AARHUS BSS
SCHOOL OF BUSINESS AND SOCIAL SCIENCES
AARHUS UNIVERSITY

Formidability and human behavior: an interdisciplinary approach

PhD dissertation

Dan Nguyen

Aarhus BSS, Aarhus University
Department of Economics and Business Economics

2016

PREFACE

This dissertation was written from October 2013 to September 2016 while I was enrolled as a PhD student on the PhD program in Social Science and Business at Aarhus University. I am grateful to the Department of Economics and Business Economics for providing me with an excellent and inspiring research environment and generous financial support.

I would like to give a special thanks to my fantastic team of supervisors Alexander Koch, Emma von Essen and Michael Bang Petersen for walking along on my interdisciplinary journey with endless support, discussions, advice, encouragement to develop and investigate my research ideas and challenging my perspectives. Thanks to Alex for encouraging me to pursue a PhD and always being open and constructive in his approach as a supervisor. I am looking forward to future collaboration.

During the academic year of 2015/2016 I visited Jim Sidanius at the Department of Psychology at Harvard University, Cambridge. I am grateful for the stay at the Sidanius Lab and the opportunity to interact with and be inspired by so many talented researchers both on an academic as well as a personal level. It has opened my eyes to social psychology and in particularly Social Dominance Theory. The opportunity to present my research at the lab provided very helpful comments and improvements to the second chapter. Thanks to Robin and Sarah for critical thoughts on my work and Lumumba for generously sharing his office with me. Also, many thanks to the rest of the people at Jim's lab: Mariska, Greg, Kiera, Sasha, Rachel and Milan for many wonderful times both at the lab and outside. Thanks to Aaron for making the time in Boston most memorable and providing a keen eye on the language of the thesis.

At Aarhus University I would like to thank my colleagues and fellow PhD students for providing me with hours of fun which all added to a very pleasant working environment. Special thanks to my office mates and the people on the C-wing: Christina, Leonie, Heidi, Mette, Christian, Anne, Max and Rune over the last three years. Without you guys, the time here would have been much less colorful and worthwhile! Also, many thanks to the MOB group as a whole, especially for the many interesting and thought provoking presentations and discussions at our seminars. Thanks to Josh and

Dan at the Cognition and Behavior Lab at Aarhus University for always providing excellent support and making sure my experiments could run as effortless as possible. Also, to Jonatan for providing many hours of technical support with z-Tree and providing valuable help in the lab. To the people at political science Lene, Lasse, Troels and Mathias and the rest of the political and evolutionary psychology crowd – many thanks for letting me take part in interesting events and discussions.

Finally, I am very thankful for the endless love and support I have received from my family – cảm ơn! Thanks to my beloved friends outside (as well as inside) of the academic world for never stop asking me about my projects and for being an outlet to both frustration and joy. I offer my most profound gratitude to Carsten for his unlimited support and encouragement – without him this dissertation would have had minimal chances at seeing the light of day!

Dan Nguyen
Aarhus, September 2016

UPDATED PREFACE

The assessment committee consists of Professor Anna Dreber Almenberg from Stockholm School of Economics, Professor Leda Cosmides from the University of California at Santa Barbara, and Professor Nabanita Datta Gupta from Aarhus University. I am thankful to the members of the committee for their thorough reading and very constructive and insightful comments. Some suggestions have already been incorporated in the current thesis, while others remain for future research.

Dan Nguyen
Aarhus, November 2016

CONTENTS

ENGLISH SUMMARY

O, it is excellent
To have a giant's strength, but it is tyrannous
To use it like a giant.

But man, proud man,
Drest in a little brief authority,
Most ignorant of what he's most assur'd,
His glassy essence, like an angry ape,
Plays such fantastic tricks before high heaven
As makes the angels weep.

-William Shakespeare, Measure for Measure

This dissertation contains four autonomous chapters. They are all motivated by observations of how evolution and ancestral life (1) influenced how the psychology functions in the human species; and (2) affected decision-making and, consequentially, behavior in modern humans.

An observable behavior is violence, which is found throughout the history of mankind. Violence is here defined as the use of physical force with the intent to injure another person or destroy property. While it does not seem to have been invented at a specific time, it is probably one of the most profound features of human nature. The continuous presence of violence has led psychologists to seek out cognitive mechanisms that would have been selected in order to aid human decision-making within this domain.

This quest to acquire knowledge about human psychology based on the challenges and environments our ancestors faced has come to be known as evolutionary psychology, which can best be described as an approach to psychology whereby testable hypotheses about human behavior are derived directly from evolutionary biology (Tooby and Cosmides, 1992). Particularly, the recent discovery of the human cognitive ability to accurately assess other humans' physical strength (formidability), a trait believed to be crucial for fighting ability, has largely inspired the course of this dissertation. If such cues of fighting ability were important to our ancestors for many thousands of generations, they would most likely still be relevant in modern society.

The first chapter, *Social Dominance and Conservatism Affect Assessment of Physical Formidability*, can be seen as a direct extension of previous research of the human cognitive ability to assess formidability (Sell et al., 2009). It is a project resulting from a fellowship with Professor Jim Sidanius at Harvard University. Professor Sidanius, the co-author of this chapter, is the founder of social dominance theory (Sidanius and Pratto, 2001), a general and theoretical framework of how group-based hierarchy and oppression is developed and sustained. One important psychological trait developed within this research is the social dominance orientation (SDO). It is a measure used to elicit individual differences in preference for hierarchical and unequal relations among groups, regardless of whose group is on the top or the bottom (Pratto et al., 2001). In an online survey experiment, we show that individuals vary in their assessments of other individuals' physical strength in a systematic way. Participants were presented with either a weak or strong person portrayed in a body silhouette and asked to rate the strength of the person. We find that individuals' evaluation of physical strength is conditioned by SDO and political conservatism such that individuals who report higher preferences on both dimensions also amplify their ratings of the persons shown. However, the effect is only present for individuals in the strong picture condition.

The second chapter, *Upper-body Strength and Human Conflict Resolution in Males*, is co-authored with Alexander Koch, Julia Nafziger and Michael Bang Petersen. In this study we apply the paradigmatic model to human males in order to explain animal conflict resolution: the asymmetric war or attrition model. The model has been examined and validated across many species in the animal kingdom. Animals coordinate conflicts based on simple physical cues of fighting ability. In the manuscript, we are the first to directly test the essence of this model in humans. To do so, we employed a non-physical, experimental game in which participants engaged in a conflict over an economic resource. One group of participants was exposed to cues about their opponents' physical strength, while the other was not. Even though strength was not objectively relevant for payoffs, participants used these cues when available to resolve the conflict (i.e., who obtained the resource and at what cost). Weaker males spontaneously cede resources to stronger males to avoid conflict escalation and do so more quickly if strength differences are larger. This key finding

demonstrates that evolutionarily relevant signals of fighting ability continue to shape the conflict resolution of modern humans. Thus, the peacefulness of humans emerges, in part, from rudimentary strategies for conflict resolution, present throughout the animal world. The chapter is currently accepted for revision and resubmission to the journal, Psychological Science.

The third chapter, *The Role of Physical Attractiveness and Strength in the Resolution of Intrasexual Female Conflicts*, is a mirror reflection of chapter two and is co-authored with Michael Bang Petersen and Alexander Koch. It is a study motivated by the observation that females who have a higher self-perceived attractiveness also report themselves as being more prone to anger, having a greater sense of entitlement and are more successful in conflict (Sell et al., 2009). In other words, for women, physical attractiveness would be directly equivalent to physical upper-body strength in males. In our study, we therefore directed our attention to female attractiveness. We employ the same paradigm as in chapter two, but instead expect physical attractiveness to be the decisive factor in female intra-sexual competition. We measure both self-reported attractiveness and waist-to-hip ratio, a measure that is often used as a marker for female attractiveness (Sing 1993, 1994; Sing et al., 2010; Streeter and McBurney, 2003). One group of participants received visual information about its opponents, allowing it to assess attractiveness, while the other group of participants had no information. We find robust evidence that contest duration is decreasing in the differences in attractiveness, and that this effect is driven by mutual assessment rather than being an artefact of internal processes that rely on self-assessment only. While the evidence suggests that less attractive females were quicker at opting out of the competition, the greater the differences in attractiveness should also prompt more attractive females to be more likely to win, these results are less clear cut. Our results provide tentative evidence that females rely on visual cues about attractiveness when engaging in intra-gender competition.

The fourth and final chapter, *Institutional Norms of Fairness and Support for Taxation*, is an investigation of how individuals' support for taxation responds to changes in the institutions. It is often argued that how the welfare state operates matters for citizens' willingness to pay tax. Yet, how institutions are shaped might well be an outcome of citizens' preferences,

thereby making causality hard to determine. Using a lab experiment, I test the effect of institutional rules on preferred tax rates. Building on Rothstein (1998; 2001), I study three variations in institutional rules that manipulate individuals' perception of substantial fairness (that only those deserving of help get help), procedural fairness (that help is given according to transparent rules), and fair burden (that the costs of distributional policies are fairly distributed between citizens). I find that perceptions of both substantial fairness and procedural fairness affect the participants' preferred tax rate. This partially corroborates the notion that how the welfare state operates matters for citizens' willingness to pay tax.

REFERENCES

Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange in Barkow J., Cosmides L., & Tooby J. (Ed.), *The adapted mind: Evolutionary psychology and the generation of culture (*pp. 206-221). New York: Oxford University Press

Pratto, F., Sidanius, J., Stallworth, L. M. & Malle, B.F. (2001). Social dominance orientation: A personality variable predicting social and political attitudes. In Michael A. Hogg, and Dominic Abrams, (Eds.). pp. 30-59. *Intergroup relations: Essential readings*. Philadelphia, PA: Psychology Press/Taylor & Francis

Rothstein, B. (1998). *Just institutions matter: the moral and political logic of the universal welfare state*. Cambridge University Press.

Rothstein, B. (2001). The universal welfare state as a social dilemma. *Rationality and society*, *13*(2), 213-233.

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.

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Singh, D. (1993). Adaptive significance of female physical attractiveness: role of waist-to-hip ratio. *Journal of personality and social psychology*, *65*(2), 293

Singh, D. (1994). Is thin really beautiful and good? Relationship between waist-to-hip ratio (WHR) and female attractiveness. *Personality and Individual Differences*, *16*(1), 123-132

Singh, D., Dixson, B. J., Jessop, T. S., Morgan, B., & Dixson, A. F. (2010). Cross-cultural consensus for waist–hip ratio and women's attractiveness.*Evolution and Human Behavior*, *31*(3), 176-181

Streeter, S. A., & McBurney, D. H. (2003). Waist–hip ratio and attractiveness: New evidence and a critique of "a critical test". *Evolution and Human Behavior*,*24*(2), 88-98

DANSK RESUMÉ

O hvor herligt
at eje jættestyrke, men hvor grumt
at bruge den som jætter gør.

men manden,
den stolte mand, udrustet med en lille,
kortvarig magt, og mest uvidende
om hvad han mest er sikker på, - sit væsen
så skørt som glas, -gør som en arrig abe
sådanne dårefagter op mod himlen
at englene må græde.

-William Shakespeare, Lige for lige

Denne afhandling består af fire selvstændige kapitler. De er baseret på iagttagelser om, hvordan vores forfædres liv og evolutionen har (1) præget menneskets psykologiske funktion og (2) påvirket hvordan det moderne menneske træffer beslutninger, der i sidste ende påvirker adfærd.

Vold er en observerbar handling, der altid har eksisteret i menneskets historie. Vold kan defineres som anvendelsen af fysisk kraft med en intention om at skade eller ødelægge en anden person eller ejendom. På trods af at ikke helt præcist ved, hvornår i menneskets historie volden er opstået, er det sandsynligvis en af de mest fundamentale elementer af menneskets natur.

Studiet af menneskets psykologi, hvor der trækkes på viden om vores forfædres liv, er kendt som *evolutionspsykologi*, der bedst kan beskrives som en tilgang, hvor testbare hypoteser om menneskelig adfærd udledes direkte fra dens biologien (Tooby and Cosmides, 1992). Min afhandling trækker på ny viden om menneskets evne til at estimere hinandens fysiske styrke. Fysisk styrke antages i denne litteratur at være direkte relateret til især mænds kampevne. Hvis signaler om kampstyrke har været vigtige for vores forfædre gennem tusindvis af generationer, er det højst sandsynligvis stadigvæk vigtigt i et moderne samfund.

