.82) was measured using the Interpersonal Personality Item Pool version of the dominance subscale from the CPI narcissism scale (Goldberg et al., 2006; Gough, 1956). A typical item was: "I impose my will on others."

Shyness (5 items, 5-point Likert scale, agreement format from 1 = "not at all" to 5 = "completely", Cronbach's α=0.84) was measured using the five item Shyness Scale (Asendorpf & Wilpers, 1998). A typical item was: "I feel inhibited when I am with other people".

Narcissism was measured using the Narcissistic Admiration and Rivalry Questionnaire (NARQ; Back et al., 2013), comprising the two dimensions Admiration (9 items, 5-point Likert scale, 1 = "not agree at all" to 6 = "agree completely", Cronbach's α=0.77) and Rivalry (9 items, 5-point Likert scale, Cronbach's α=0.73). Typical items were: "I will someday be famous" (Admiration) and "I react annoyed if another person steals the show from me." (Rivalry).

Sociosexual Orientation (9 items, 5-point response scales, Cronbach's α=0.88) was measured using the revised Sociosexual Orientation Inventory (SOI-R; Penke & Asendorpf, 2008), comprising the three facets Attitude (3 items, Cronbach's α=0.88), Behavior (3 items, Cronbach's α=0.85), and Desire (3 items, Cronbach's α=0.83). Typical items were: "With how many different partners have you had sex within the past 12 months?" (Behavior), "Sex without

love is OK." (Attitude), "In everyday life, how often do you have spontaneous fantasies about having sex with someone you have just met?" (Desire).

2.1.2.2 Measures of Physical Attractiveness

Self-perceived physical attractiveness was measured as a 3-item aggregate (Cronbach's α=0.82) based on questions similar to those used in previous studies (cf. Lukaszewski & Roney, 2011; Price et al., 2012; Sell et al., 2009): "I am more attractive than __% of others of my sex"; "On a scale from 1 to 10, how physically attractive are you?", "Compared to others I'm a very attractive person" (on a Likert scale from 1="strongly disagree" to 5="strongly agree").

Since a tighter link in a attractiveness-personality-relationship could be expected for measures that are more directly related to mating success, we computed *self-perceived mating success* as a 3-item aggregate (Cronbach's α=0.82) based on 7-point Likert scale items (1= "not at all" to 7 = "very") from the self-perceived Mate Value Scale MVS (Landolt, Lalumière, & Quinsey, 1995): "Members of the opposite sex are attracted to me.", "Members of the opposite sex notice me.", "I do not receive many compliments from members of the opposite sex.".

Other-rated physical attractiveness was judged by eight male and eight female raters, mostly undergraduate students of the University of Edinburgh (age *M* = 22.1 years, *SD* = 1.1). Raters saw rotating animations ('beauty turns') of the body scans, with heads removed to focus attention on body attractiveness and uniform grey color (so free of skin color cues; similar to Smith et al., 2007). Relative height differences of the stimulus subjects were maintained in the presentations. Beauty turns were displayed on a computer screen using the Eprime software, and evaluated them individually on a (7-point Likert scale from 1= "not attractive" to 7 = "very attractive", interrater agreement Cronbach's α=0.92).

Residual self-perceived physical attractiveness was computed as the residuals from a regression of self-perceived on other-rated physical attractiveness. This additional index of self-perceived physical attractiveness is controlled for the consensual, arguably more objective outside perception of one's body attractiveness, thereby getting closer the subjective component of self-perceived attractiveness.

2.1.2.3 Formidability Measures

Strength was assessed as hand grip strength for both hands and upper body strength, all measured with a Saehan SH500 dynamometer following the procedure described in the Appendix of Sell et al. (2009). Handedness was assessed using the Edinburgh Handedness Inventory (Oldfield, 1971). Each strength measure was taken three times and the maximum performance was used for further analyses. An overall strength variable was computed as the mean of z-standardized measures of dominant hand grip and upper body strength (which were highly correlated; $r=0.83$, $p<0.001$).

Lung function was measured with a spirometer (MicroPlus, CareFusion). Lung forced vital capacity (FVC) and forced expiratory volume per 1 second (FEV) were assessed three times, and the maximum performance values for each measure were z-standardized and averaged into an overall lung function variable. We included lung function as an additional measure of formidability indicating a person's aerobic fitness and therefore potentially physical competitiveness.

Body masculinity was calculated, following Price et al. (2012), as the regression score of the first unrotated principal component extracted across sexes from the following sexually dimorphic variables (effect sizes of group differences between sexes are given in parentheses): height (d=2.17), dominant arm flexed biceps circumference (d=1.84), and indices calculated from automatic measurements taken by the TC^2 NX-16 body scanner software (average of measures from three scans), including shoulder breadth (d=1.69), forearm circumference

(*d*=1.73), chest circumference (*d*=1.65), waist-to-hip ratio (*d*=1.34), bust-to-underbust ratio (*d*=-2.80), and leg-to-body ratio (*d*=-1.07). The PCA explained 63.48% of the variables` variance.

BMI-controlled body masculinity was computed as the residuals from a regression of *Body masculinity* on BMI. This measure was intended as a robustness check to separate the index of *Body masculinity* from the influence of the BMI, approximatively controlling for differences in body measurements that are not due to sexual dimorphisms in body shape or muscularity but due to idiosyncratic differences in body fat (although BMI is correlated to muscularity as well; Heymsfield, Scherzer, Pietrobelli, Lewis, & Grunfeld, 2009).

Facial masculinity was computed sensu Penton-Voak et al. (2001) from 14 landmarks placed on the 3D facial photographs using Morphanalyser (Tiddeman, Duffy, & Rabey, 2000). For the facial masculinity index, z-standardized measures of face width to lower face height, eye size, and cheekbone prominence were subtracted from the z-standardized ratio of lower face to face height.

2.1.3 Statistical Analyses

We z-standardized all variables (except for the dichotomous control variable ethnicity, white vs. non-white) prior to analysis and graphically inspected the data for normality and outliers. From the original sample (119 males, 124 females) we excluded 3 influential cases (2 males, 1 female) based on outliers in body appearance (values beyond 1.5 * interquartile range above or below the third or first quartile, respectively) that showed high leverage in regression diagnostics (graphical inspection of the bivariate distribution between residuals and leverage of data points as well as distribution of Cook's distances; see figures S1.1 and S1.2 in the supplementary material). For the resulting sample (117 males, 123 females) we computed zero-order Pearson correlation coefficients between all variables and their bootstrapped 95% confidence intervals. To account for multiple testing we also adjusted all p-values of the

correlations in the result section for false discovery rate (FDR; Benjamini & Hochberg, 1995), i.e. controlled for the expected proportion of falsely rejected hypotheses among all rejected hypotheses, using the R package "psych" (Revelle, 2016). As a robustness analysis we additionally computed partial Pearson correlation coefficients controlling for age (since age was an important influence on the results of Price et al., 2012) and ethnicity of participants (separating between white and non-white). Then we computed the congruency coefficient Rc (Abdi, 2010) between the correlation matrices with and without control variables, where a value of Rc=1 would indicate complete congruency between both matrices, and checked whether changes in statistical significance (in terms of a threshold $p<0.05$) occurred.

2.2 Results

First, we assessed how self-perceived attractiveness and self-perceived mating success were related to each other, as well as to objective body measures and other-rated attractiveness. Whereas self-rated attractiveness and mating success were strongly correlated in men ($r=0.69$, $p<0.001$) and women ($r=0.61$, $p<0.001$), self-perceived attractiveness and mating success correlated with objective indicators of physical appearance only among men (other-rated attractiveness, strength, and height; see Table 2).

Self-perceived attractiveness was, as expected, positively correlated with the personality measures Extraversion, Dominance (only in men), and Narcissistic Admiration, as well as negatively with Shyness. Furthermore it showed a positive relationship with Sociosexual Behavior and also partly with Anger or Aggressiveness in both sexes. Other-rated attractiveness was overall not correlated with any personality measures (except for Success in Conflict in women; see Table 3). For men we even found a negative correlation between other-rated attractiveness and Proneness to Anger, which was contrary to the expected direction (Table 3).

Extraversion was completely unrelated to objective measures of formidability such as physical masculinity and strength on the domain level. However, in the male subsample, the

facets Assertiveness and Activity positively correlated with body masculinity and physical strength, as well as height with Gregariousness (see Table 3). With one exception (lung function positively related to Excitement Seeking) we did not find any such correlations for women. Sociosexual Orientation and Anger Proneness were also largely unrelated to objective body measures in both sexes, however with a few exceptions: Positive correlations occurred between body masculinity, upper body strength and Utility of Personal Aggression, between physical strength and Narcissistic Admiration, and between height and short-term mating behavior in men. In addition, we found an unpredicted negative correlation between lung function and Narcissistic Rivalry in men.

The similarity of partial correlations controlling for age and ethnicity with zero-order correlations was highly significant (men: Rc=0.999; 95% CI= [0.999;0.999]; p<0.001; women: Rc=0.998; 95% CI= [0.996;0.999]; p<0.001). Changes of statistical significance between partial- and zero-order correlations occurred in six cases for the male and five cases for the female sample. All changes, except the correlation among BMI-controlled Body Masculinity and History of Fighting in men, reflected correlations either among two personality variables or two formidability variables and did thus not affect the interpretation of relationships subject to facultative calibration. The results of correlations based on single anthropometric measures of the body and the face of individuals (which were combined into measures of body masculinity and facial masculinity in Table 3) can be found in the supplementary material (S3).