Det første kapitel i afhandlingen, *Social Dominance and Conservatism Affect Assessment of Physical Formidability,* er en

videreførelse af den seneste forskning om menneskets evne til at vurdere fysisk styrke (Sell et al., 2009). Det er et projekt, der blev til i forbindelse med mit besøg hos professor Jim Sidanius på Harvard University, der også er medforfatter på kapitlet. Jim Sidanius er grundlæggeren af *socialdominansteorien,* en generel rammeteori, der søger at forklare, hvordan gruppebaseret hierarkier og undertrykkelse udvikles og fastholdes. Inden for dette område har man udviklet det vigtige personlighedstrækmål *socialdominansorientering (SDO)*, en skala, som måler individers præferencer for hierarkier imellem grupper (Sidanius et al., 2010). I et online spørgeskemaeksperiment viser vi, at individer varierer i deres vurderinger af andre individers fysiske styrke. I eksperimentet bliver deltagerne stillet overfor et kropssilhuet af enten en svag eller stærk mand og bedt om at vurdere, hvor stærk personen på billedet er. Vi viser, at individers vurdering af andres fysiske styrke er betinget af SDO og hvor politisk konservative de er, således at dem, der scorer højt på begge dimensioner, forstærker deres trusselvurdering. Effekten er kun til stede i den stærke betingelse.

Det andet kapitel, *Upper-body Strength and Human Conflict Resolution in Males*, er forfattet sammen med Alexander Koch, Julia Nafziger og Michael Bang Petersen. I studiet applicerer vi den paradigmatiske model, *the asymmetric war of attrition model, der* benyttes til at forklare konfliktresolution i dyreverden. Modellen er veludforsket og valideret på tværs af mange arter i dyreverden. Dyr koordinerer konflikter baseret på simple signaler om kampevne. Vi er de første til at applicere modellen på mennesker. For at kunne gøre dette anvendte vi et ikke-fysisk eksperimentelt spil, hvori deltagerne deltog i en konflikt over en monetær ressource. En gruppe blev udsat for signaler relateret til deres modstanders fysiske styrke, imens en anden gruppe ikke fik tilsvarende signaler. Deltagerne anvendte signalerne selvom fysisk styrke ikke direkte påvirkede de monetære udfald i spillet (dvs. hvem der vand og med hvilke omkostninger). Svagere mænd overlod spontant ressourcer til stærkere mænd for at undgå en konfliktoptrapning, og de gjorde det hurtigere desto større forskellen i styrke var. Hovedresultatet demonstrerer, hvordan (evolutionære relevante) signaler om kampevne fortsat påvirker

konfliktresolution i moderne mennesker. Fred i blandt mennesker er således baseret på meget grundliggende strategier for konfliktløsning, der også er til stede i gennem dyreverden. Dette kapitel er i skrivende stund inviteret til *revise* og *resubmission* ved tidsskriftet *Psychological Science*.

Det tredje kapitel, *The Role of Physical Attractiveness and Strength in the Resolution of Intrasexual Female Conflicts*, skal ses om et spejlbillede på kapitel 2 og er medforfattet af Michael Bang Petersen og Alexander Koch. Studiet er motiveret af tidligere forskning, der har vist at kvinder der også anså sig selv for at være mere fysisk attraktive også var mere tilbøjelige til at anvende vrede, føle sig mere berettiget og rapportere højere grad af succes i konflikter (Sell et al., 2009). Med andre ord er det at være fysisk attraktiv for en kvinde den direkte ækvivalent til overkropsstyrke for mænd. I dette kapitel rettede vi fokus mod kvindens attraktivitet. Vi anvendte det samme paradigme som i kapitel 2, men forventede at fysisk attraktivitet måtte være en betydningsfuld determinant i intraseksuelle konflikter. Vi målte både selvrapporteret attraktivitet og talje-hofte ratio, et mål der ofte anvendes som en markør for kvindelig attraktivitet (Sing, 1993, 1994; Sing et al., 2010, Streeter and McBurney, 2003). En af deltagergrupperne fik visuel information om deres modstander, og kunne dermed vurdere deres attraktivitet, mens den anden gruppe ikke modtog nogen information. Vores resultater viser at konfliktens længde er aftagende med forskellen i attraktivitet, og at denne effekt er drevet af gensidig vurdering og dermed ikke et resultat af selvopfattelser og interne processer. Da resultaterne antyder at mindre attraktive kvinder var hurtigere til at trække sig fra konflikten, ville man også forvente at mere attraktive kvinder skulle være mere tilbøjelige til at vinde, men resultaterne her er mere tvetydige. Samlet set tyder vores resultater på, at kvinder benytter visuelle signaler om hinandens attraktivitet, når de indgår i intraseksuel konkurrence.

Det fjerde, og sidste kapitel, *Institutional Norms of Fairness and Support for Taxation*, undersøger, hvordan individers opbakning til beskatning påvirkes af institutioner. Det argumenteres ofte for, at måden hvormed velfærdsstaten fungerer på, har indflydelse på borgernes villighed til at betale skat. Men hvordan institutionerne er udformet kan lige så vel være et resultat af borgernes præferencer, hvilket gør det svært at determinere den

basale sammenhæng mellem institutioner og præferencer. Gennem et laboratorieforsøg undersøger jeg påstanden om, at institutioners udformning påvirker den foretrukne skatterate. På baggrund af Rothstein (1998; 2001) undersøger jeg tre variationer i de institutionelle regler, der manipulerer individers opfattelse af *substantiv fairness* (at kun personer, der fortjener hjælp, modtager hjælp), *procedural fairness* (at hjælp tildeles gennem transparente regler), og *fair burden* (at omkostningerne af de distributive politikker er ligeligt fordelt mellem borgerne). Undersøgelsen viser, at både opfattelse af *substantial fairness* og *procedural fairness* påvirker deltagernes foretrukne skatterate. Dette understøtter idéen om, at hvordan velfærdsstaten fungerer, har betydning for borgenes villighed til at betale skat.

LITTERATURLISTE

Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange in Barkow J., Cosmides L., & Tooby J. (Ed.), *The adapted mind: Evolutionary psychology and the generation of culture (*pp. 206-221). New York: Oxford University Press

Pratto, F., Sidanius, J., Stallworth, L. M. & Malle, B.F. (2001). Social dominance orientation: A personality variable predicting social and political attitudes. In Michael A. Hogg, and Dominic Abrams, (Eds.). pp. 30-59. *Intergroup relations: Essential readings*. Philadelphia, PA: Psychology Press/Taylor & Francis

Rothstein, B. (1998). *Just institutions matter: the moral and political logic of the universal welfare state*. Cambridge University Press.

Rothstein, B. (2001). The universal welfare state as a social dilemma. *Rationality and society*, *13*(2), 213-233.

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.

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Singh, D. (1993). Adaptive significance of female physical attractiveness: role of waist-to-hip ratio. *Journal of personality and social psychology*, *65*(2), 293

Singh, D. (1994). Is thin really beautiful and good? Relationship between waist-to-hip ratio (WHR) and female attractiveness. *Personality and Individual Differences*, *16*(1), 123-132

Singh, D., Dixson, B. J., Jessop, T. S., Morgan, B., & Dixson, A. F. (2010). Cross-cultural consensus for waist–hip ratio and women's attractiveness.*Evolution and Human Behavior*, *31*(3), 176-181

Streeter, S. A., & McBurney, D. H. (2003). Waist–hip ratio and attractiveness: New evidence and a critique of "a critical test". *Evolution and Human Behavior*,*24*(2), 88-98

Chapter 1

Social Dominance Orientation and Conservatism Affect Assessment of Physical Formidability

Dan Nguyen
Aarhus University

Jim Sidanius
Harvard University

ABSTRACT[†]

Physical prowess is an important status indicator in dominance hierarchies. Decisions on engagement in hierarchy-disruptive behavior are affected by assessments of the costs and benefits of engaging in such challenging behaviors. We investigate whether individual assessment of formidability is conditioned by social dominance orientation (SDO) and general socio-political conservatism. In an online survey experiment we asked individuals (N = 269) to rate the physical strength of either a weak or a strong person. Our main findings are as follows. First, individuals' evaluation of physical strength is conditioned by conservatism and SDO. Second, this effect is only found for individuals exposed to the body silhouette of a strong male target person opposed to the exposure of a weaker male target person. Third, we find that our results are explained by

[†] **Acknowledgments:** The authors wish to thank Michael Bang Petersen, Alexander Koch, Lasse Laustsen, and Emma von Essen for valuable comments and the Harvard Digital Lab for Social Sciences for providing assistance on the study design.

the SDO-Dominance scale, the predictor of aggressive social dominance orientation. Our results suggest that individual differences in the perception of the formidability of male targets are associated with systematic individual differences in the desire for group-based social hierarchy and social order.

KEYWORDS
evolutionary psychology, social dominance orientation, conservatism, formidability, assessment

INTRODUCTION

The ability to assess physical strength (formidability) is important in the evaluation of interpersonal status (Sell et al., 2009). In particular, such an evolutionarily developed cognitive mechanism is crucial for making decisions about inter-group conflicts (Tooby & Cosmides, 2010). Nonetheless, substantial evidence exists of individual differences on preferences for inter-group conflict and group-based dominance (Duckitt & Sibley, 2010; Ho et al., 2012;, 2015; Pratto et al, 1994; Sidanius & Pratto, 1999). In this research article, we ask whether individual differences in assessment of physical threat are associated with political conservatism, which has been argued to reflect a more fundamental psychological phenomenon, namely negativity bias (Hibbing et al. 2014). We also explore the possibility that individual differences might be related to differences in preferences for inter-group hierarchies.

In a nutshell, we hypothesize an interaction exists between social dominance orientation (SDO) and political conservatism because conservatism entails that individuals with a preference for social hierarchies will tend to feel that these hierarchies are under threat. This is particularly true when one valuates the existence of inter-group hierarchies. Because formidability is an evolutionary understood indicator of disruptive potential, the individuals who combine high SDO and conservatism should be particularly sensitive to such threats. While the cognitive mechanisms for judging upper-body strength are activated when participants are explicitly asked to rate physical strength, we would also expect such mechanisms to be context dependent. Thus, people with high SDO levels and conservative socio-political ideology will increase their attention to the physical strength of others.

Consistent with recent developments within the social dominance theory literature and the bifurcation of SDO into two subdimensions, we argue that judgment of physical formidability will only be relevant when actors have a preference for group-based inequality. Our study presents findings of a systematic bias in people's assessment of physical strength. We show that judgment of physical strength is systematically affected by people's political

orientation such that evaluations are moderated by SDO (Pratto et al., 1994; Sidanius & Pratto, 2001).

CONCEPTS: FORMIDABILITY, SDO AND POLITICAL IDEOLOGY

Formidability, or fighting ability, is defined as the potential to inflict costs on others (Sell, Cosmides and Tooby, 2009). As such, formidability can be seen as the capacity to cause disorder to established hierarchies: it is an evolutionarily conditioned determinant of the individual's bargaining position (Tooby and Cosmides, 1988). Across non-human animals, evidence suggests that individuals' decisions whether or not to engage in physical conflict over resources – and, hence, attempt to challenge the relative status among its conspecifics – is based on various strategies that lead to accurate assessment of cues determinative of fighting ability (Archer, 2009; Arnott and Elwood, 2009). For humans, this also seems to be the case: pre-verbal infants and children expect larger individuals to prevail in conflicts (Pietraszewski & Shaw, 2015; Thomsen et al., 2011); stronger individuals are more likely to apply aggression and anger (Sell, Tooby & Cosmides, 2009); and when facing a physically superior opponent, individuals seem to withdraw more quickly in conflicts (Nguyen et al., in preparation).

The notion that different political ideologies reflect deeper distinct psychological responses is well-documented. For example, conservatives compared to liberals, gaze longer on threatening stimuli (Dodd et al., 2012), have stronger physiological reactions to threatening images or sounds (Oxley et al, 2008), and are more easily fixated on negative exposures (Carraro et al., 2011; McLean et al., 2014) Such persistent aversion to negative stimuli in political conservatives across many domains has been captured in the term 'negativity bias'. In the words of Hibbing, Smith and Alford (2014: 297), 'compared to liberals, conservatives tend to register greater physiological responses to such [negative] stimuli and also to devote more psychological resources to them.' Given that formidability is a measure of potential disorder, conservatives with their greater negativity bias should be more alert to cues about other individuals' formidability.

Conservatives also seem to differentiate themselves from liberals in terms of how they perceive social settings. The Dual Process Model of

Duckitt and Sibley (2010) argues that 1) conservatives tend to see outgroups as potentially much more harmful to the in-group stability than liberals, and 2) they are more competitive-driven in their pursuit of maintaining or establishing group superiority. The latter is also the psychological feature that links conservatism with Social Dominance Orientation. SDO assesses the degree to which people prefer hierarchical and unequal relations among groups, regardless of whose group is on the top or bottom (Pratto et al., 2001). SDO has been shown able to predict a wide array of social attitudes and group relevant behaviors associated with the intensity of group-based social inequality (Ho et al., 2012, Sidanius et al, in press). As such, it maps individuals' preferences in their pursuit of ways to sustain social hierarchies.

We believe that assessment of formidability is linked to the evolutionary nature of dominance in both intra-group status formation (von Rueden, Gurven and Kaplan, 2011) as well as inter-group conflicts (Thomsen, Green & Sidanius, 2008). Recently, the SDO scale has been suggested to contain two underlying sub-dimensions: SDO-Dominance (SDO-D) and SDO-Egalitarianism (SDO-E) (Ho et al., 2012; 2015). This need for further distinction is based on the dual nature of SDO capturing both group-based dominance and general anti-egalitarianism in the scale and was drawn to attention by Jost and Thompson (2000). They noted, that the initial SDO-scale consists of 16-items in which 8 statements directly elicits group-based dominance, aggression or control such as "In getting what your group wants, it is sometimes necessary to use force against other groups". The remaining eight questions however, all but one, includes the word "equal" or "equality" in statements like "It would be good if all groups could be equal" – these questions constitute the SDO-E scale.

Another important insight from that study, and particularly relevant for ours, is that whereas SDO-E correlated with political conservatism across two samples, the SDO-D scale did not (Jost and Thompson, 2010, p. 227). Whereas SDO-E is understood as a more non-violent opposition to group-based equality, SDO-D is defined as: "…support for group-based dominance hierarchies in which dominant groups actively oppress subordinate groups" (Ho et al., 2012, p. 585). We stress the words "actively oppress" since this must, at some level, involve decision-making and

evaluation of the cost-benefits in pursuing those activities. We argue, that SDO-D, compared to SDO-E, is a better indicator and more representative for the notion that high SDO individuals perceive the world as a "…competitive jungle in which might is right, the strong win, and the weak lose, as opposed to a place of cooperative harmony" (Duckitt & Sibley, 2010, p. 1869).

We argue that individual differences in threat perception may be reflected in biased estimates of physical strength in others. Specifically, we hypothesize that individuals with right-leaning ideologies and high SDO levels are likely to increase their psychological attention to the physical strength of others, especially the strength of other men. The logic is that a high SDO enhances conservatives' tendency to feel threatened.

As hypothesized and tested by Holbrook et al (in press), and contrary to intuition, conservative leaning individuals should envision potential enemies as less physically formidable. Their argument for this relies on the notion that conservatives also carry a predisposed preference for aggressive responses to potential threats hence eliminating the potential danger before it realizes itself. In turn, conservatives see the threats as less formidable. This is coined the "Gulliver" effect. Instead of asking participants to envision the strength of a person, we ask our participants to evaluate a visual stimulus in a direct encounter. We therefore hypothesize that the "Gulliver"-effect should only be found when one is exposed to a relatively strong individual because such an individual is a potentially greater physical threat than is a weaker individual. We hypothesize that individuals with high SDO through their competitive pursuit they possess in establishing group-dominance will be more likely to rate potentially threatening stimulus as being more formidable. As such we assert that the model presented by Holbrook et al. (in press) is underspecified by not taking the actual formidability of the envisioned target as well as the participants' level of SDO into account. Our analysis, thus, integrates the effects of conservatism and SDO across strong and weak targets, respectively. Finally, following our discussion of the distinction between SDO-E and SDO-D, we hypothesize that this effect should be particularly pronounced for SDO-D rather than for SDO-E.

In sum, this reasoning led us to the following three hypotheses:

H1: SDO and conservatism interact regarding the perceived physical formidability of target males such that subjects high on SDO and political conservatism tend to perceive others as more formidable

H2: The effect of SDO is driven by the SDO-D component rather than the SDO-E component of SDO. Hence, SDO-D and not SDO-E should be a significant predictor of rated strength.

H3: The interaction effect of SDO-D and conservatism on perceived formidability will manifest when people are confronted with more physically formidable male targets and not when one is confronted with less physically formidable targets.

METHODS

*Participants***.** We employed 150 participants for each of two conditions in a one-way ANOVA design. All participants were recruited from the Harvard Digital Lab for Social Sciences (DLABSS) subject pool. Based on the age distribution, our sample did not seem to consist purely of students (see Figure S1). Twenty-eight participants failed to complete the survey and another ten participants did not provide answers (or answered "Don't know") and are coded as missing values. This left us with a sample size of 269 participants, 52% of whom were female. Each participant received a lottery ticket for a monthly drawn gift card ($50) for Amazon administered by the DLABBS.

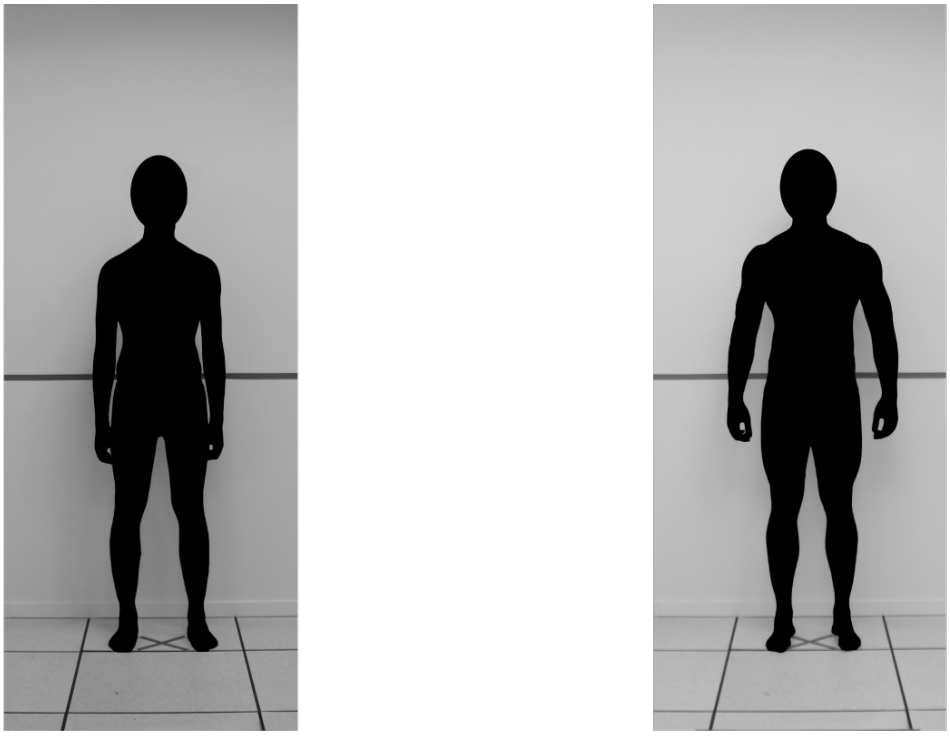
*Procedures***.** We examined the relationships between SDO, ideology and assessment of physical strength through an online survey in which participants were asked to rate the physical strength of a male target based on a silhouetted picture of his body (see Figure 1).

The silhouettes were selected from a picture pool collected during 2013 and 2014 (Nguyen et al., in preparation). The main advantage of using silhouettes is the masking of possible cues about race and body markings

(scars, tattoos etc.). Furthermore, people's estimates based on body silhouettes will on average correlate with actual physical strength (Nguyen et al., in preparation). Based on a standardized composite measure of physical strength (hand grip strength, flexed bicep circumference, chest strength, self-reported strength – see Sell et al. (2009)) we selected the most and least physically formidable male of the sample to be shown the survey.

Participants were asked to evaluate the physical strength of the body silhouette shown to them on a 7-point Likert scale (1 = very weak and 7 = very strong). They were randomly shown either a body silhouette of the weak male figure (Weak Condition) or the strong male figure (Strong Condition). Immediately after, the participants filled out a questionnaire.

Figure 1 **Body silhouettes**



On the left is the weak person and on the right, the strong person. Participants were asked to stand in a neutral position with arms along the side of their body. Feet lightly spread on the "X". Participants were shirtless but wore shorts (same model but in three different sizes). The line on the wall indicates 1m above the floor.

*Measures***.** Participants started the questionnaire by answering the 16-items *Social Dominance Orientation* (SDO_7) scale (Ho et al., 2016). All the items were 7-point Likert-scales. We employed both the total SDO scale ($M = 2.51;\ SD = 1.06;\ \alpha = 0.88;\ n = 269$) and the two subdimensions of SDO, the SDO-Dominance ($M = 2.40;\ SD = 1.10;\ \alpha = 0.79;\ n = 269$) and SDO-Egalitarianism ($M = 2.51;\ SD = 1.06;\ \alpha = 0.88;\ n = 269$). As expected, the various SDO-scales had relatively modest kurtosis of 2.72 were right-skewed (Skewness = 0.63). Furthermore, we collected participants' ideological orientation by asking them to indicate their political views on a scale from '1' (Extremely liberal) to '7' (Extremely conservative) ($M = 3.10;\ SD = 1.54;\ n = 260$). We employed additional variables to check for possible confounds. In particular, we controlled for age, fighting history and self-rated physical strength. Older persons might have more experience in assessing others, and hence better calibrated, and are perhaps also less motivated, to assess physical strength (Wilson and Daly, 1995). Fighting history, measured by the number of physical aggressive events a person has been experiencing over the last four years, possibly confounds evaluations of others' formidability since more aggression prone persons might have more to gain and experience in making more frequent and possibly accurate predictions (Sell, Tooby and Cosmides, 2009; Romero, Pham and Goetz, 2014). Finally, stronger persons might calibrate their personalities to be more extroverted (Tooby and Cosmides, 1990; Lukaszewski and Roney, 2011), be more likely to make spontaneous assessments of other men's formidability (Goetz, 2014), and support the use of military force in intergroup conflicts (Sell, Tooby and Cosmides, 2009).

Table 1 summarizes the descriptive statistics for the study. As a manipulation check, the strong male target was perceived as being significantly stronger than the weak male (two-sided two samples t-test in means:$t = -13.81;\ p < .001$). None of the other independent variables differed significantly across the two strength conditions, indicating that randomization was successful (two-side t-tests showed all the values on the variables did not differ, $p > .29$ as the lowest).

Table 1 **Descriptive statistics**

	Weak Condition				Strong Condition			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Formidability rating of target	3.95	0.95	2.00	7.00	5.37	0.82	3.00	7.00
SDO	2.55	1.14	1.00	5.38	2.49	0.99	1.00	5.25
SDO-D	2.45	1.15	1.00	5.25	2.36	1.06	1.00	6.50
SDO-E	2.65	1.30	1.00	7.00	2.62	1.14	1.00	6.25
Conservatism	3.11	1.64	1.00	6.00	3.11	1.46	1.00	6.00
Self-rated strength	53.14	21.66	6.00	94.00	51.90	22.44	2.00	97.00
Fighting history	1.14	0.54	1.00	4.00	1.21	0.65	1.00	4.00
Observations	132				137			

Table 2 **Correlation matrix of IVs**

	1	2	3	4	5	6	7
1. SDO	-						
2. SDO-D	.91***	-					
3. SDO-E	.92***	.67***	-				
4. Conservatism	.44***	.40***	.41***	-			
5. Self-rated strength	.12**	.10	.12**	-.01	-		
6. Age	.03	-.05	.10*	.31***	.05	-	
7. Fighting history	.00	.05	-.04	-.01	.11*	-.30***	-

* p<0.1, ** p<0.05, *** p<0.01. Listwise N = 259.