Table 2

Pearson correlation coefficients between self-rated and objective formidability measures

Women	SP attractiveness	SP mating success

Other-rated attractiveness	0.14 [-0.09;0.34]	0.16 [-0.03;0.35]
Body masculinity	-0.15 [-0.30;0.03]	-0.16 [-0.28;-0.02]
BMI-controlled body masculinity	0.03 [-0.12;0.22]	-0.03 [-0.16;0.12]
Upper body size	-0.13 [-0.29;0.06]	-0.14 [-0.26;0.03]
Facial masculinity	-0.16 [-0.34;0.02]	-0.12 [-0.29;0.05]
Strength	0.06 [-0.15;0.21]	0.08 [-0.09;0.23]
Lung function	-0.01 [-0.16;0.13]	-0.02 [-0.18;0.13]
Height	-0.02 [-0.16;0.14]	0.01 [-0.12;0.19]
BMI	*-0.23 [-0.43;-0.01]*	-0.19 [-0.34;0.00]
Men	SP attractiveness	SP mating success
Other-rated attractiveness	***0.33 [0.16;0.52]***	***0.23 [0.07;0.41]***
Body masculinity	0.14 [-0.02;0.34]	0.02 [-0.14;0.24]
BMI-controlled body masculinity	***0.30 [0.17;0.44]***	*0.17 [0.03;0.35]*
Upper body size	0.13 [-0.06;0.36]	-0.02 [-0.2;0.20]
Facial masculinity	0.05 [-0.15;0.19]	-0.03 [-0.18;0.12]
Strength	***0.26 [0.10;0.41]***	0.12 [-0.05;0.29]
Lung function	0.10 [-0.02;0.27]	0.10 [-0.05;0.25]
Height	***0.26 [0.12;0.44]***	***0.25 [0.08;0.40]***
BMI	-0.06 [-0.28;0.16]	-0.11 [-0.29;0.12]

Significant results in terms of FDR-adjusted p-values <.05 are displayed in bold, confidence intervals not containing the value 0 in italics.

Table 3

Pearson correlation coefficients between formidability and personality measures from the Study 1, female subsample

	Anger		**Extraversion**										**Additional Personality**				**Sociosexual Orientation**			
	UA	AP	SC	FH	V	E	E1	E2	E3	E4	E5	E6	D	S	NR	NA	SOI-R	S-A	S-B	S-D
Self-perceived attractiveness	-0.06 [-0.28; 0.11]	-0.04 [-0.26; 0.16]	***0.30*** ***[0.13; 0.47]***	0.17 [-0.02; 0.32]	0.08 [-0.10; 0.25]	***0.35*** ***[0.23; 0.50]***	***0.24*** ***[0.06; 0.42]***	***0.26*** ***[0.08; 0.42]***	0.17 [0.00; 0.34]	*0.19* *[0.04; 0.36]*	***0.25*** ***[0.11; 0.38]***	***0.30*** ***[0.16; 0.45]***	0.13 [-0.11; 0.33]	***-0.35*** ***[-0.51; -0.21]***	0.06 [-0.15; 0.24]	***0.58*** ***[0.46; 0.69]***	*0.17* *[0.03; 0.35]*	0.07 [-0.09; 0.26]	***0.25*** ***[0.11; 0.38]***	0.10 [-0.0 0.29
Residual self-perceived attractiveness	-0.07 [-0.28; 0.09]	-0.06 [-0.29; 0.15]	*0.21* *[0.03; 0.42]*	0.17 [-0.01; 0.33]	0.05 [-0.15; 0.23]	***0.31*** ***[0.15; 0.47]***	*0.21* *[0.00; 0.41]*	***0.25*** ***[0.10; 0.41]***	0.14 [-0.05; 0.36]	0.12 [-0.04; 0.29]	***0.29*** ***[0.15; 0.44]***	***0.26*** ***[0.09; 0.44]***	0.10 [-0.15; 0.32]	***-0.33*** ***[-0.5; -0.18]***	0.05 [-0.15; 0.21]	***0.53*** ***[0.41; 0.67]***	*0.18* *[0.02; 0.36]*	0.05 [-0.13; 0.26]	***0.28*** ***[0.14; 0.41]***	0.13 [0.0 30]
Self-perceived mating success	-0.05 [-0.23; 0.10]	-0.08 [-0.28; 0.14]	***0.38*** ***[0.21; 0.53]***	0.08[-0.10; 0.24]	0.03 [-0.16; 0.20]	***0.38*** ***[0.23; 0.51]***	***0.28*** ***[0.07; 0.49]***	***0.30*** ***[0.13; 0.44]***	0.19 [-0.05; 0.34]	*0.20* *[0.04; 0.36]*	***0.20*** ***[0.04; 0.35]***	***0.39*** ***[0.26; 0.54]***	0.10 [-0.17; 0.25]	***-0.40*** ***[-0.52; -0.24]***	-0.11 [-0.29; 0.05]	***0.50*** ***[0.37; 0.61]***	0.14 [-0.02; 0.32]	0.10 [-0.07; 0.27]	***0.29*** ***[0.13; 0.43]***	-0.0 [-0.2 0.16
Other-rated attractiveness	0.11 [-0.04; 0.22]	0.10 [-0.08; 0.26]	***0.21*** ***[0.05; 0.35]***	-0.04 [-0.23; 0.11]	*0.17* *[0.01; 0.31]*	0.01 [-0.19; 0.19]	-0.05 [-0.20; 0.10]	-0.08 [-0.27; 0.13]	0.02 [-0.16; 0.17]	0.14 [-0.04; 0.33]	0.00 [-0.19; 0.22]	0.02 [-0.15; 0.19]	0.13 [-0.03; 0.28]	0.06 [-0.09; 0.24]	0.09 [-0.10; 0.25]	0.08 [-0.08; 0.22]	-0.10 [-0.25; 0.08]	-0.05 [-0.21; 0.10]	-0.14 [-0.31; 0.05]	-0.0 [-0.2 0.13
Body masculinity	-0.09 [-0.21; 0.04]	-0.08 [-0.22; 0.09]	-0.05 [-0.21; 0.10]	0.00 [-0.17; 0.15]	-0.13 [-0.29; 0.05]	0.09 [-0.05; 0.26]	0.04 [-0.08; 0.18]	0.01 [-0.14; 0.19]	0.13 [-0.03; 0.29]	0.03 [-0.12; 0.21]	0.06 [-0.10; 0.25]	0.09 [-0.04; 0.26]	-0.01 [-0.16; 0.13]	-0.06 [-0.20; 0.08]	-0.17 [-0.33; 0.03]	-0.10 [-0.27; 0.11]	0.11 [-0.02; 0.27]	0.08 [-0.05; 0.21]	0.11 [-0.05; 0.28]	0.10 [-0.0 0.24
BMI-controlled body masculinity	-0.09 [-0.20; 0.05]	-0.02 [-0.22, 0.15]	0.03 [-0.14; 0.16]	-0.01 [-0.17; 0.15]	-0.11 [-0.23; 0.04]	0.10 [-0.03; 0.27]	-0.04 [-0.15; 0.13]	-0.02 [-0.18; 0.13]	0.10 [-0.07; 0.26]	*0.14* *[0.01; 0.31]*	0.15 [-0.03; 0.31]	0.08 [-0.10; 0.26]	0.10 [-0.06; 0.22]	-0.01 [-0.16; 0.12]	-0.09 [-0.22, 0.06]	0.02 [-0.13; 0.19]	0.06 [-0.08; 0.24]	0.06 [-0.10; 0.27]	0.03 [-0.08; 0.18]	0.06 [-0.0 0.21
Upper body size	-0.08 [-0.20; 0.06]	-0.12 [-0.29; 0.03]	-0.04 [-0.22; 0.12]	0.02 [-0.14; 0.18]	-0.09 [-0.26; 0.05]	0.03 [-0.10; 0.21]	0.02 [-0.10; 0.17]	0.00 [-0.12; 0.17]	0.06 [-0.09; 0.24]	-0.04 [-0.19; 0.15]	0.01 [-0.16; 0.19]	0.07 [-0.10; 0.25]	-0.07 [-0.21; 0.10]	-0.04 [-0.17; 0.09]	-0.12 [-0.28; 0.03]	-0.11 [-0.29; 0.13]	*0.17* *[0.01; 0.32]*	0.12 [0.00; 0.28]	*0.20* *[0.02; 0.35]*	0.10 [-0.1 0.25

Facial masculinity	*0.23 [0.03; 0.39]*	0.02 [-0.17; 0.2]	0.07 [-0.16; 0.26]	0.01 [-0.19; 0.15]	-0.04 [-0.22; 0.18]	0.06 [-0.17; 0.20]	-0.10 [-0.26; 0.07]	0.02 [-0.14; 0.19]	0.01 [-0.15; 0.14]	0.20 [-0.04; 0.36]	0.12 [-0.10; 0.25]	-0.02 [-0.21; 0.13]	0.09 [-0.12; 0.27]	-0.15 [-0.30; 0.02]	*0.13 [0.01; 0.29]*	0.02 [-0.17; 0.16]	-0.02 [-0.16; 0.12]	0.00 [-0.16; 0.17]	-0.13 [-0.26; 0.02]	0.09 [-0.0 0.25
Strength	*-0.17 [-0.34; -0.02]*	-0.17 [-0.34; 0.06]	0.02 [-0.12; 0.19]	-0.10 [-0.33; 0.10]	*-0.18 [-0.33; -0.02]*	0.13 [-0.03; 0.28]	0.06 [-0.06; 0.23]	0.00 [-0.16; 0.20]	-0.03 [-0.18; 0.11]	*0.19 [0.07; 0.35]*	0.14 [-0.05; 0.30]	*0.18 [0.01; 0.33]*	-0.07 [-0.18; 0.10]	-0.05 [-0.20; 0.12]	*-0.14 [-0.30; -0.01]*	-0.01 [-0.16; 0.16]	0.00 [-0.14; 0.13]	0.02 [-0.15; 0.22]	0.12 [-0.03; 0.26]	*-0.1 [-0.2 -0.0*
Lung function	-0.16 [-0.32; 0.02]	-0.08 [-0.25; 0.11]	-0.03 [-0.20; 0.13]	-0.13 [-0.34; 0.04]	-0.08 [-0.3; 0.07]	0.10 [-0.08; 0.26]	-0.04 [-0.20; 0.17]	0.04 [-0.11; 0.20]	-0.04 [-0.19; 0.13]	0.13 [-0.07; 0.31]	***0.25 [0.13; 0.39]***	0.06 [-0.17; 0.20]	0.10 [-0.09; 0.33]	0.02 [-0.13; 0.18]	0.05 [-0.11; 0.23]	-0.02 [-0.14; 0.14]	*0.16 [0.01; 0.35]*	0.09 [-0.08; 0.28]	0.16 [-0.01; 0.32]	0.14 [-0.0 0.33
Height	-0.21 [-0.36; -0.02]	-0.14 [-0.34; 0.06]	-0.06 [-0.17; 0.15]	*-0.16 [-0.34; -0.01]*	-0.07 [-0.23; 0.10]	0.00 [-0.14; 0.17]	-0.07 [-0.21; 0.13]	-0.05 [-0.20; 0.12]	0.03 [-0.11; 0.25]	0.01 [-0.12; 0.16]	-0.01 [-0.16; 0.17]	0.06 [-0.08; 0.21]	-0.01 [-0.20; 0.2]	0.09 [-0.10; 0.21]	-0.04 [-0.19; 0.12]	-0.07 [-0.19; 0.13]	0.02 [-0.14; 0.16]	-0.02 [-0.14; 0.14]	0.01 [-0.15; 0.18]	0.06 [-0.1 0.20
BMI	-0.03 [-0.17; 0.10]	-0.10 [-0.27; 0.07]	-0.10 [-0.27; 0.08]	0.01 [-0.13; 0.18]	-0.09 [-0.26; 0.07]	0.02 [-0.14; 0.19]	0.08 [-0.03; 0.21]	0.03 [-0.13; 0.20]	0.08 [-0.10; 0.22]	-0.10 [-0.27; 0.08]	-0.07 [-0.28; 0.10]	0.04 [-0.13; 0.25]	-0.12 [-0.26; 0.04]	-0.06 [-0.21; 0.05]	*-0.15 [-0.26; -0.01]*	-0.16 [-0.36; 0.07]	0.09 [-0.07; 0.26]	0.04 [-0.11; 0.20]	0.12 [-0.05; 0.28]	0.06 [-0.1 0.22

Significant results in terms of FDR-adjusted p-values <.05 are displayed in bold, confidence intervals not containing the value 0 in italics. UA= Utility of Personal Aggression, AP= Proneness to Anger, SC= Success in Conflict, FH= History of Fighting, V= Vengefulness, E= Extraversion, E1= Warmth, E2= Gregariousness, E3= Assertiveness, E4= Activity, E5= Excitement-Seeking, E6= Positive Emotions, D= Dominance, S= Shyness, NR= Narcissistic Rivalry, NA= Narcissistic Admiration, SOI-R= revised Sociosexual Orientation Inventory, S-A= Sociosexual Attitudes, S-B= Sociosexual Behavior, S-D= Sociosexual Desire

Table 4

Pearson correlation coefficients between formidability and personality measures from the Study 1, male subsample

	Anger					**Extraversion**							**Additional Personality**				**Sociosexual Orientation**			
	UA	AP	SC	FH	V	E	E1	E2	E3	E4	E5	E6	D	S	NR	NA	SOI-R	S-A	S-B	S-D
Self-perceived attractiveness	***0.36 [0.23; 0.50]***	*0.20 [0.06; 0.38]*	***0.58 [0.45; 0.71]***	***0.39 [0.24; 0.53]***	***0.24 [0.09; 0.38]***	***0.46 [0.36; 0.56]***	*0.17 [0.05; 0.30]*	***0.38 [0.23; 0.55]***	***0.46 [0.28; 0.59]***	***0.50 [0.39; 0.60]***	***0.31 [0.14; 0.46]***	***0.30 [0.14; 0.46]***	***0.24 [0.12; 0.38]***	***-0.45 [-0.57; -0.33]***	-0.03 [-0.17; 0.14]	***0.52 [0.39; 0.61]***	*0.18 [0.05; 0.33]*	0.16 [0.00; 0.35]	***0.23 [0.11; 0.38]***	0.02 [-0.13 0.21]
Residual self-perceived attractiveness	***0.40 [0.29; 0.53]***	***0.27[0.13; 0.43]***	***0.56 [0.43; 0.69]***	***0.45 [0.32; 0.59]***	***0.30 [0.14; 0.44]***	***0.44 [0.32; 0.55]***	*0.19 [0.07; 0.34]*	***0.38 [0.23; 0.57]***	***0.43 [0.28; 0.54]***	***0.46 [0.31; 0.57]***	***0.28 [0.11; 0.47]***	***0.27 [0.11; 0.46]***	***0.28 [0.14; 0.42]***	***-0.42 [-0.56 ;-0.31]***	0.01 [-0.18; 0.21]	***0.48 [0.35; 0.58]***	*0.18 [0.03; 0.34]*	0.14 [-0.03; 0.35]	***0.22 [0.07; 0.41]***	0.05 [-0.08 0.23]
Self-perceived mating success	*0.18 [0.03; 0.37]*	0.04 [-0.14; 0.27]	***0.43 [0.24; 0.59]***	***0.27 [0.10; 0.45]***	*0.18 [0.04; 0.34]*	***0.40 [0.27; 0.57]***	***0.22 [0.07; 0.38]***	***0.37 [0.22; 0.54]***	***0.27 [0.07; 0.45]***	***0.36 [0.19; 0.56]***	***0.32 [0.14; 0.51]***	***0.29 [0.07; 0.46]***	*0.18 [0.03; 0.35]*	***-0.38 [-0.56; -0.23]***	-0.08 [-0.25; 0.14]	***0.37 [0.23; 0.52]***	***0.32 [0.20; 0.47]***	***0.33 [0.18; 0.50]***	***0.34 [0.20; 0.48]***	0.10 [-0.02 0.29]
Other-rated attractiveness	-0.01 [-0.16; 0.17]	***-0.22 [-0.37; -0.07]***	0.12 [-0.06; 0.3]	-0.07 [-0.22; 0.08]	-0.19 [-0.39; 0.00]	0.11 [-0.06; 0.27]	-0.08 [-0.24; 0.07]	0.07 [-0.10; 0.24]	0.09 [-0.11; 0.26]	*0.17 [0.02; 0.36]*	0.12 [-0.05; 0.28]	0.15 [0.00; 0.32]	-0.09 [-0.25; 0.07]	-0.07 [-0.24; 0.09]	-0.18 [-0.34; 0.02]	0.14 [-0.01; 0.31]	0.09 [-0.11; 0.23]	0.11 [-0.07; 0.27]	0.11 [-0.09; 0.24]	-0.01 [-0.18 0.17]
Body masculinity	***0.27 [0.13; 0.41]***	0.04 [-0.15; 0.19]	0.15 [-0.05; 0.31]	0.19 [-0.01; 0.31]	0.10 [-0.07; 0.26]	0.14 [-0.01; 0.31]	0.08 [-0.06; 0.24]	-0.01 [-0.12; 0.18]	***0.25 [0.04; 0.42]***	***0.24 [0.07; 0.42]***	0.05 [-0.08; 0.22]	0.03 [-0.16; 0.21]	0.13 [-0.08; 0.28]	0.00 [-0.20; 0.15]	-0.01 [-0.20; 0.15]	0.11 [-0.15; 0.31]	0.10 [-0.06; 0.24]	0.02 [-0.14; 0.2]	0.12 [-0.01; 0.27]	0.09 [-0.06 0.22]
BMI-controlled body masculinity	***0.23 [0.05; 0.37]***	-0.03 [-0.19, 0.14]	0.17 [-0.02; 0.41]	*0.18 [0.04; 0.31]*	0.05 [-0.10; 0.21]	*0.16 [0.02; 0.32]*	0.01 [-0.18; 0.20]	0.05 [-0.10; 0.21]	***024 [0.12; 0.38]***	***0.26 [0.09; 0.41]***	0.10 [-0.03; 0.29]	0.08 [-0.09; 0.29]	0.04 [-0.16; 0.20]	-0.04 [-0.22; 0.11]	-0.15 [-0.32; 0.05]	0.14 [-0.03; 0.28]	0.14 [-0.02; 0.32]	0.10 [-0.06; 0.24]	*0.20 [0.04; 0.36]*	0.04 [-0.15 0.24]
Upper body size	***0.23 [0.08; 0.36]***	0.04 [-0.16; 0.22]	0.15 [-0.05; 0.30]	0.13 [-0.08; 0.30]	0.10 [-0.04; 0.24]	0.14 [-0.02; 0.33]	0.13 [-0.01; 0.31]	-0.02 [-0.16; 0.20]	***0.24 [0.04; 0.42]***	*0.20 [0.01; 0.39]*	0.05 [-0.08; 0.22]	0.05[-0.13; 0.22]	0.13 [-0.06; 0.27]	-0.01 [-0.19; 0.14]	0.00 [-0.15; 0.14]	0.09 [-0.16; 0.27]	0.00 [-0.14; 0.15]	-0.08 [-0.21; 0.09]	0.06 [-0.08; 0.2]	0.02 [-0.17 0.19]
Facial masculinity	0.06 [-0.08; 0.22]	0.05 [-0.12; 0.26]	0.04 [-0.14; 0.28]	0.10 [-0.07; 0.29]	-0.06 [-0.22; 0.10]	0.06 [-0.18; 0.33]	0.11 [-0.11; 0.36]	-0.01 [-0.21; 0.18]	0.04 [-0.12; 0.23]	0.01 [-0.19; 0.22]	0.12 [-0.04; 0.37]	0.02 [-0.20; 0.24]	-0.01 [-0.19; 0.20]	-0.05 [-0.26; 0.15]	0.07 [-0.13; 0.24]	0.09 [-0.07; 0.27]	-0.02 [-0.22; 0.18]	-0.12 [-0.27; 0.06]	0.10 [-0.08; 0.33]	-0.03 [-0.22 0.14]
Strength	0.09 [-0.11; 0.26]	-0.12 [-0.37; 0.06]	0.20 [-0.04; 0.39]	0.09 [-0.08; 0.29]	-0.08 [-0.24; 0.10]	*0.18 [0.02; 0.34]*	0.08 [-0.07; 0.23]	0.02 [-0.11; 0.20]	***0.28 [0.13; 0.44]***	***0.26 [0.08; 0.44]***	0.08 [-0.06; 0.26]	0.12 [-0.04; 0.31]	0.10 [-0.05; 0.25]	-0.12 [-0.30; 0.03]	0.04 [-0.11; 0.22]	***0.28 [0.15; 0.43]***	0.01 [-0.17; 0.18]	0.01 [-0.15; 0.21]	0.04 [-0.13; 0.21]	-0.03 [-0.17 0.11]
Lung function	0.12	-0.09	0.05	0.13	-0.08	0.06	-0.03	0.04	0.08	0.13	0.02	0.02	-0.01	0.02	***-0.26***	-0.08	0.05	0.01	0.07	0.03