Table 2 shows the correlations between the independent variables. It is worth noting that although there was a significant correlation between conservatism and SDO, SDO-D and SDO-E, these correlations are of a moderate magnitude.

We then analyzed the data using multiple linear regression analysis. We attenuated the non-normality in the distributions of the SDO scales by use of log-transformation of these scales.

RESULTS

We start out by validating our treatment. We regress a treatment dummy (1 = strong picture treatment) on the perceived physical formidability in Model 1, Table 3 and we check for possible confounders by adding the following control variables to our models: age category, fighting history and self-rated strength[1]. It is confirmed that participants exposed to a stronger person rate that person as being stronger.

Next, we test the "Gulliver"-effect which states that participants holding more political conservative standpoints, in contrast to liberals, perceive ambiguous strangers as being potentially more dangerous, but "vaquishable through force" and hence less formidable (Holbrook et al., in press). As hypothesized, we did not find any effect of political orientation when the rated target was weak (Model 2), however, once the activation of a potential threat becomes salient, we find that being politically more conservative leads to a diminished formidability rating (Model 3). It is worth noting that individuals with conservative orientation shift their perception across formidability. In the weak condition, compared to liberals, ideologically conservative participants evaluate the targets to be stronger whereas in the strong condition, targets were seen as weaker.

Turning to Social Dominance Orientation, we hypothesized that conservative individuals who also are high on SDO will be more likely to evaluate threats as more dangerous. We explore this in Model 4. The three-way interaction between treatment, conservatism, and SDO is, however, only borderline significant ($p = .105$), but in the expected direction. It is worth remembering that this three-way interaction is estimated with the combined SDO scale that includes both SDO-D and SDO-E. We expect that the results will be very different when running the estimations for SDO-E and SDO-D, respectively. Testing hypothesis 2, we ran the same regression

[1] We also estimate a model consisting only of the control variables age, fighting history and self-reported strength to check for our randomization. The overall model is significant ($F(3, 263) = 2.63$, $p = 0.051$). Consistently with the other models reported in Table 3, age does come out as a significant and positive predictor (although small in magnitude, $b = 0.11$, $p < 0.01$). Taken together, our randomization was successful.

models but with SDO-D (Model 5) and SDO-E (Model 6), respectively. The three-way interaction is only significant (p = .065) and positive in the case of SDO-D (Model 5) but not for SDO-E (Model 6).

Table 3 Regression models

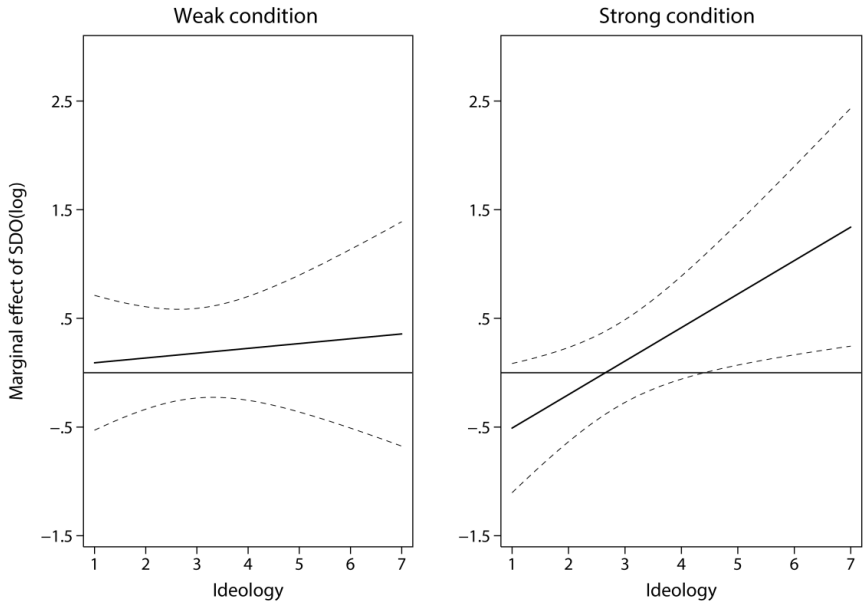
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Treatment	1.444***	1.444***	2.173***	2.868***	2.897***	2.545***
	(0.108)	(0.111)	(0.247)	(0.507)	(0.453)	(0.469)
Conservatism		0.113	0.113**	0.066	0.109	0.064
		(0.037)	(0.048)	(0.115)	(0.101)	(0.107)
Treatment*Conservatism			-0.234***	-0.496***	-0.499***	-0.389**
			(0.071)	(0.180)	(0.157)	(0.166)
SDO				0.011		
				(0.381)		
Treatment*SDO				-0.821		
				(0.583)		
Conservatism*SDO				0.033		
				(0.110)		
Treatment*Conservatism*SDO				0.284		
				(0.175)		
SDO-D					0.123	
					(0.361)	
Treatment*SDO-D					-0.945*	
					(0.550)	
Conservatism*SDO-D					-0.013	
					(0.102)	
Treatment*Conservatism*SDO-D					0.317*	
					(0.162)	
SDO-E						-0.054
						(0.333)
Treatment*SDO-E						-0.405
						(0.508)
Conservatism*SDO-E						0.041
						(0.097)
Treatment*Conservatism*SDO-E						0.156
						(0.153)
Age	0.096***	0.092***	0.098***	0.106***	0.108***	0.102***
	(0.032)	(0.038)	(0.034)	(0.034)	(0.035)	(0.034)
Fighting history	0.072	0.124	0.110	0.112	0.101	0.120
	(0.095)	(0.107)	(0.105)	(0.105)	(0.105)	(0.107)
Self-reported strength (0-100)	0.002	0.002	0.002	0.001	0.002	0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	3.209***	3.124***	2.794***	2.801***	2.696***	2.853***
	(0.271)	(0.285)	(0.298)	(0.399)	(0.380)	(0.384)
N	267	259	259	259	259	259
R^2	0.422	0.422	0.446	0.459	0.459	0.453

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

As highlighted by Brambor et al. (2005), when analyzing the interaction term of two continuous variables one must consider the conditional relationships between these variables. To do so, we considered the marginal effects of SDO under various values of conservatism on the rating of physical formidability. For illustrational purposes, we split now the sample into two subsamples by treatment (the regression models can be seen in the appendix).

Figure 2 illustrates the three-way interaction between real physical threat (i.e. strong condition), conservative ideology and increased preference for intergroup-hierarchies. We observe that there never is a statistically significant marginal effect of SDO in the weak treatment. Yet, for the strong treatment we in fact find that SDO has a significant effect for above-average levels of conservatism.

Figure 2 **The effect of SDO on the assessment of physical strength as a function of ideology.**

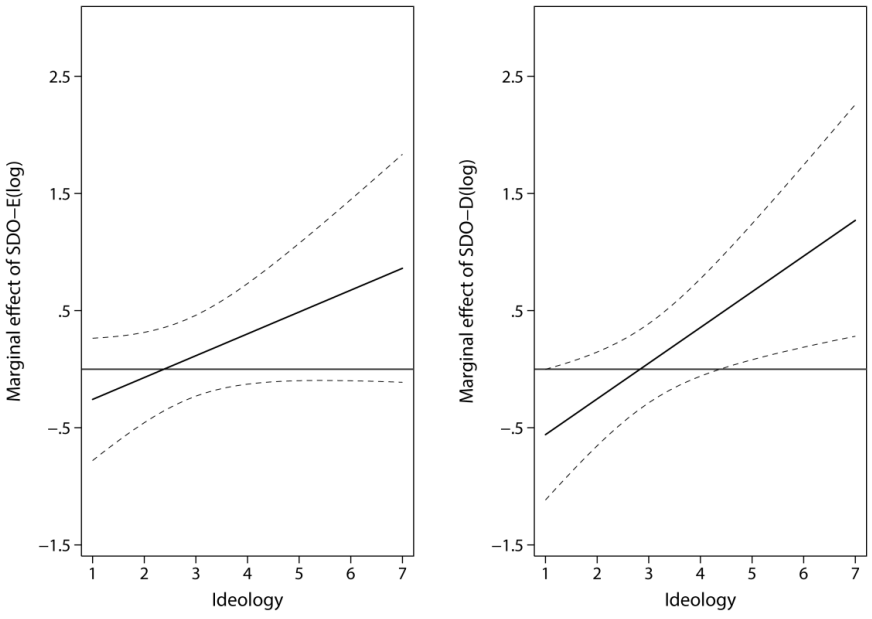


Results are shown for the weak condition (left side) and strong condition (right side), respectively. Solid lines represent the predicted effect sizes, and estimated confidence intervals at 95% of the given effect sizes are shown by the dashed lines. Effect sizes are interpreted as a 1%-increase on the SDO-scale on the average rating of physical strength on the body silhouettes. Marginal effects are estimated based on Model 2 and Model 4 in Table S1..

It follows from the arguments above that the effects found in Figure 2 should be strongest for the SDO-D sub-dimension. Furthermore, as shown in the left-hand side of Figure 3, in the strong condition, the interaction between SDO-E and conservatism was insignificant ($b = 0.186; p > .05$). In contrast, the relationship between SDO-D and assessed physical strength, however, was clearly moderated by conservatism ($b = 0.305; p < .01$); see the right-hand side in Figure 3. In the weak condition, neither the interaction between conservatism and SDO-D nor SDO-E had any significant effect (SDO-D*Conservatism: $b = 0.002; p = .982$; Model 1 in table S2, and SDO-E*Conservatism: $b = 0.034; p = .745$; Model 1 in table S3).

In sum, we found evidence that participants with a conservative ideology tend to rate the physical strength of other relatively high, but only (1) when confronted with a picture of a strong male and (2) when the participants had a high degree of SDO-Dominance.

Figure 3 **The effect of SDO-E (left panel) and SDO-D (right panel) on the assessment of physical strength as a function of ideology.**



Solid lines represent the predicted effect sizes, and estimated confidence intervals at 95% of the given effect sizes are shown by the dashed lines. Effect sizes are interpreted as a 1%-increase on the SDO-scale on the average rating of physical strength on the body silhouettes. Both models were estimated with control variables.

DISCUSSION

Research on human formidability suggests humans are quite adept at assessing the fighting ability of other humans (Sell et al., 2009). We tested if there is systematic individual variation in how humans make assessment of males' physical formidability. We hypothesized that such underlying differences would be driven by the interplay between SDO and political conservatism.

At the core of our argument, we assume that assessment of physical strength has an intergroup purpose (Lukaszewski et al., 2016), and perhaps also an intragroup purpose and as such should be more salient for persons who are more motivated to preserve intergroup-hierarchies and those who are also more likely to perceive threats negatively. Based on recent developments in the SDO-literature, we also hypothesized that because assessment of formidability facilitates decision-making at an individual level to actively defend or pursue status-upsetting activities, from an inter-group perspective, this relationship should depend on SDO-D and not SDO-E. SDO-D would be the main contributor explaining assessment of physical prowess due to its direct relationship to people with more support for violent and aggressive behaviour towards out-groups. Perceived physical formidability was mainly found when this group-based dominance orientation goes together with the negativity bias of individuals with conservative ideology.

We can ask why individuals who are high on SDO and conservatism are likely to display heightened perceived formidability. Error Management Theory (EMT) suggests that human cognition is biased so as to effectively minimize the cost of inferential errors (Haselton and Buss, 2000; Haselton and Nettle, 2006). First, the perception of an outgroup-member will possibly lead to an increased sensitivity to physical threats. Therefore perceivers of threats will increase their accuracy in estimating the potential physical threat when exposed to cues of out-group members' formidability. Second, and in accordance with the EMT, perceivers might constantly over-infer outgroup-members' formidability in a "better safe than sorry" way. We did not compare the rated strength with actual strength and, hence we cannot directly speak to accuracy. Our results, however, do seem in line with a

"better safe than sorry" type of explanation for individuals whose objective functions might be negatively influenced by risky events or threats and positively by higher degrees of inter-group status. One fruitful avenue for future work would be to include accuracy measures.