	[-0.02; 0.27]	[-0.25; 0.07]	[-0.11; 0.20]	[-0.05; 0.27]	[-0.23; 0.09]	[-0.08; 0.21]	[-0.18; 0.14]	[-0.07; 0.17]	[-0.05; 0.22]	[-0.02; 0.28]	[-0.16; 0.17]	[-0.11; 0.15]	[-0.18; 0.11]	[-0.13; 0.15]	***[-0.38; -0.08]***	[-0.2; 0.08]	[-0.13; 0.25]	[-0.16; 0.18]	[-0.10; 0.27]	[-0.14 0.25]
Height	0.06	-0.11	0.10	*0.16*	0.04	*0.20*	0.07	***0.25***	0.10	*0.18*	0.16	0.13	-0.04	-0.17	-0.16	0.03	*0.15*	0.04	***0.25***	0.07
	[-0.12; 0.26]	[-0.27; 0.05]	[-0.08; 0.30]	*[0.04; 0.31]*	[-0.15; 0.22]	*[0.05; 0.35]*	[-0.03; 0.26]	***[0.09; 0.41]***	[-0.05; 0.28]	*[0.02; 0.38]*	[-0.02; 0.34]	[-0.02; 0.31]	[-0.19; 0.16]	[-0.33; -0.02]	[-0.30; 0.05]	[-0.10; 0.25]	*[0.02; 0.29]*	[-0.12; 0.19]	***[0.13; 0.38]***	[-0.07 0.23]
BMI	*0.19*	0.08	0.07	0.12	0.10	0.06	0.10	-0.05	0.15	0.13	-0.01	-0.03	0.16	0.03	0.11	0.04	0.02	-0.05	0.00	0.10
	[0.06; 0.33]	[-0.08; 0.24]	[-0.12; 0.24]	[-0.10; 0.29]	[-0.04; 0.26]	[-0.08; 0.24]	[-0.04; 0.26]	[-0.17; 0.2]	[-0.03; 0.34]	[-0.03; 0.3]	[-0.13; 0.17]	[-0.21; 0.15]	[0.00; 0.28]	[-0.17; 0.19]	[-0.04; 0.27]	[-0.22; 0.23]	[-0.14; 0.18]	[-0.20; 0.11]	[-0.14; 0.17]	[-0.08 0.27]

Significant results in terms of FDR-adjusted p-values <.05 are displayed in bold, confidence intervals not containing the value 0 in italics. UA= Utility of Personal Aggression, AP= Proneness to Anger, SC= Success in Conflict, FH= History of Fighting, V= Vengefulness, E= Extraversion, E1= Warmth, E2= Gregariousness, E3= Assertiveness, E4= Activity, E5= Excitement-Seeking, E6= Positive Emotions, D= Dominance, S= Shyness, NR= Narcissistic Rivalry, NA= Narcissistic Admiration, SOI-R= revised Sociosexual Orientation Inventory, S-A= Sociosexual Attitudes, S-B= Sociosexual Behavior, S-D= Sociosexual Desire

3. Study 2

3.1 Methods

Note that since Study 2 was overall very similar to Study 1, we do only describe those methodological aspects of Study 2 in detail that differed from Study 1.

3.1.1 Participants and Procedure

165 men (age: *M*=24.3, *SD*=3.2 years) from the local community of Göttingen (Germany), mostly university undergraduates, were preselected for being heterosexual (to satisfy demands of another study based on these data; heterosexual orientation on a 7-point Kinsey scale; Kinsey et al., 1948) and recruited via an online database, in exchange for monetary compensation. The participants were asked to fill out a battery of questionnaires (described below) and subsequently were body-scanned using a Vitus^{smart}XXL 3D bodyscanner (Human Solutions GmbH, Kaiserslautern, Germany) and a 3dMD face scanner. All participants were scanned three times, while only wearing standardized tight underwear in the size of their choice (small to extra-extra-large). Participants were instructed to stand still in a standardized posture (standing upright with legs hip-widely apart, arms stretched out and held slightly away from the body, hands making a fist with thumbs showing forward, and head positioned in accordance with the Frankfort Horizontal) and breathe normally during the scanning process (ca. 10 secs. each). Additionally, body height (in cm) was measured twice using a statiometer. The participants stood upright without wearing shoes; the two values were averaged. Weight (in kg) was measured as part of each body scanning process with the integrated scale SECA 635 (SECA, Hamburg, Germany); the three values were averaged.

3.1.2 Measures

3.1.2.1 Personality Measures

Extraversion (8 items, Cronbach's α=0.87) was measured with the German version of the Big Five Inventory (BFI; Lang et al., 2001; 5-point Likert scale from 1 = "strongly disagree" to 5 = "strongly agree"). A typical items was: "I see myself as someone who is talkative".

In an exploratory manner we also included the remaining Big Five personality domains *Neuroticism* (7 items, Cronbach's α=0.81), *Openness to Experience* (10 items, Cronbach's α=0.78), *Agreeableness* (8 items, Cronbach's α=0.73), and *Conscientiousness* (9 items, Cronbach's α=0.84), which were measured with the German version of the Big Five Inventory as well (5-point Likert scale from 1 = "strongly disagree" to 5 = "strongly agree"). Typical items were as follows: "I see myself as someone who: is depressed, blue" (Neuroticism); "is original, comes up with new ideas" (Openness to Experience); "has a forgiving nature" (Agreeableness); "does a thorough job" (Conscientiousness).

Dominance was measured using five out of the eight items of the Interpersonal Adjective List (Jacobs & Scholl, 2005) that assess the PA facet (dominance-assured) of the Interpersonal Circumplex (Wiggins et al., 1988; 8-point Likert scale, 1= "extremely inaccurate" to 8 = "extremely accurate", Cronbach's α=0.74). A typical item was: "I am forceful."

Shyness was measured using five out of the eight items of the Interpersonal Adjective List (Jacobs & Scholl, 2005) that assess the HI facet (unassured-submissive; 8-point Likert scale, 1= "extremely inaccurate" to 8 = "extremely accurate", Cronbach's α=0.75). A typical item was: "I am timid."

Narcissism was measured using the short version of the Narcissistic Admiration and Rivalry Questionnaire (NARQ; Back et al., 2013), comprising the two dimensions Admiration (3 items, 6-point Likert scale, Cronbach's α=0.71) and Rivalry (3 items, 6-point Likert scale, Cronbach's α=0.56). For typical items see methods of Study 1.

Sociosexual Orientation was measured using the revised Sociosexual Orientation Inventory (SOI-R; Penke & Asendorpf, 2008) on a 9-point scale comprising the three facets Attitude (3 items, Cronbach's α=0.71), Behavior (3 items, Cronbach's α=0.86), and Desire (3 items, Cronbach's α=0.85). For typical items see methods of Study 1.

Anger Proneness and *Vengefulness* were not included in Study 2.

3.1.2.2 Measures of Physical Attractiveness

Self-perceived physical attractiveness was based on an aggregate of three items (attractive, sexy, appealing) rated on a 5-point Likert scale from 1 = "very unattractive" to 5 = "very attractive" (Cronbach's α=0.85).

Other-rated body attractiveness was rated by 31 females (the stimuli sample was divided into two sets in order to avoid rater fatigue effects, we had 15 and 16 raters for sets 1 and 2, respectively) on an 11-point scale (from -5 = "very unattractive" to +5 = "very attractive") in response to the question "How attractive is this man?" (set 1: α= 0.93, set 2: α= 0.94). Animated videos of a body scan turning around its vertical axis (similar to Smith et al., 2007) were created using AnthroScan software ("beauty turns", duration: 8 sec. each; Human Solutions GmbH, Kaiserslautern, Germany) and used as stimuli.

Other-rated facial attractiveness was based on facial photographs of the study`s participants and rated by 12 independent female raters on an 11-point scale (from -5 = "very unattractive" to +5 = "very attractive") and included as the mean of the responses to the two questions "How sexually attractive do you find this man?" (α=0.86) and "How attractive is this man for a long-term relationship?" (α=0.85).

Residual self-perceived physical attractiveness was calculated as in Study 1, but divided into two measures, the first representing residuals resulting from a regression of self-perceived

attractiveness on other-rated body attractiveness, the second being residuals from a regression on other-rated facial attractiveness.

Self-perceived mating success was measured exactly as in Study 1.

3.1.2.3 Formidability Measures

Strength was assessed as in Study 1, but testing hand grip strength for participants' self-reported dominant hand only (88.2% used their right, the remaining 11.8% their left hand). The correlation between dominant hand grip and upper body strength was $r=0.41$ ($p<0.001$).

Lung function was assessed as in Study 1, only assessing forced expiratory volume per 1 second (FEV), but not forced vital capacity (FVC).

Body masculinity was calculated similar to Study 1. The variables forearm circumference, biceps circumference, chest circumference, waist-to-hip ratio, bust-to-underbust ratio, and leg-to-body ratio were extracted as automatic measurements from the body scanner (according to ISO 20685:200) and included as the average of the z-standardized values from three body scans. For forearm and biceps circumference we included the maximum value of measurements from the left and right arm. Shoulder breadth and height were measured manually. The PCA over all variables explained 43% of the total variance.

Facial masculinity and *BMI-controlled body masculinity* were assessed exactly as in Study 1.

3.1.3 Statistical Analyses

We z-standardized all variables prior to analysis and graphically inspected the data for normality and outliers. From the original sample (165 males) we excluded 1 influential case based on outlier values in body appearance (values beyond 1.5 * interquartile range above or below the third or first quartile) that showed high leverage in regression diagnostics (graphical inspection of the bivariate distribution between residuals and leverage of data points;

distribution of Cook's distances; see figure S1.3 in the supplementary material S1). For the resulting sample (164 males) we computed zero-order Pearson correlation coefficients between all variables and their bootstrapped 95% confidence intervals. Again, we adjusted all p-values for false discovery rate (FDR, Benjamini & Hochberg, 1995) and additionally computed partial Pearson correlation coefficients controlling for age (ethnicity was not a varying factor in this sample). We checked whether changes in statistical significance (in terms of a threshold $p<0.05$) occurred between the matrices of zero-order and partial correlations and assessed their similarity via the congruency coefficient Rc (see above).

3.2 Results

Self-perceived attractiveness and self-perceived mating success were significantly correlated ($r=0.47$, $p<0.001$). Self-perceived mating success was stronger correlated to objective measures of the body of participants (body masculinity, lung function, and height) and other-rated attractiveness than self-perceived attractiveness. Notably, other people's ratings of the body were unrelated to self-perceived measures of both attractiveness and mating success (see Table 4) and other-rated facial attractiveness only showed a significant correlation with self-perceived mating success. Measures of other-rated body and facial attractiveness were however significantly correlated ($r=0.40$, $p<0.001$), which is consistent with the one ornament hypothesis (Grammer et al., 2003; Thornhill & Grammer, 1999; but see Honekopp et al., 2007) and thus indicative of the validity of the ratings.

Measures of self-perceived attractiveness were positively correlated with extraversion and inversely with neuroticism, while self-perceived mating success was correlated with conscientiousness. Extraversion was furthermore positively related to various measures of physical masculinity, while neuroticism showed a negative correlation with height (see Table 5). Measures of Dominance, Shyness, and Narcissism were largely unrelated to physical

masculinity or strength, apart from a negative correlation between Shyness and physical strength and a positive correlation among Dominance and BMI-controlled body masculinity. Sociosexual Orientation, on the other hand, correlated positively with a variety of physical masculinity and strength measures, notably only the facets of Attitude and Behavior.

The similarity between partial correlations controlling for age and zero-order correlations was highly significant (Rc=0.999; 95% CI= [0.999; 1]; p<0.001). Changes of statistical significance between partial- and zero-order correlations occurred in six cases, four cases reflecting relationships between personality and formidability / physical attractiveness, namely among Sociosexual Behavior and self-perceived attractiveness, Shyness and strength, Neuroticism and residual self-perceived attractiveness (corrected for rated facial attractiveness), and Neuroticism and height. Results of correlations based on single anthropometric measures of the body and the face of individuals can be found in the supplementary material (S4).

Table 4

Pearson correlation coefficients between self-rated and objective formidability measures

	SP attractiveness	SP mating success
Other-rated facial attractiveness	0.15 [-0.03;0.33]	***0.30 [0.18;0.42]***
Other-rated body attractiveness	0.07 [-0.10;0.22]	0.11 [-0.07;27]
Body masculinity	0.09 [-0.06;0.20]	***0.20 [0.06;0.33]***

BMI-controlled body masculinity	***0.25 [0.05;0.38]***	***0.35 [0.19;0.50]***
Upper body size	0.08 [-0.10;0.18]	***0.20 [0.07;0.31]***
Facial masculinity	-0.01 [-0.20;0.23]	0.09 [-0.06;0.24]
Strength	0.13 [-0.08;0.23]	*0.18 [0.04;0.31]*
Lung function	***0.19 [0.02,0.31]***	***0.32 [0.21;0.45]***
Height	0.09 [-0.07;0.22]	***0.25 [0.12;0.38]***
BMI	-0.05 [-0.20;0.05]	0.01 [-0.13;0.15]

Significant results in terms of FDR-adjusted p-values <.05 are displayed in bold, confidence intervals not containing the value 0 in italics. SP = Self-perceived.

Table 5

Pearson correlation coefficients between formidability and personality measures from Study 2

	Big 5					**Additional Personality**				**Sociosexual Orientation**			
	E	A	N	O	C	D	S	NA	NR	SOI-R	S-A	S-B	S-D
Self-perceived attractiveness	***0.25*** ***[0.13; 0.40]***	0.05 [-0.07; 0.21]	***-0.22*** ***[-0.36; -0.06]***	0.04 [-0.11; 0.16]	0.13 [-0.06; 0.25]	***0.35*** ***[0.18; 0.47]***	*-0.18* *[-0.31; -0.01]*	***0.33*** ***[0.19; 0.44]***	0.00 [-0.15; 0.17]	***0.28*** ***[0.14; 0.43]***	***0.23*** ***[0.07; 0.39]***	***0.19*** ***[0.05; 0.33]***	***0.20*** ***[0.02; 0.34]***
Residual self-perceived attractiveness (corrected for rated body attractiveness)	***0.22*** ***[0.08; 0.38]***	0.07 [-0.06; 0.23]	***-0.20*** ***[-0.34; -0.05]***	0.05 [-0.09; 0.16]	0.16 [-0.03; 0.30]	***0.31*** ***[0.16; 0.42]***	-0.14 [-0.27; 0.02]	***0.35*** ***[0.22; 0.44]***	0.00 [-0.17; 0.18]	***0.25*** ***[0.09; 0.41]***	***0.19*** ***[0.02; 0.36]***	*0.17* *[0.01; 0.30]*	***0.20*** ***[0.05; 0.35]***
Residual self-perceived attractiveness (corrected for rated facial attractiveness)	***0.23*** ***[0.08; 0.37]***	0.03 [-0.09; 0.17]	*-0.19* *[-0.32; -0.03]*	0.08 [-0.07; 0.19]	0.13 [-0.05; 0.26]	***0.32*** ***[0.16; 0.40]***	-0.15 [-0.28; 0.02]	***0.35*** ***[0.20; 0.45]***	0.02 [-0.14; 0.19]	***0.25*** ***[0.10; 0.40]***	***0.19*** ***[0.03; 0.34]***	*0.16* *[0.02; 0.30]*	***0.21*** ***[0.05; 0.35]***
Self-perceived mating success	***0.40*** ***[0.27; 0.55]***	0.04 [-0.12; 0.14]	***-0.29*** ***[-0.43; -0.11]***	0.05 [-0.10; 0.18]	***0.20*** ***[0.04; 0.34]***	***0.48*** ***[0.34; 0.58]***	***-0.35*** ***[-0.49; -0.20]***	***0.31*** ***[0.14; 0.44]***	-0.02 [-0.16; 0.18]	***0.36*** ***[0.21; 0.49]***	***0.30*** ***[0.15; 0.43]***	***0.35*** ***[0.19; 0.47]***	*0.17* *[0.02; 0.33]*
Other-rated facial attractiveness	0.09 [-0.04; 0.23]	0.17 [0.00; 0.35]	-0.10 [-0.24; 0.05]	-0.14 [-0.28; 0.01]	0.03 [-0.11; 0.20]	0.05 [-0.08; 0.17]	-0.05 [-0.18; 0.08]	-0.05 [-0.17; 0.09]	-0.08 [-0.21; 0.07]	0.10 [-0.04; 0.24]	0.05 [-0.11; 0.15]	0.10 [-0.03; 0.24]	0.07 [-0.09; 0.23]
Other-rated body attractiveness	0.01 [-0.10; 0.17]	-0.06 [-0.22; 0.05]	0.05 [-0.06; 0.19]	-0.14 [-0.28; 0.05]	0.01 [-0.14; 0.20]	-0.04 [-0.18; 0.08]	-0.04 [-0.16; 0.09]	0.10 [-0.07; 0.24]	0.05 [-0.11; 0.18]	0.00 [-0.17; 0.15]	-0.04 [-0.18; 0.11]	-0.05 [-0.25; 0.13]	0.09 [-0.08; 0.23]
Body masculinity	***0.28*** ***[0.14; 0.42]***	0.02 [-0.12; 0.13]	-0.10 [-0.26; 0.01]	-0.07 [-0.17; 0.06]	-0.03 [-0.18; 0.09]	*0.17* *[0.04; 0.31]*	*-0.15* *[-0.28; -0.05]*	0.03 [-0.11; 0.19]	0.03 [-0.14; 0.18]	***0.26*** ***[0.12; 0.41]***	***0.24*** ***[0.12; 0.40]***	***0.33*** ***[0.19; 0.48]***	0.02 [-0.13; 0.18]
BMI-controlled body masculinity	***0.20*** ***[0.04; 0.33]***	0.06 [-0.08; 0.16]	*-0.18* *[-0.28; -0.02]*	-0.09 [-0.23; 0.03]	-0.07 [-0.20; 0.06]	***0.22*** ***[0.03; 0.36]***	-0.12 [-0.26; 0.04]	0.10 [-0.04; 0.22]	0.00 [-0.12; 0.16]	***0.28*** ***[0.15; 0.41]***	***0.29*** ***[0.12; 0.44]***	***0.27*** ***[0.16; 0.39]***	0.07 [-0.05; 0.24]
Upper body size	***0.25*** ***[0.13; 0.38]***	0.02 [-0.15; 0.14]	-0.07 [-0.24; 0.08]	-0.03 [-0.14; 0.11]	-0.01 [-0.15; 0.10]	0.14 [-0.01; 0.29]	-0.11 [-0.23; 0.01]	0.03 [-0.11; 0.19]	0.01 [-0.19; 0.19]	***0.26*** ***[0.12; 0.41]***	***0.19*** ***[0.09; 0.33]***	***0.32*** ***[0.18; 0.46]***	0.07 [-0.11; 0.23]
Facial masculinity	0.06 [-0.10; 0.24]	-0.11 [-0.25; 0.05]	-0.07 [-0.27; 0.13]	-0.17 [-0.32; 0.00]	-0.07 [-0.21; 0.11]	0.09 [-0.11; 0.28]	-0.11 [-0.30; 0.04]	-0.10 [-0.29; 0.06]	0.03 [-0.17; 0.20]	0.09 [-0.04; 0.24]	0.07 [-0.07; 0.24]	0.13 [-0.02; 0.29]	0.01 [-0.13; 0.19]