REFERENCES

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.

Arnott, G., & Elwood, R. W. (2009). Assessment of fighting ability in animal contests. *Animal Behaviour*, *77*(5), 991-1004.

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political analysis*, *14*(1), 63-82.

Carraro, L., Castelli, L., & Macchiella, C. (2011). The automatic conservative: Ideology-based attentional asymmetries in the processing of valenced information. *PLoS One*, *6*(11), e26456.

Dodd, M. D., Balzer, A., Jacobs, C. M., Gruszczynski, M. W., Smith, K. B., & Hibbing, J. R. (2012). The political left rolls with the good and the political right confronts the bad: connecting physiology and cognition to preferences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1589), 640-649.

Duckitt, J., & Sibley, C. G. (2010). Personality, Ideology, Prejudice, and Politics: A Dual-Process Motivational Model. *Journal of personality*, *78*(6), 1861-1894.

Goetz, A. (2015). Individual Differences in Men's Spontaneous Assessments of Formidability: Support for Facultative Calibration. Presented at the Annual Meeting of the Human Behavior and Evolution Association, 2015.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: a new perspective on biases in cross-sex mind reading. *Journal of personality and social psychology*, *78*(1), 81.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. Personality and Social Psychology Review, 10, 47-66.

Hibbing, J. R., Smith, K. B., & Alford, J. R. (2014). Differences in negativity bias underlie variations in political ideology. *Behavioral and Brain Sciences*,*37*(03), 297-307.

Ho, A. K., Sidanius, J., Pratto, F., Levin, S., Thomsen, L., Kteily, N., & Sheehy-Skeffington, J. (2012). Social dominance orientation revisiting the structure and function of a variable predicting social

and political attitudes.*Personality and Social Psychology Bulletin*, *38*(5), 583-606.

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K., ... & Stewart, A. (2015). The nature of social dominance orientation: Introducing the Social Dominance Orientation-7 Scale. *Journal of Personality and Social Psychology*.

Holbrook, C., López-Rodríguez, L., Fessler, D.M.T., Vázquez, A., Gómez. A. (in press). Gulliver's politics: Conservatives envision potential enemies as readily vanquished and physically small. *Social Psychological and Personality Science.*

Jost, J. T., & Thompson, E. P. (2000). Group-based dominance and opposition to equality as independent predictors of self-esteem, ethnocentrism, and social policy attitudes among African Americans and European Americans. *Journal of Experimental Social Psychology*, *36*(3), 209-232.

Lee, I. C., Pratto, F., & Johnson, B. T. (2011). Intergroup consensus/disagreement in support of group-based hierarchy: an examination of socio-structural and psycho-cultural factors. *Psychological bulletin*, *137*(6), 1029.

Lukaszewski, A. W., Simmons, Z. L., Anderson, C., & Roney, J. R. (2016). The role of physical formidability in human social status allocation. *Journal of Personality and Social Psychology, 110, 385-406*

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.

McLean, S. P., Garza, J. P., Wiebe, S. A., Dodd, M. D., Smith, K. B., Hibbing, J. R., & Espy, K. A. (2014). Applying the flanker task to political psychology: A research note. *Political Psychology*, *35*(6), 831-840.

Nguyen, D., Petersen, M.B., Nafziger, J., & Koch, A. (2016). Upper-body strength and conflict resolution in human males. Working paper.

Oxley, D. R., Smith, K. B., Alford, J. R., Hibbing, M. V., Miller, J. L., Scalora, M., Hatemi, P. K. & Hibbing, J. R. (2008). Political attitudes vary with physiological traits. *Science*, *321*(5896), 1667-1670.

Pietraszewski, D., & Shaw, A. (2015). Not by strength alone: Children's conflict expectations follow the logic of the asymmetric war of attrition**.** *Human Nature, 26* (1).

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of personality and social psychology*, *67*(4), 741.

Pratto, F., Sidanius, J., Stallworth, L. M. & Malle, B.F. (2001). Social dominance orientation: A personality variable predicting social and political attitudes. In Michael A. Hogg, and Dominic Abrams, (Eds.). pp. 30-59. *Intergroup relations: Essential readings*. Philadelphia, PA: Psychology Press/Taylor & Francis

Romero, G. A., Pham, M. N., & Goetz, A. T. (2014). The implicit rules of combat. *Human Nature*, *25*(4), 496-516.

Sell, A., Hone, L. S., & Pound, N. (2012). The importance of physical strength to human males. *Human Nature*, *23*(1), 30-44.

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.

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society of London B: Biological Sciences*, *276*(1656), 575-584.

Sidanius, J**.** & Pratto, F. (1999). *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression.* New York: Cambridge University Press.Thomsen, L., Green, E. G., & Sidanius, J. (2008). We will hunt them down: How social dominance orientation and right-wing authoritarianism fuel ethnic persecution of immigrants in fundamentally different ways. *Journal of Experimental Social Psychology*, *44*(6), 1455-1464.

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. Science, 331, 477–480.

Tooby, J., & Cosmides, L. (1988). The evolution of war and its cognitive foundations. *Institute for evolutionary studies technical report*, *88*(1), 1-15.Tooby, J., & Cosmides, L. (2010). Groups in mind: The coalitional roots of war and morality. *Human morality and sociality: Evolutionary and comparative perspectives*, 91-234.

von Rueden, C., Gurven, M., & Kaplan, H. (2011). Why do men seek status? Fitness payoffs to dominance and prestige. *Proceedings of the Royal Society of London B: Biological Sciences*, 278, 2223-2232.

SUPPLEMENTARY MATERIAL

Figure S1 **Histogram of age**

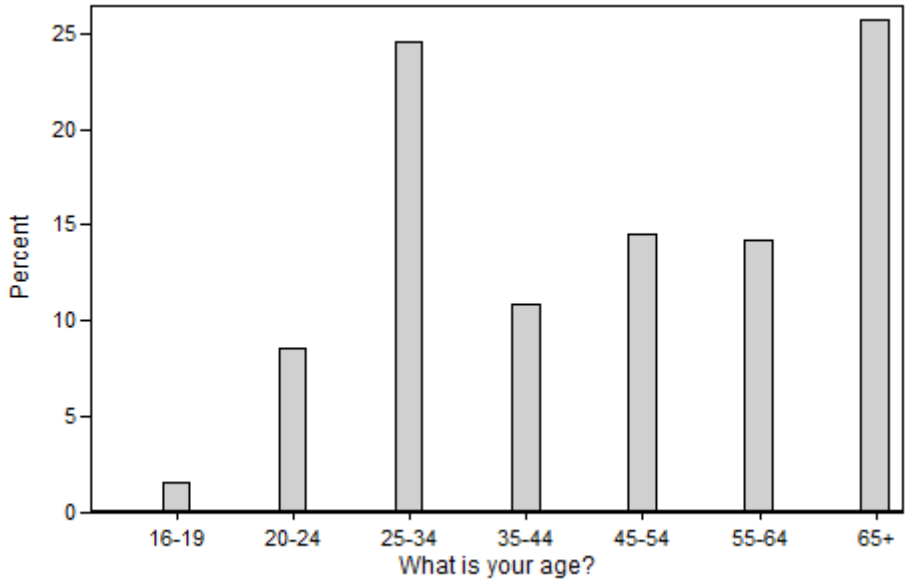


Table S1 **Regression models for the ratings of formidability (DV)**

	Weak condition		Strong condition	
	Model 1	Model 2	Model 3	Model 4
SDO(log)	-0.013	0.047	-0.680*	-0.818**
	(0.411)	(0.410)	(0.396)	(0.405)
Ideology	0.103	0.033	-0.322**	-0.404***
	(0.123)	(0.125)	(0.126)	(0.131)
SDO*Ideology	0.032	0.044	0.260**	0.308**
	(0.119)	(0.119)	(0.123)	(0.125)
Age		0.150***		0.077*
		(0.052)		(0.045)
Self-rated strength		-0.004		0.006*
		(0.004)		(0.003)
Fighting		0.082		0.105
		(0.178)		(0.126)
Constant	3.519***	2.924***	6.211***	5.561***
	(0.345)	(0.492)	(0.348)	(0.487)
N	129	129	131	130
R^2	0.061	0.123	0.052	0.109

Beta coefficients are reported. Standard errors in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table S2 **Regression models for the ratings of formidability (DV)**

	Weak condition		Strong condition	
	Model 1	Model 2	Model 3	Model 4
SDO-D (log)	0.015	0.178	-0.690*	-0.864**
	(0.389)	(0.391)	(0.375)	(0.380)
Ideology	0.138	0.081	-0.286***	-0.372***
	(0.108)	(0.109)	(0.109)	(0.114)
SDO-D*Ideology	0.002	-0.007	0.246**	0.305***
	(0.110)	(0.110)	(0.114)	(0.116)
Age		0.153***		0.078*
		(0.053)		(0.045)
Self-rated strength		-0.004		0.007**
		(0.004)		(0.003)
Fighting		0.070		0.095
		(0.178)		(0.125)
Constant	3.474***	2.808***	6.157***	5.500***
	(0.306)	(0.481)	(0.314)	(0.451)
N	129	129	131	130
R^2	0.059	0.122	0.053	0.115

Beta coefficients are reported. Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table S3 Regression models for the ratings of formidability (DV)

	Weak condition		Strong condition	
	Model 1	Model 2	Model 3	Model 4
SDO-E (log)	-0.026	-0.054	-0.407	-0.444
	(0.359)	(0.357)	(0.347)	(0.354)
Ideology	0.101	0.033	-0.246**	-0.298**
	(0.114)	(0.116)	(0.118)	(0.122)
SDO-E*Ideology	0.034	0.055	0.167	0.186*
	(0.104)	(0.104)	(0.109)	(0.110)
Age		0.147***		0.071
		(0.052)		(0.045)
Self-rated strength		-0.004		0.006*
		(0.004)		(0.003)
Fighting		0.089		0.115
		(0.178)		(0.129)
Constant	3.528***	2.982***	5.998***	5.291***
	(0.323)	(0.475)	(0.322)	(0.480)
N	129	129	131	130
R^2	0.061	0.122	0.037	0.087

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Chapter 2

Upper-body Strength and Conflict Resolution in Human Males

REVISE AND RESUBMIT AT PSYCHOLOGICAL SCIENCE

Dan Nguyen
Aarhus University

Michael Bang Petersen
Aarhus University

Julia Nafziger
Aarhus University

Alexander Koch
Aarhus University

ABSTRACT[†]

Conflict is universal across all living species. Among non-human animals, a key strategy is to resolve conflicts without fighting by merely assessing relative fighting ability. We demonstrate the existence of the same system for conflict resolution in humans by providing experimental evidence that human male contestants spontaneously coordinate conflict behavior on the basis of differences in upper-body strength. We do so by applying a non-physical, anonymous, economic game - the war-of-attrition -

[†] **Acknowledgments:** We thank J. Midtgaard, M. Skov, H. Thyssen and S. Knudsen for excellent research assistance. We also thank H. Schumacher, S. Hirth, E. von Essen, L. Gerhards, Christian Ellermann-Aarslev, Sarah Cotterill, Robin Bergh, participants at the Max Planck Institute for Human Development, participants at the Sidanius Lab at Harvard, and at the Interacting Minds Seminars at Aarhus University for valuable comments.

in which contestants compete by means of perseverance to win a monetary prize. Though strength differences are not relevant for payoffs in this game, weaker males spontaneously cede resources to stronger males to avoid conflict escalation and do so quicker if strength differences are larger. The results show that the complex conflict-resolution strategies of humans are founded on more rudimentary strategies, present throughout the animal world.

KEYWORDS
evolutionary psychology, decision making, physical appearance, violence

INTRODUCTION

Conflict pervades the human condition: nations go to war, co-workers compete for promotions, and neighbors quarrel over fences. A prerequisite for any well-functioning social group consequently is the existence of strategies for conflict resolution among its members. Conflict resolution has accordingly constituted a focal object of study across the social sciences, which have identified a number of crucial resolution strategies facilitated by distinctly human features such as verbal communication (McCullough et al., 1997), culture (Ross, 1993), and institutions (Llewellyn and Howse, 1999).