Strength	0.12 [-0.05; 0.26]	-0.08 [-0.21; 0.08]	*-0.18 [-0.33; -0.02]*	0.01 [-0.13; 0.16]	-0.02 [-0.19; 0.13]	0.13 [-0.03; 0.25]	**-0.19 [-0.37; 0.00]**	0.03 [-0.10; 0.17]	-0.06 [-0.23; 0.13]	*0.17 [0.05; 0.33]*	***0.19 [0.06; 0.33]***	***0.22 [0.08; 0.37]***	-0.03 [-0.16; 0.12]
Lung function	***0.21 [0.02; 0.32]***	0.13 [0.00; 0.26]	-0.14 [-0.27; 0.00]	-0.06 [-0.20; 0.07]	-0.03 [-0.17; 0.10]	0.14 [-0.02; 0.30]	-0.11 [-0.25; 0.04]	0.10 [-0.08; 0.27]	-0.07 [-0.21; 0.07]	0.10 [-0.08; 0.27]	0.13 [-0.07; 0.29]	0.11 [-0.04; 0.24]	-0.03 [-0.17; 0.1]
Height	0.12 [-0.03; 0.23]	0.17 [0.03; 0.29]	***-0.18 [-0.32; -0.01]***	-0.08 [-0.26; 0.09]	-0.10 [-0.23; -0.01]	0.10 [-0.06; 0.26]	-0.06 [-0.19; 0.08]	-0.08 [-0.24; 0.06]	-0.08 [-0.20; 0.04]	*0.16 [0.03; 0.26]*	***0.22 [0.08; 0.34]***	0.11 [-0.02; 0.23]	0.01 [-0.12; 0.14]
BMI	***0.21 [0.06; 0.34]***	-0.02 [-0.12; 0.11]	-0.01 [-0.15; 0.13]	-0.02 [-0.14; 0.10]	0.01 [-0.14; 0.13]	0.07 [-0.10; 0.22]	-0.10 [-0.22; 0.02]	-0.03 [-0.16; 0.13]	0.03 [-0.12; 0.18]	0.13 [-0.03; 0.30]	0.10 [-0.08; 0.29]	***0.22 [0.06; 0.37]***	-0.03 [-0.16; 0.15]

Significant results in terms of FDR-adjusted p-values <.05 are displayed in bold, confidence intervals not containing the value 0 in italics. E= Extraversion, A= Agreeableness, N= Neuroticism, O= Openness, C= Conscientiousness, D= Dominance, S= Shyness, NA= Narcissistic Admiration, NR= Narcissistic Rivalry, SOI-R= revised Sociosexual Orientation Inventory, S-A= Sociosexual Attitudes, S-B= Sociosexual Behavior, S-D= Sociosexual Desire

4. Discussion

Consistent with previous research, we found evidence for a relationship between self-rated attractiveness and various personality measures. With regard to objective indicators of formidability and other-rated physical attractiveness we mostly attained null findings, thereby not supporting the notion of reactive heritability of personality. Additional indices accounting for shared variance among self-perceived and other-rated physical attractiveness and potential confounding influences of BMI on anthropometric measurements corroborated this pattern of results. Given the limits of our sample sizes, we can, however, not rule out small effects of facultative calibration in the tested domains. Apart from the overall trend, some personality measures also did show relationships with objective trait measurements consistent with the theory of facultative calibration. This calls for a differentiated discussion of our findings. We will rest this discussion upon an integrated summary of the results of both studies given in Table 6.

Table 6

An overview of the results from Studies 1 and 2.

Women	Self-perceived attractiveness	Other-rated attractiveness	Physical strength	Anthropometric Measurements
Anger	positive association only for Success in Conflict	positive association only for Success in Conflict	no association	no association

Extraversion	positive association	no association	no association	no association
Narcissism	positive association only for the facet of Admiration	no association	no association	no association
Sociosexual orientation	positive association only for the facet of Behavior	no association	no association	no association
Men	Self-perceived attractiveness	Other-rated attractiveness	Physical strength	Anthropometric Measurements
Anger	positive association	no association	no association	positive association only for Utility of Personal Aggression
Extraversion	positive association	no association	mixed results	mixed results
Narcissism	positive association only for the	no association	no association	no association

	facet of Admiration			
Sociosexual orientation	positive association	no association	mixed results	mixed results

With high consistency among both studies, we did hardly find any evidence for more objective, that is other-rated, physical attractiveness being a trait driving adaptive calibration of personality in men or women (for a summary, see Figure 1). Referring to substantial correlations between self- and other-rated attractiveness, some previous studies merged these measures (Lukaszewski et al., 2014) or integrated them in a path model (Lukaszewski, 2013). Therefore we additionally assessed the effect of self-perceived attractiveness when controlling for different measures of other-rated attractiveness (i.e. residual self-perceived physical attractiveness). We found only miniscule changes in the correlations between self-rated attractiveness and all behavioral measures when controlling for other-rated attractiveness. This suggests that the relationship between physical attractiveness and personality was mainly not due to shared variance of self-perceived with more objective measures of physical attractiveness, but almost solely relied on how participants perceived themselves, regardless of how other people judged their looks. In direct comparison to Lukaszewski et al.'s (2014) result of a correlation between a composite measure of physical attractiveness and Sociosexual Orientation in men, we additionally computed composite measures of attractiveness for our samples (their z-standardized mean). Overall, the results turned out nonsignificant (supplementary material S2), although we cannot entirely rule out a small effect of composite indices based on body attractiveness. However, as we showed that rated body attractiveness was not related to Sociosexual Orientation (or had a very small effect at best), we conclude that our effect of composite physical attractiveness, and perhaps the effect of Lukaszewski (2014)

as well, is foremost driven by self-perceived attractiveness. This could call for entirely different theories to explain such covariation. Haysom et al. (2015) already discussed the potential role of a positivity bias in the relationship among self-perceived attractiveness and extraversion. Also, a mediating role of self-esteem, one of the strongest correlates of self-perceived attractiveness (Feingold, 1992), could explain attractiveness-personality relationships without a mechanism of facultative calibration, as it is entailed in or correlates with personality constructs such as extraversion (Robins et al., 2001), narcissism (Morf & Rhodewalt, 2001), and sociosexuality (Jonason, Teicher, & Schmitt, 2011). At last, the correlations between self-perceived and other-rated attractiveness were also not particularly strong in our samples, especially for women in Study 1 and men in Study 2. In fact, we would not expect them to be, given a long-known meta-analytical effect size of r=0.24 for both sexes (Feingold, 1988). This could possibly be due to a flawed operationalization of physical attractiveness by external information like anthropometric measurements or other people's judgements. An impoverishment of judgments of body attractiveness could have occurred from only rating the participants` body scans, which lack features such as skin tone or texture. As however ratings of facial attractiveness were in line with the results of body attractiveness, a disparity among self- and other-rated attractiveness could also be, as argued above, due to biases in the assessment of one's own attractiveness. These issues taken aside, effects of both, internal and external, representations of physical attractiveness or their shared variation would have been more compelling results in line with facultative calibration than a solitary effect of self-perceived attractiveness.

Consistent with the theory of facultative calibration, physical strength and physical masculinity were related to extraverted behavior in men. Interestingly, on a facet level analysis of Extraversion we found a relationship with being active and assertive, attributes that by definition require physical ability or imply a proneness to conflict. More prosocially orientated

facets such as Gregariousness or Warmth, however, were unrelated to formidability. Hence and perhaps unsurprisingly, male formidability may be only related to aspects of extraverted behavior that are inherently linked to strength and body condition. This goes at hand with Mõttus' (2016) suggestion of a stronger consideration of personality trait diversity that is using facets or even items as predictor, when linking personality to specific outcomes. In Study 2 we also found a domain-level correlation between BFI Extraversion and male formidability, suggesting overall calibration of Extraversion. Interestingly, the NEO Extraversion facets Assertiveness and Activity are the only two facets (out of six) that are clearly represented in the item pool of the BFI (Soto & John, 2009). The domain-level relationship from Study 2 could thus be foremost driven by facultative calibration of these specific aspects of extraverted personality. More studies are needed to secure a detailed knowledge of a formidability-extraversion-relationship.

Both studies differed substantially in their results for Sociosexual Orientation. We found some, although not consistent, evidence for a relationship of Sociosexuality with physical masculinity and physical strength in men in Study 2. As to be expected, significant correlations appeared in the facets of *Attitude* and *Behavior*, not in *Desire*. Contrary to that, the results of Study 1 did overall not support a relationship between formidability and Sociosexual Orientation. Thus, we remain with mixed evidence of whether markers of genetic quality (Gangestad & Simpson, 2000) may influence orientation towards and success in short-term mating for men. Again, more studies are needed to clarify the theoretical applicability of facultative calibration in this domain.