All living organisms face conflicts, and strikingly similar conflict-resolution strategies exist throughout the animal world (De Waal, 2000). One of the best-validated models of general animal behavior - the Asymmetric War of Attrition (AWA) - captures how animals in competition over resources coordinate to avoid costly fighting (Maynard Smith and Price, 1973; Hammerstein and Parker, 1982; Maynard Smith, 1974; Maynard Smith and Parker, 1976). Because a number of easy-to-determine physical differences between animals are causally related to their fighting ability – e.g. general size of body, anglers, or horns (Arnott and Elwood, 2009) - mutual assessment of these features allows contestants to forecast who would prevail in a physical fight and coordinate their behavior accordingly without resorting to physical aggression (Parker, 1974). The AWA-model has been validated in species as diverse as ants, flies, fish, spiders, butterflies, crickets, chameleons, deer, elephants, and many others (Briffa and Hardy, 2013). Yet, no experimental evidence yet exists for this model in higher primates including humans.

Drawing on these observations of conflict resolution among animals we investigate the existence of corresponding strategies in human males, the more aggressive and conflict-prone sex (Daly and Wilson, 1988; Sell et al., 2012). If selection pressures underlying conflict resolution in the animal world have similarly acted on humans, human conflict resolution should also be shaped by individual differences associated with fighting ability as predicted by the AWA-model. Empirically, we focus on upper-body strength, which has been reliably associated with fighting ability throughout human evolutionary history (Puts, 2010; Sell et al., 2009a). We predict that

humans resolve conflicts in a coordinated fashion under the implicit cognitive assumption that relative upper-body strength is decisive for the outcome, as was the case over predominant periods of human evolution. If valid, this suggests that the unique and highly complex conflict-resolution strategies of humans are built on a basis of more rudimentary strategies, present throughout the animal world.

The mathematical structure of the AWA-model has previously been applied to human conflicts such as competition between firms, military arms races, and strikes (see Hörisch and Kirchkamp, 2010). Moreover, previous studies provide tentative evidence for a role of upper-body strength in human conflict behavior: humans are highly apt at judging males' upper-body strength (Sell et al., 2009a; Sell et al., 2009b; Sell et al., 2010); preverbal infants and children use cues of fighting ability, such as size differences, to predict conflict outcomes (Thomsen et al., 2011; Pietraszweski and Shaw, 2015); computer gamers are sensitive to relative fighting abilities of the avatars they operate (DeScioli and Wilson, 2011); physical strength in males correlates with self-reported measures of both aggressiveness and conflict success (Sell et al., 2009a, Archer and Thanzami, 2007; Price et al., 2012). While suggestive, these later studies have focused on the effect of *absolute* differences in upper-body strength on (self-reported) psychological *traits*. The AWA-model, in contrast, predicts that the psychological *states* involved in conflict behavior are flexibly modulated by the *relative* strength of contestants. In this way, the previous literature does not test whether the kernel of the AWA-model of conflict resolution among animals applies to humans: Are humans influenced by cues to relative upper-body strength such that actual conflicts are intuitively resolved in favor of the contestant most likely to prevail if the conflict escalated into a physical confrontation?

We investigate this possibility in an experiment where anonymous participants interact in a non-physical, economic conflict game - the war-of-attrition. In this game, two contestants compete over a prize. The winner is the opponent who persists longer and persisting comes with a cost that rises with contest duration. From an *objective* perspective, experimental contests are symmetric in that the induced cost of persisting and the benefit of

winning are the same for both contestants and, hence, independent of the contestants' physical strength. The experimental contest, however, is set up to investigate if participants operate under a *subjective* assumption that greater relative strength translates into an advantage in conflict situations. Hence, in one of two experimental treatments (named the *Mutual Assessment Treatment*), the participants were given an opportunity to mutually assess physical differences before they compete. If participants intuitively assume that greater relative strength translates into an advantage in conflict situations, the weaker participants should be most likely to withdraw and they should be quicker to withdraw the easier differences in strength are detected, i.e., the larger the strength asymmetry (Parker, 1974). In the other experimental treatment (the *Self-Assessment Treatment*), participants competed without prior assessment of mutual physical differences. In this treatment, participants only have information available about their own physical strength and, hence, cannot coordinate on the basis of relative strength. If the mind is designed to utilize estimates of relative formidability, strength should matter much less (and, potentially, not at all) in this treatment relative to the *Mutual Assessment Treatment*.

PRESTUDY: CREATING STIMULUS MATERIAL

To isolate the effect of strength, we were required to develop stimulus materials for the Mutual Assessment Treatment that reliably transfer information about the strength of an individual while retaining anonymity: the picture of a body silhouette (Fig. 1). The stimulus was selected based on a pre-study.

Methods

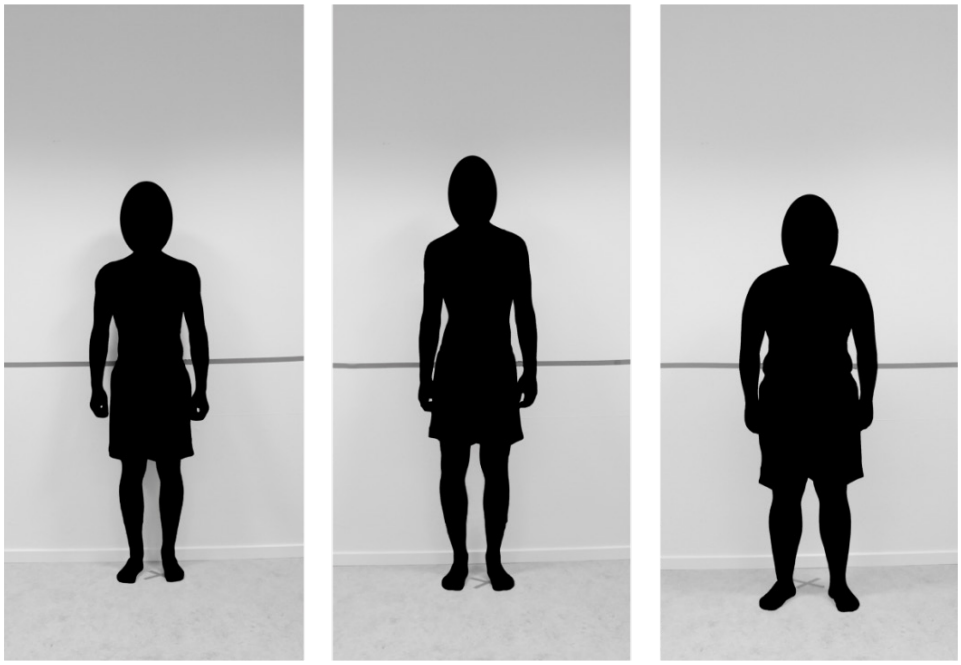
Participants. Fifty male participants (age: M = 22.84, SD = 1.81) were recruited around the campus area of Aarhus University. Upon consent, they were photographed and had their physical measures taken while also completing a questionnaire. Description of the measurements and procedures can be found in the Supplementary Online Materials. Photographs were then edited using Adobe Photoshop CS6 and through Qualtrics 400 US male raters (age*:* M = 26.99, SD = 6.27) were recruited to

evaluate the various body pictures on both physical strength and physical attractiveness on a 7-point scale (1 = Very weak/homely and 7 = Very strong/handsome) (see Supplementary Material for Methods). We obtained a total of 5000 unique ratings of physical strength on our 50 body silhouette pictures. Although a body silhouette picture may convey additional information, the purpose here is solely to investigate whether participants are able to reliably predict the physical strength of the presented body silhouette pictures. We assessed rater accuracy following the procedure of Sell et al., 2009b.

Results

Averaging the ratings of a particular silhouette across the raters provides a correlation between the rated and the actual physical strength of $r = 0.55$ ($p < .001$, $n = 50$). Hence male strangers' average rating of strength from body silhouettes of other men strongly correlated with the actual strength of those shown. To determine the average individual accuracy, we estimated an OLS regression model with the actual strength as the dependent variable and all the ratings from each rater as the independent variable. We applied clustered robust standard errors at rater level (200 clusters) to account for the repeated interactions (rating multiple pictures). We obtained a significant average individual accuracy of $\beta = 0.11$ ($p < .001$, $n = 5000$). Thus, accurate assessment does not only occur when data are aggregated across raters, but can also be found at an individual level.

Figure 1 **Examples of body silhouettes used in the Mutual Assessment**



MAIN STUDY: STRENGTH AND CONFLICT-RESOLUTION IN THE WAR-OF-ATTRITION GAME

On the basis of the findings of the pre-study, we conducted the main study using the war-of-attrition game. The main study was designed to test the following set of hypotheses: To the extent that resource conflict implicitly activates strength-based strategies for conflict-resolution in the contestants, a greater - and, hence, more easily detectable - asymmetry in upper-body strength between them should (i) increase the probability that the stronger contestant prevails and (ii) decrease the duration of the contest under conditions of mutual assessment in the war-of-attrition game. Under conditions of pure self-assessment (i.e., where no information about the opponent is available) these effects should be weaker or non-existing. In addition, we performed a number of auxiliary analyses serving as robustness tests of these main hypotheses including identifying the specific physical cues utilized for coordination.

Methods

Participants. Sixty-six male participants (age*:* M = 22.34, SD = 2.56) were recruited to "an economic experiment" at the Cognition and Behavior lab at Aarhus University. Each participant was randomly assigned into one of the two treatment groups: the *Mutual Assessment Treatment* (n = 42) and the *Self-Assessment Treatment* (n = 24). We oversampled the main treatment, the Mutual Assessment Treatment. We had no a priori evidence on likely effect sizes to formally determine the optimal sample size. For that reason we aimed for at least 100 contests in each treatment and oversampled the main treatment to increase power to detect the main effects in a regression framework. Table 1 provides the descriptive statistics for the individual characteristics of participants in our main study. The most important thing to notice is that randomization into treatments was successful in that none of the individual characteristics differ significantly across the two treatment groups.

Table 1 **Descriptive statistics for the main study (non-standardized)**

	Mean [95% confidence interval]		**Min**		**Max**	
	MAT *n=42*	SAT *n=24*	MAT	SAT	MAT	SAT
Age (years)	22.83 [20.61 ; 22.38]	21.5 [21.99 ; 23.67]	19	19	30	27
Height (cm)	184.19 [182.55 ; 185.82]	182.33 [179.22 ; 185.44]	172	169	200	196
Weight (kg)	79.92 [76.48 ; 83.35]	77.23 [73.07 ; 81.38]	63.3	57.1	103.2	97.10
BMI	23.57 [22.57 ; 24.57]	23.23 [22.06 ; 24.40]	17.9	18.4	31.1	29.7
Grip strength (kg/F)	49.6 [47.05 ; 52.13]	52.42 [47.57 ; 57.25]	36	34	77	75
Chest strength (kg/F)	50.79 [46.35 ; 55.20]	53.04 [48.82 ; 57.25]	20	35	90	70
Flexed biceps (cm)	35.5 [32.63 ; 34.36]	33.43 [32.35 ; 34.51]	27	27.5	40	39.5
Self-reported strength (0-100)	45.57 [38.26 ; 52.87]	55.88 [46.88 ; 64.86]	2	10	95	90
Risk tolerance (0-10)	6.47 [5.82 ; 7.12]	6.20 [5.08 ; 7.33]	2	0	10	10

MAT: Mutual Assessment Treatment, SAT: Self-Assessment Treatment.

Experimental Procedure. The experiment consisted of two parts. In the first part, participants had their pictures taken using the same procedure as

in stage 1 of the pre-study. However, strength measures were postponed until the end of the second part to avoid experimenter demand effects. The photographs of the participants were then edited to only reveal the outline of their body silhouette (detailed description and procedures are available online in the Supplementary Materials). In the second part, which took place within six weeks after the first part, participants were invited to sessions in an experimental lab where they completed several rounds of a two-person war-of-attrition game programmed with the z-Tree experimental software (Fishbacher, 2007). The instructions can be found in the Appendix. Absolute stranger-matching was applied, that is, each participant played against the same opponent only once, in a random order. Each session consisted of six to twelve participants.