In addition to previous studies, we explored personality-formidability relationships for other traits of the Big Five personality spectrum. Whereas Openness to Experience, Agreeableness, and Conscientiousness were unrelated to physical formidability and other-rated attractiveness, we did find some, although again inconsistent, evidence for a negative

relationship between physical formidability and Neuroticism. Furthermore, there was a relationship between physical strength and Shyness in men. Linking this back to the correlations between Assertiveness and formidability, this could reflect that stronger and more masculine men are less fearful of potential conflicts. These results are furthermore consistent with previous findings showing a negative relationship between handgrip strength and Neuroticism in men (Fink, Weege, Pham, & Shackelford, 2016).

More recent empirical work on the theory of facultative calibration has extended its scope to a relationship between Aggressiveness and Coalitional Strength in adolescents (Sell et al., 2016). As these authors point out, Coalitional Strength could be influenced by a variety of traits that are rather unrelated to body condition (e.g. specialized knowledge, skills, or mutual interests). However, opening the idea of facultative calibration to indicators of social relationships increases the potential of finding links where directions of causality are hard to identify (especially in cross-sectional data) and potentially even circular. Although Coalitional Strength could still be causing facultative calibration, effects of behavior on Coalitional Strength are equally likely, even more so than from behavior to body condition. In our study we purposely focused on testing the idea of behavioral calibration to phenotypic traits related to bodily and facial appearance and body condition. The absence of correlations among personality and facial attractiveness or facial masculinity in our results thereby matches a recent study testing facultative calibration of egalitarianism (Price et al., 2017), which did only find significant correlations for bodily formidability as well.

Future Directions

A caveat of using cross-sectional data to test facultative calibration (which applies to our studies as well) is that some phenotypic traits discussed as anchors of recalibration, e.g. physical strength, are somewhat plastic. Thus, the mechanism of recalibration could be, at least in some cases, reversed. For example, individuals with greater genetic dispositions to be

extraverted, dominant, narcissistic, or aggressive might select their environments or elicit reactions from others in a way that they find themselves in competitive situations more often. In response to these experiences they might learn that a higher physical formidability would be beneficial to them and decide to work out in order to increase it. Similarly, more extraverted or narcissistic individuals might actively or reactively encounter social situations more often where higher attractiveness is more beneficial and as a response work out more, increase their grooming behavior or even become more likely to seek out cosmetic surgery. Such cases would explain an inverted causal direction, and they are indistinguishable in cross-sectional correlational studies. Indeed, Holtzman and Strube (2013) found stronger relationships between narcissism and effective adornment than natural beauty, and von Soest and colleagues (2009) showed that female patients undergoing cosmetic surgery were already more extraverted prior to surgery than females from a representative control sample, indicating another possibility of a reversed causal direction related to body attractiveness. Both examples are possible scenarios of a gene-environment-correlation, where a genetically influenced personality trait leads to the selection of and adaption to specific environments (Bleidorn et al., 2014). Another example of reverse causation could occur via positive ontogenetic feedback among personality traits such as Extraversion or Aggressiveness and self-esteem, which in turn may influence levels of self-perceived attractiveness.

As facultative calibration has not been proposed to be the one and only mechanism driving personality variation, a control of genetic confounders will gather advocates and opponents of the theory of facultative calibration most likely under the same umbrella (see also Lukaszewski & Roney, 2015, on this matter). The common method of assessing reactive heritability in genetically informed studies that have so far tested facultative calibration was to compare the heritability of the personality trait before and after removing the shared genetic variance with another correlated phenotypic trait (Haysom et al., 2015; von Rueden et al.,

2015). This is an interesting approach, since it indicates whether the genetic components of traits are independent or not. However, a substantial amount of shared genetic variation is still not a proof of facultative calibration, since such pleiotropy can still be due to various different mechanisms, including reverse causality and biological pleiotropy (Johnson et al., 2011; Solovieff et al., 2013). The approach of Lukaszewski and Roney (2011; see above) to assess the independence of effects due to calibration and a genetic polymorphism in the androgen receptor gene was thus commendable, however to be improved in complexity and sample size, since a single polymorphism is highly unlikely to reflect the genetic origin of complex trait variation (which is known to consist of a large number of very small additive genetic effects, see Munafò & Flint, 2011; Zietsch, 2016). In the animal literature, correlations between physiological and behavioral traits have already been theoretically, though not consistently empirically, differentiated into either trans-generational genetic effects (a pace-of-life syndrome; Réale et al., 2010) or ontogenetic adaptations (a state-behavior feedback loop; Sih et al., 2015; Wolf & Weissing, 2010). Empirically dissecting correlations based on such a two-fold theoretical classification may be of advantage for the human literature as well. A suited study design to further explore the origin of phenotypic covariation in a quasi-experimental way would be a longitudinal twin difference study (McGue et al., 2010), that could test the effect of differences in formidability or physical attractiveness among twins while controlling for genetic confounders and reversed causality due to ontogenetic changes in physical traits.

Acknowledgements

We would like to thank Hazel Hamilton, Sanni Kujala, Chih-Hao Chang, Amanda Ewen, Sven Radcke, Carmen Ponce, and Tanya Frostestad Salvesen for helping with the collection of the data of Study I. We acknowledge support by the Leibniz Association through funding for the

Leibniz ScienceCampus Primate Cognition and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of the Project number 254142454 / GRK 2070.

Data Availability

The data associated with this research are available at https://osf.io/vyw8b/.

Funding

None.

References

Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What Does the Narcissistic Personality Inventory Really Measure? *Assessment*, *18*(1), 67–87. https://doi.org/10.1177/1073191110382845

Archer, J., & Thanzami, V. (2007). The relation between physical aggression, size and strength, among a sample of young Indian men. *Personality and Individual Differences*, *43*(3), 627–633. https://doi.org/10.1016/j.paid.2007.01.005

Asendorpf, J.B., & Wilpers, S. (1998). Personality effects on social relationships. J*ournal of Personality and Social Psychology*, *74*, 1531-1544.

Back, M. D., Küfner, A. C. P., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. A. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology*, *105*(6), 1013–1037. https://doi.org/10.1037/a0034431

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Bleidorn, W., Kandler, C., & Caspi, A. (2014). The Behavioural Genetics of Personality Development in Adulthood-Classic, Contemporary, and Future Trends: Behavioural genetics of personality development. *European Journal of Personality*, *28*(3), 244–255. https://doi.org/10.1002/per.1957

Buss, D.M. & Hawley, P.H. (Eds.) (2010). *The evolution of personality and individual differences* (pp. 243-279). New York: Oxford University Press.

Buss, D. M., & Penke, L. (2015). Evolutionary personality psychology. In M. Mikulincer, P. R. Shaver, M. L. Cooper & R. J. Larsen (Hrsg.): *APA Handbook of Personality and*

Social Psychology, Vol. 4: Personality Processes and Individual Differences (S. 3-29). *Washington, DC:* American Psychological Association.

Campbell, W. K., Bonacci, A. M., Shelton, J., Exline, J. J., & Bushman, B. J. (2004). Psychological Entitlement: Interpersonal Consequences and Validation of a Self-Report Measure. *Journal of Personality Assessment*, *83*(1), 29–45. https://doi.org/10.1207/s15327752jpa8301_04

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, Fl.: Psychological Assessment.

Feingold, A. (1988). Matching for attractiveness in romantic partners and same-sex friends: A meta-analysis and theoretical critique. *Psychological Bulletin*, *104*(2), 226–235. https://doi.org/10.1037/0033-2909.104.2.226

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, *111*(2), 304–341. https://doi.org/10.1037/0033-2909.111.2.304

Fink, B., Weege, B., Pham, M. N., & Shackelford, T. K. (2016). Handgrip strength and the Big Five personality factors in men and women. *Personality and Individual Differences*, *88*, 175–177. https://doi.org/10.1016/j.paid.2015.09.013

Gabriel, M. T., Critelli, J. W., & Ee, J. S. (1994). Narcissistic Illusions in Self-Evaluations of Intelligence and Attractiveness. *Journal of Personality*, *62*(1), 143–155. https://doi.org/10.1111/j.1467-6494.1994.tb00798.x

Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: trade-offs and strategic pluralism. *The Behavioral and Brain Sciences*, *23*(4), 573-587-644.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Gough, H. G. (1956). *California Psychological Inventory*. Palo Alto, Ca.: Consulting Psychologists Press.

Grammer, K., Fink, B., Mller, A. P., & Thornhill, R. (2003). Darwinian aesthetics: sexual selection and the biology of beauty. *Biological Reviews*, *78*(3), 385–407. https://doi.org/10.1017/S1464793102006085

Haysom, H. J., Mitchem, D. G., Lee, A. J., Wright, M. J., Martin, N. G., Keller, M. C., & Zietsch, B. P. (2015). A test of the facultative calibration/reactive heritability model of extraversion. *Evolution and Human Behavior*, *36*(5), 414–419. https://doi.org/10.1016/j.evolhumbehav.2015.03.002

Heymsfield, S. B., Scherzer, R., Pietrobelli, A., Lewis, C. E., & Grunfeld, C. (2009). Body mass index as a phenotypic expression of adiposity: quantitative contribution of muscularity in a population-based sample. *International Journal of Obesity*, *33*(12), 1363–1373. https://doi.org/10.1038/ijo.2009.184

Holtzman, N. S., & Strube, M. J. (2010). Narcissism and attractiveness. *Journal of Research in Personality*, *44*(1), 133–136. https://doi.org/10.1016/j.jrp.2009.10.004

Holtzman, N. S., & Strube, M. J. (2013). People With Dark Personalities Tend to Create a Physically Attractive Veneer. *Social Psychological and Personality Science*, *4*(4), 461–467. https://doi.org/10.1177/1948550612461284

Honekopp, J., Rudolph, U., Beier, L., Liebert, A., & Muller, C. (2007). Physical attractiveness of face and body as indicators of physical fitness in men. *Evolution and Human Behavior*, *28*(2), 106–111. https://doi.org/10.1016/j.evolhumbehav.2006.09.001

Isen, J. D., McGue, M. K., & Iacono, W. G. (2015). Aggressive-Antisocial Boys Develop Into Physically Strong Young Men. *Psychological Science*, *26*(4), 444–455. https://doi.org/10.1177/0956797614567718

Jacobs, I., & Scholl, W. (2005). Interpersonale Adjektivliste (IAL). *Diagnostica*, *51*(3), 145–155. https://doi.org/10.1026/0012-1924.51.3.145

Johnson, W., Penke, L., & Spinath, F. M. (2011). Heritability in the era of molecular genetics: Some thoughts for understanding genetic influences on behavioral traits. *European Journal of Personality*, *25*, 254-266.

Jonason, P. K., Teicher, E. A., & Schmitt, D. P. (2011). The TIPI's validity confirmed: Associations with sociosexuality and self-esteem. *Individual Differences Research*, *9*(1), 52-60.

Kinsey, A. C., Pomeroy, W. R., & Martin, C. E. (1948). *Sexual Behavior in the Human Male*. Philadelphia, PA: W.B. Saunders.

Landolt, M. A., Lalumière, M. L., & Quinsey, V. L. (1995). Sex differences in intra-sex variations in human mating tactics: An evolutionary approach. *Ethology and Sociobiology*, *16*(1), 3–23. https://doi.org/10.1016/0162-3095(94)00012-V

Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. *Diagnostica*, *47*(3), 111–121. https://doi.org/10.1026//0012-1924.47.3.111

Lukaszewski, A. W. (2013). Testing an Adaptationist Theory of Trait Covariation: Relative Bargaining Power as a Common Calibrator of an Interpersonal Syndrome: Adaptationist theory of trait covariation. *European Journal of Personality*, *27*(4), 328–345. https://doi.org/10.1002/per.1908

Lukaszewski, A. W., Larson, C. M., Gildersleeve, K. A., Roney, J. R., & Haselton, M. G. (2014). Condition-dependent calibration of men's uncommitted mating orientation: evidence from multiple samples. *Evolution and Human Behavior*, *35*(4), 319–326. https://doi.org/10.1016/j.evolhumbehav.2014.03.002

Lukaszewski, A. W., & Roney, J. R. (2011). The Origins of Extraversion: Joint Effects of Facultative Calibration and Genetic Polymorphism. *Personality and Social Psychology Bulletin*, *37*(3), 409–421. https://doi.org/10.1177/0146167210397209

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, *36*(1), 1–15. https://doi.org/10.1017/S0140525X11002160

McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology*, *75*(6), 1586–1603.

McGue, M., Osler, M., & Christensen, K. (2010). Causal Inference and Observational Research: The Utility of Twins. *Perspectives on Psychological Science*, *5*(5), 546–556. https://doi.org/10.1177/1745691610383511

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the Paradoxes of Narcissism: A Dynamic Self-Regulatory Processing Model. *Psychological Inquiry*, *12*(4), 177–196. https://doi.org/10.1207/S15327965PLI1204_1

Mõttus, R. (2016). Towards More Rigorous Personality Trait-Outcome Research: Personality traits and outcomes. *European Journal of Personality*, *30*(4), 292–303. https://doi.org/10.1002/per.2041

Munafò, M. R., & Flint, J. (2011). Dissecting the genetic architecture of human personality. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.07.007

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Parker, G. A. (1974). Assessment strategy and the evolution of fighting behaviour. *Journal of Theoretical Biology*, *47*(1), 223–243.

Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*(6), 556–563. https://doi.org/10.1016/S0092-6566(02)00505-6

Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: a more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, *95*(5), 1113–1135. https://doi.org/10.1037/0022-3514.95.5.1113

Penke, L., Denissen, J. J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality*, *21*(5), 549–587. https://doi.org/10.1002/per.629

Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B., Burt, D. M., & Perrett, D. I. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1476), 1617–1623. https://doi.org/10.1098/rspb.2001.1703

Petersen, M., Sell, A., Tooby, J., and Cosmides, L. (2010). Evolutionary psychology and criminal justice: A recalibrational theory of punishment and reconciliation**.** In H. Høgh-Olesen (Ed.): *Human Morality and Sociality: Evolutionary and Comparative Perspectives* (pp. 72-131). New York: Palgrave MacMillan.

Price, M. E., Dunn, J., Hopkins, S., & Kang, J. (2012). Anthropometric correlates of human anger. *Evolution and Human Behavior*, *33*(3), 174–181. https://doi.org/10.1016/j.evolhumbehav.2011.08.004

Price, M. E., Sheehy-Skeffington, J., Sidnaius, J., & Pound, N. (2017). Is sociopolitical egalitarianism related to bodily and facial formidability in men? *Evolution and Human Behavior*, *38*(5), 626–634. https://doi.org/10.1016/j.evolhumbehav.2017.04.001

Réale, D., Garant, D., Humphries, M. M., Bergeron, P., Careau, V., & Montiglio, P.-O. (2010). Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1560), 4051–4063. https://doi.org/10.1098/rstb.2010.0208

Revelle, W. (2016). psych: Procedures for Personality and Psychological Research (Version 1.6.9). Evanston, Illinois, USA: Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych

Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. D. (2001). Personality Correlates of Self-Esteem. *Journal of Research in Personality*, *35*(4), 463–482. https://doi.org/10.1006/jrpe.2001.2324

Sell, A., Eisner, M., & Ribeaud, D. (2016). Bargaining power and adolescent aggression: the role of fighting ability, coalitional strength, and mate value. *Evolution and Human Behavior*, *37*(2), 105–116. https://doi.org/10.1016/j.evolhumbehav.2015.09.003

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, *106*(35), 15073–15078. https://doi.org/10.1073/pnas.0904312106

Sih, A., Mathot, K. J., Moirón, M., Montiglio, P.-O., Wolf, M., & Dingemanse, N. J. (2015). Animal personality and state–behaviour feedbacks: a review and guide for empiricists. *Trends in Ecology & Evolution*, *30*(1), 50–60. https://doi.org/10.1016/j.tree.2014.11.004

Smith, K. L., Cornelissen, P. L., & Tovée, M. J. (2007). Color 3D bodies and judgements of human female attractiveness. *Evolution and Human Behavior*, *28*(1), 48–54. https://doi.org/10.1016/j.evolhumbehav.2006.05.007

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics*, *14*, 483-495.

Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, *43*(1), 84–90. https://doi.org/10.1016/j.jrp.2008.10.002

Stulp, G., Kordsmeyer, T., Buunk, A. P., & Verhulst, S. (2012). Increased aggression during human group contests when competitive ability is more similar. *Biology Letters*, *8*(6), 921–923. https://doi.org/10.1098/rsbl.2012.0591

Thornhill, R., & Grammer, K. (1999). The Body and Face of Woman. *Evolution and Human Behavior*, *20*(2), 105–120. https://doi.org/10.1016/S1090-5138(98)00044-0

Tiddeman, B. P., Duffy, N., & Rabey, G. (2000). Construction and visualisation of three-dimensional facial statistics. *Computer Methods and Programs in Biomedicine*, *63*, 9-20.

Tooby, J., & Cosmides, L. (1990). On the universality of human nature and the uniqueness of the individual: the role of genetics and adaptation. *Journal of Personality*, *58*(1), 17–67.

Tooby, J., Cosmides, L., Sell, A., Lieberman, D., & Sznycer, D. (2008). Internal regulatory variables and the design of human motivation: A computational and evolutionary approach. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 251–271). New York, NY, USA: Psychology Press.

von Rueden, C. R., Lukaszewski, A. W., & Gurven, M. (2015). Adaptive personality calibration in a human society: effects of embodied capital on prosocial traits. *Behavioral Ecology*, *26*(4), 1071–1082. https://doi.org/10.1093/beheco/arv051

von Soest, T., Kvalem, I. L., Skolleborg, K. C., & Roald, H. E. (2009). Cosmetic surgery and the relationship between appearance satisfaction and extraversion: Testing a transactional model of personality. *Journal of Research in Personality*, *43*(6), 1017–1025. https://doi.org/10.1016/j.jrp.2009.07.001

Wetzel, E., Leckelt, M., Gerlach, T. M., & Back, M. D. (2016). Distinguishing Subgroups of Narcissists with Latent Class Analysis: Subgroups of narcissists. *European Journal of Personality*, *30*(4), 374–389. https://doi.org/10.1002/per.2062

Wiggins, J. S., Trapnell, P., & Phillips, N. (1988). Psychometric and Geometric Characteristics of the Revised Interpersonal Adjective Scales (IAS-R). *Multivariate Behavioral Research*, *23*(4), 517–530. https://doi.org/10.1207/s15327906mbr2304_8

Wolf, M., & Weissing, F. J. (2010). An explanatory framework for adaptive personality differences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1560), 3959–3968. https://doi.org/10.1098/rstb.2010.0215

Zietsch, B. P. (2016). Individual differences as the output of evolved calibration mechanisms: does the theory make sense in view of empirical observations? *Current Opinion in Psychology*, *7*, 71–75. https://doi.org/10.1016/j.copsyc.2015.08.014

Figure captions

Figure 1: Pearson correlation coefficient r between combined personality measures and different types of attractiveness measures for the two samples. The dashed lines indicate the threshold of statistical significance based on $\alpha=0.05$ for the respective sample. If correlations were a priori hypothesized to be negative (concerning the personality variables shyness and neuroticism), they were multiplied by -1 to match the expected direction of the other correlations. The whiskers of the boxplots represent minimum or maximum values that do not exceed the range of 1.5 * interquartile range (vertical size of the box).

Figures

Figure 1