In our main treatment, the *Mutual Assessment Treatment* (205 contests), participants saw a picture of the body silhouette of their opponent before playing a round of the war-of-attrition game. In this treatment, all rounds began with participants being instructed to look-up the body silhouette of their otherwise anonymous opponent in a booklet of pictures. The *Self-Assessment Treatment* (106 contests) mirrors the design of the *Mutual Assessment Treatment*, with the exception that there are no booklets with pictures and participants do not get to see pictures of the body silhouette of their opponent.

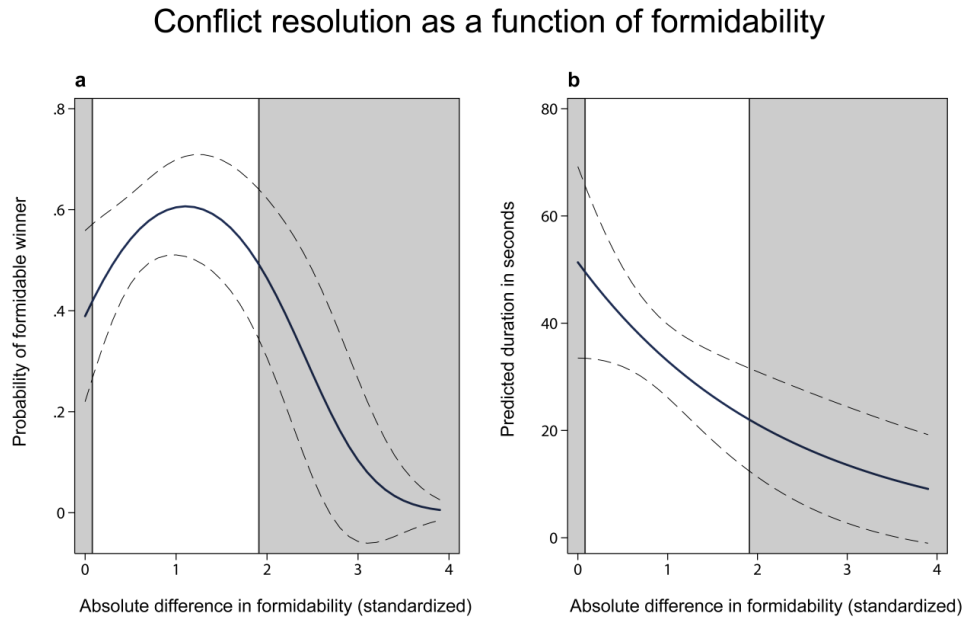
Independently of the treatment, every game started with each participant having 225 Experimental Currency Units (1 ECU=0.35 Danish kroners ≈ 5 US cents) from which 1 ECU got deducted for each second both participants continued the game.[*] By pressing a button, each participant could withdraw from the game at any time with a payoff of [225 – duration of contest in seconds] ECU for him and would leave the other contestant, who would receive a prize of 100 ECU for persisting longest, with a payoff of [100 + 225 – duration of contest in seconds] ECU. In case no one withdrew before

[*] Independently of treatment, players were instructed to press an "OK"-button when ready. When all players had done so, the actual game screen was shown and a 10-seconds countdown started. After the countdown, the "STOP"-button would turn from grey to red indicating that the round had begun. This feature allowed participants to press the button exactly at the start of the game if they so wished.

225 seconds elapsed the prize was split, leaving both contestants with a payoff of 50 ECU. After each game, a payoff-screen showed a participant the duration of the contests and whether the participant won or lost (i.e., withdrew from) the contest. The unit of analysis is a single round of the war-of-attrition game between two participants (N = 311). To assess strength, participants were brought to an adjacent room in order to record physical measures as those described in the pre-study design (see Table 1) and then received their payment in cash. Similar to Sell et al. (2009a) we combined standardized scores for chest and handgrip strength, bicep circumference, and a self-reported measure of physical strength into a composite measure of physical strength (Cronbach's $\alpha = .79$). In addition, we obtained a number of self-reported measures. Immediately after the games, participants completed a brief questionnaire on basic background data and on their risk tolerance using a validated (Dohmen et al., 2011) survey measure used to reliably predict actual behavior in lotteries: "*How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'unwilling to take risks' and the value 10 means: 'fully prepared to take risks'.* It appears plausible that risk preferences affect individuals' competitive decisions (see discussion on risk preferences in Borghans et al., 2006).

Results

Figure 2 **The effect of differences in strength on conflict outcomes in the Mutual Assessment Treatment.**



An increased absolute difference in strength between the contestants (**a**) increases the predicted probability that the winner is the stronger contestant (Table S1, model 2), (**b**) decreases the predicted duration of the contest (Table S2, model 1). The dashed lines show 95% confidence intervals. The white area shows the 5th to 95th percentile range in the data.

We start out by testing whether winners are more formidable in the *Mutual Assessment Treatment*. That is, does being the strongest person in a pair increase the probability of winning the contest? If it does, we should expect the probability of winning to rise if differences in strength become more apparent, i.e. if there are greater asymmetries between contestants. A logistic regression of the absolute difference in strength and its quadratic term on the binary outcome variable, whether the strongest in the pair won the game, reveals a statistically significant concave relationship – first increasing and then decreasing - of relative strength on the probability of the stronger contestant winning (joint test of linear and quadratic terms, $\chi^2(2) = 8.26$, $p = .016$), as illustrated in Fig. 2a. All estimated models include period and session fixed effects and apply robust standard errors. That is, as the absolute difference in formidability between two contestants increased, it

was more likely that the most formidable individual of a pair won. Even small differences in upper-body strength were enough to tilt the contest in favor of the stronger opponent (Fig. 2a) confirming the first hypothesis.

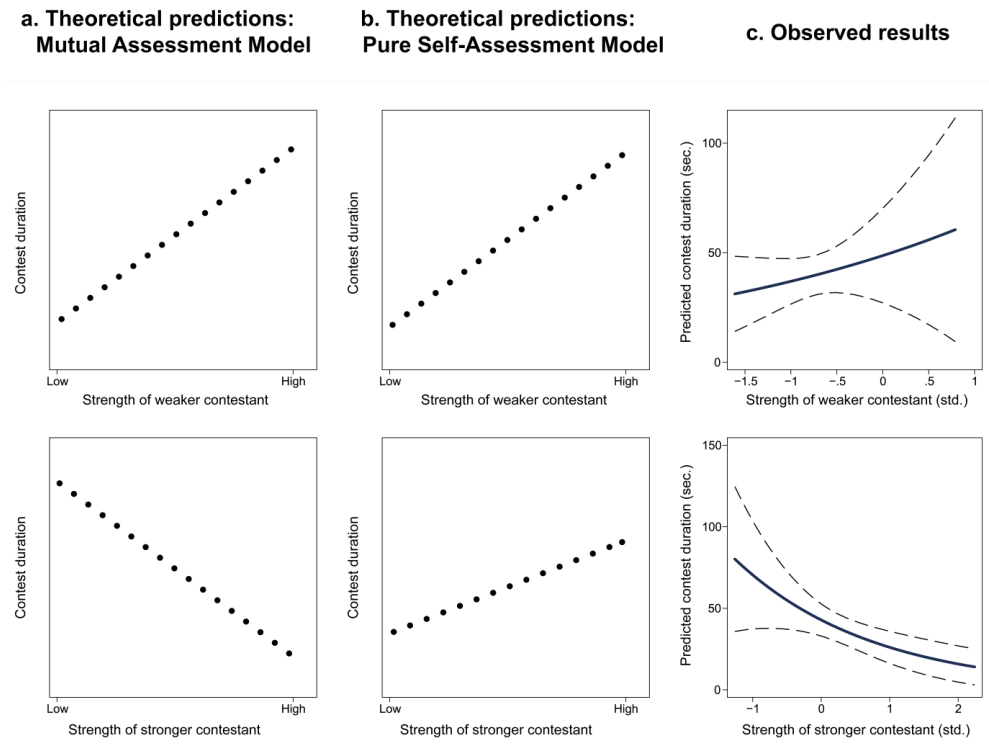
The first set of results shows that conflicts in the Mutual Assessment Treatment are resolved in the favor of the stronger subject. The next question we address relates to the duration of contests: are conflicts resolved more quickly when the difference in upper-body strength is larger and easier to determine? To test whether contest duration is inversely related to the difference in upper-body strength of the participants, we employ negative binomial regressions. This empirical model is well-suited to handle both the count structure of our data (Hilbe, 2011) – contest duration is measured in seconds - and the exponentially declining distribution of contest durations predicted by models of the war of attrition (Hammerstein and Parker, 1982; Bishop and Cannings, 1978; see Fudenberg & Tirole, 1991, pp. 191-121). A regression of the absolute difference in upper-body strength on contest duration in the Mutual Assessment Treatment (duration: M = 39.01, SD = 61.14) reveals a significant negative effect of relative strength on contest duration (β = -0.93, $p < .0031$). Contrary to the results for the winner of the contest, relative strength had a linear effect on contest duration. Fig. 2b illustrates how one standard deviation in relative strength reduces the expected contest duration by around 60%.

A key theoretical concern is whether these results really reflect coordination behavior. Are subjects mutually computing relative formidability in order to coordinate behavior in a manner consistent with the AWA model? Or do the results simply reflect a combination of independent effects that do not rely on relative strength assessment? Specifically, from the literature on animal contests we know that a negative relation between asymmetry in strength and contest duration may also emerge if the endurance of a contestant increases with his own strength and cues about the strength of the rival play no role (Taylor and Elwood, 2003). We address this issue in two ways.

First, we apply a method developed for studies of animal contests to distinguish between the two above explanations (Taylor and Elwood, 2003). If contestants mutually assess their relative strength and modulate their

decisions accordingly (referred to as the mutual assessment model), contest duration increases (decreases) with the weaker (stronger) rival's strength because this reduces (increases) the asymmetry between contestants (Fig. 3a). If contestants' endurance depends only on own strength (referred to as the pure self-assessment model), contest duration increases with both the weaker contestant's strength and, to a lesser degree (because the weaker contestant is likely to give up first), the stronger contestant's strength (Fig. 3b). The observed empirical results are displayed in Fig 3c. They reject the self-assessment model in favor of the mutual assessment model.

Figure 3 **Evidence for mutual assessment**



(a) The Mutual Assessment model predicts that contest duration is determined by the difference in strength, which increases (decreases) with the strength of the stronger (weaker) contestant. The dotted lines indicate the theoretical directions. **(b)** The Pure Self-Assessment Model predicts that contest duration is determined by a contestant's own strength. The key difference in predictions is the effect that the strength of the stronger contestant has. **(c)** The predicted duration of the contests derived from the negative binomial regressions for the Mutual Assessment Treatment (Table S2, model 4). The top (bottom) panel holds the strength of the stronger (weaker) contestant fixed at the average observed value. The dashed lines show 95% confidence intervals.

Second, we ran an additional treatment identical to our main treatment except that contestants saw no pictures of their opponent: the *Self-Assessment Treatment* (106 contests). If contest behavior is determined by processes within each participant, qualitatively similar data patterns as in our main treatment should emerge. In contrast, if the above results reflect behavior that relies on simultaneous information about both the self and the opponent, then the effects should disappear in the Self-Assessment Treatment, where no information on the strength of the opponent is available. Consistent with the latter hypothesis, analyses of the Self-

Assessment Treatment show (i) that the stronger contestant in a pair was not most likely to win (Table S3); (ii) that the duration of a contest in the Self-Assessment Treatment (duration: M = 27.23, SD = 48.85) was neither significantly related to the difference in strength between the contestants nor the strength of the weaker contestant within each pair (Table S4). The strength-related coefficients also differed significantly from those under the Mutual-Assessment Treatment. Pooling the data across the Self-Assessment and Mutual Assessment Treatments and interacting the absolute difference in strength and its square with treatment dummies, respectively, we reject that the strength-related coefficients under the Self-Assessment Treatment are equal to those under the Mutual Assessment Treatment (negative binomial regression model 1: slope coefficients $\chi^2(1) = 4.84$, $p = .027$; logistic regression model 2: slope coefficients $\chi^2(1) = 7.70$, $p < .001$, coefficients on squared difference in strength $\chi^2(1) = 3.64$, $p = .056$). In sum, when we compare the results of the Self-Assessment Treatment with the results of the Mutual Assessment Treatment, we find that removing cues to the formidability of the opponent removes the effects of upper-body strength.

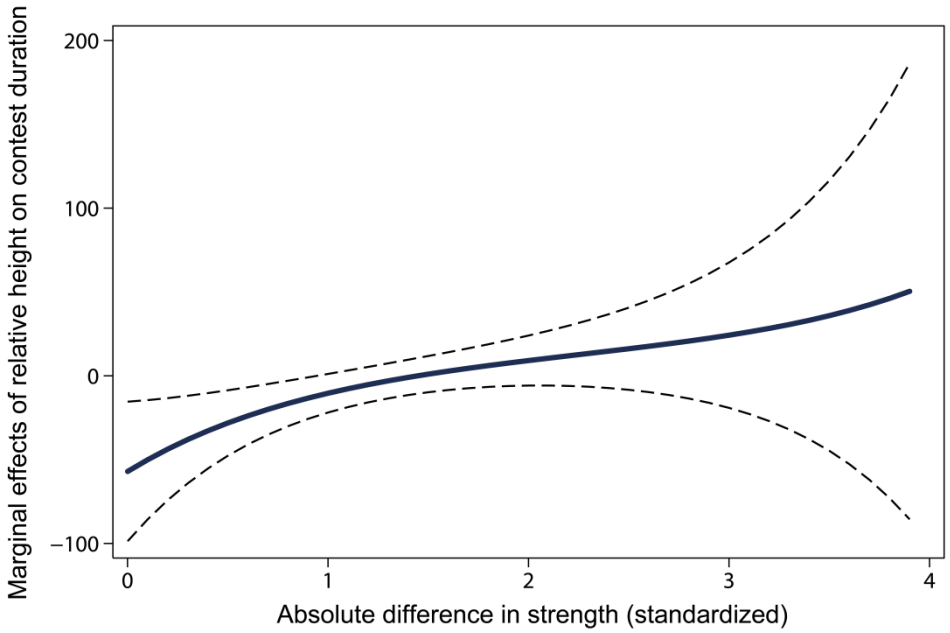
The pictures of body silhouettes used in the Mutual Assessment Treatment conveyed additional cues to height and body mass index. Potentially, any of these cues could be used for coordination. Yet, because relative upper-body strength for humans is the most important determinant of prevailing in physical contests (Sell et al., 2009a) the cognitive architecture of conflict resolution should prompt contestants to primarily coordinate on upper-body strength. In line with the prediction of (Taylor and Elwood, 2003) for the mutual assessment model (Fig. 3c), the absolute value of the coefficient on the stronger contestant's strength equals that of the weaker contestant's strength ($\chi^2(1) = 0.70$, $p = .41$; Table S2, model 4). Additional analyses confirm that relative strength functions as the primary coordination device in the experimental contests of the Mutual Assessment Treatment. For the likelihood to win, there are no significant effects for either height or body mass index (BMI) (Table S5, Table S6). We measure likelihood to win under the hypothesis that coordination is on relative height or relative BMI, respectively; that is, the outcome measure for the model

with height is whether the tallest among the two contestants wins the contest, whereas the outcome measure for the model with BMI is whether the leanest among the two contestants wins.

For contest duration, we find a significant effect of relative height in the direction predicted by the AWA-model, that is, size differences decrease contest durations, but no significant effect of relative BMI (β = -0.063, p = .69). The effect of height is biggest when the absolute difference in strength is close to zero, that is, when coordination on relative strength is most difficult and quickly becomes insignificant (Fig. 4). Height appears to be assessed mainly in the absence of other clear cues of differences in strength, which is consistent with previous research where - when actual strength is held constant - taller men were perceived to have better fighting ability (Sell et al., 2009b). Moreover, differences in strength matter for contest duration even if we control for differences in height, differences in BMI, and contestants' willingness to take risks (Table S2, model 7).

These analyses also suggest that height may be used as a secondary coordination device when strength differences are small and difficult to discern. Finally, stronger people are often believed to be more willing to take risks (see Ball et al., 2010). While a validated measure of risk tolerance does affect contest duration, we find no correlation between strength and risk tolerance (r = 0.01, p = .95, n = 90). This is in line with previous findings (Ball et al., 2010) and suggests that strength affects contest outcomes unmediated by individual differences in risk-taking.

Figure 4 **Marginal effects of the difference in height on contest duration**



An increase in the absolute difference in height (standardised) lowers the duration when cues of strength are absent. The dashed lines show the 95% confidence interval. The figure is based on Table S2, model 3.

DISCUSSION

We provide the first experimental evidence that human males intuitively cede resources to stronger rivals in non-physical conflict situations to avoid costly escalation. In spiders, frogs, deer, crickets, lions and many, many other species, a cognitive architecture has been identified that allow these organisms to avoid conflicts by assessing physical cues to the relative fighting ability of contestants and, on the basis of this assessment, resolve the conflict in the favor of the predicted winner without engaging in costly fights. Our findings demonstrate that human males share a zoologically widespread strategy for conflict resolution.

The observation that conflict resolution is organized around ancestrally relevant cues of fighting ability offers insight into the anomaly that physical characteristics shape success in modern life. Height, for example, increases labor market earnings beyond what can be explained in terms of productivity advantages (Persico et al., 2004; Case and Paxson, 2008). Our findings suggest that the repeated negotiations of interest underlying social

success are implicitly biased such that individuals with strength-related physical characteristics are more likely to get their way - even if strength is not rationally relevant to the task or negotiation at hand. Consistent with this argument, studies have shown that variation in handgrip strength explains much of the earnings-related height premium (Böckerman et al., 2010; Lundborg et al., 2014).

The identification of a basic cognitive system for conflict-resolution is of key importance because conflicts of interest pervade each and every niche of human sociality. The existence of such a system complements theories of the evolution of cooperation (Axelrod, 1984) in explaining one of the biggest puzzles of human sociality: given that human groups consist of individuals designed to propagate their own genes, why is overt aggression so relatively rare? We do not suggest that coordination on the basis of strength captures all conflict-resolution strategies available to humans. We do, however, suggest that the cognitive architecture that we have identified functions *both* as a basis for the evolution of more advanced conflict-resolution strategies *and* as a counter-weight to these strategies. Specifically, the existence of a deep intuition that advantages in fighting ability translate into a right to better treatment counter-weights other conflict-resolution mechanisms based on fairness and morality (DeScioli and Kurzban, 2009). At the same time, an intuition that it is beneficial to resolve conflicts in favor of those who are likely to win physical fights could in itself have paved the way for more advanced and fairness-based mechanisms. Indeed, for humans, the chance of prevailing in physical conflicts is not only predicted by individual upper-body strength but also by strength in numbers (Tooby and Cosmides, 2010). Through formations of alliances against the strong, the masses have been able to impose regimes that reflect wider concerns about fairness and the common good.

REFERENCES

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.

Arnott, G., & Elwood, R. W. (2009). Assessment of fighting ability in animal contests. *Animal Behaviour*, *77*(5), 991-1004.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*,*211*(4489), 1390-1396.

Ball, S., Eckel, C. C., & Heracleous, M. (2010). Risk aversion and physical prowess: Prediction, choice and bias. *Journal of Risk and Uncertainty*, *41*(3), 167-193

Bishop, D. T., & Cannings, C. (1978). A generalized war of attrition. *Journal of Theoretical Biology*, *70*(1), 85-124.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of human Resources*, *43*(4), 972-1059.

Case, A., & Paxson, C. (2008). Stature and Status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy*, *116*(3).

Daly, M., & Wilson, M. (1988). *Homicide*. Transaction Publishers.

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition*, *112*(2), 281-299.

DeScioli, P., & Wilson, B. J. (2011). The territorial foundations of human property. *Evolution and Human Behavior*, *32*(5), 297-304.

De Waal, F. B. (2000). Primates--a natural heritage of conflict resolution. *Science*, *289*(5479), 586-590.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522-550.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, *10*(2), 171-178.

Hammerstein, P., & Parker, G. A. (1982). The asymmetric war of attrition.*Journal of Theoretical Biology*, *96*(4), 647-682.

Hardy, I. C., & Briffa, M. (Eds.). (2013). *Animal contests*. Cambridge University Press.

Hörisch, H., & Kirchkamp, O. (2010). Less fighting than expected. *Public Choice*, *144*(1-2), 347-367.

Llewellyn, J. J., & Howse, R. (1999). Institutions for restorative justice: The South African truth and reconciliation commission. *University of Toronto Law Journal*, 355-388.

Lundborg, P., Nystedt, P., & Rooth, D. O. (2014). Height and earnings: The role of cognitive and noncognitive skills. *Journal of Human Resources*, *49*(1), 141-166.

McCullough, M. E., Worthington Jr, E. L., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of personality and social psychology, 73(2)*, 321.

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

Persico, N., Postlewaite, A., & Silverman, D. (2004). The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height. *Journal of Political Economy*, *112*(5).

Pietraszewski, D., & Shaw, A. (2015). Not by Strength Alone. *Human Nature*,*26*(1), 44-72.

Price, M. E., Dunn, J., Hopkins, S., & Kang, J. (2012). Anthropometric correlates of human anger. *Evolution and Human Behavior*, *33*(3), 174-181.

Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior*, *31*(3), 157-175.

Ross, M. H. (1993). *The culture of conflict: Interpretations and interests in comparative perspective*. Yale University Press.

Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., Von Rueden, C. & Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society of London B: Biological Sciences*, *277*(1699), 3509-3518.

Sell, A., Hone, L. S., & Pound, N. (2012). The importance of physical strength to human males. *Human Nature*, *23*(1), 30-44.

Sell, A., Tooby, J., & Cosmides, L. (2009a). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, *106*(35), 15073-15078.

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009b). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society of London B: Biological Sciences*, *276*(1656), 575-584.

Smith, J. M. (1974). The theory of games and the evolution of animal conflicts.*Journal of theoretical biology*, *47*(1), 209-221.

Smith, J. M., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal behaviour*, *24*(1), 159-175.

Smith, J. M., & Price, G. R. (1973). The Logic of Animal Conflict. *Nature*, *246*, 15.

Taylor, P. W., & Elwood, R. W. (2003). The mismeasure of animal contests.*Animal Behaviour*, *65*(6), 1195-1202.

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. *Science*,*331*(6016), 477-480.

Tooby, J., & Cosmides, L. (2010). Groups in mind: The coalitional roots of war and morality. *Human morality and sociality: Evolutionary and comparative perspectives*, 91-234.

SUPPLEMENTARY ONLINE MATERIALS

Participants in the pre-study. The pre-study consisted of two stages. The first stage entailed gathering material for the second stage, which was an online, survey-based experiment. Participants in the first stage are referred to as participants, while responders in the second stage are referred to as raters. For the first stage, 50 male students were recruited at the campus of Aarhus University (School of Business and Social Sciences). Participants were given a coffee-voucher (worth 7 Danish kroners (DKK) ≈ 1 USD) for their participation. Furthermore, they entered a competition to win one of three gift cards worth 200 DKK, where the chance of winning depended on how they performed on the strength tests explained below. We took the following measures: their height (M=182.86 cm, SD=6.44), their weight (M=81.08kg, SD=6.44), their chest strength (mean = 54.12 kg/F, sd=14.46), their handgrip strength (M=55.66 kg/F, SD=8.83), the circumference of their flexed biceps (M=34.95 cm, SD=2.50), and their self-reported strength measure on a scale of 0 to 100* (M=57.40, SD=20.53). From the standardized chest and handgrip strength, biceps circumference and self-reported strength we computed a composite measure, (physical) strength, for each subject (Cronbach's $\alpha = 0.79$) to be compared with the ratings in stage 2. For the second stage, 400 raters were recruited through a Qualtrics Panel from the USA with men aged between 15 and 40 years. Each rater was paid 5 USD for his participation given that he completed the survey. We conducted four online surveys using the Qualtrics survey software where raters assessed edited pictures from stage 1. By targeting men in the USA we ensured that raters would not know any of the persons in the photographs from stage 1 that were taken in Denmark.

* The precise question asked was "How physically strong are you? Out of 100 randomly picked men of your age, please state how many you would on average think are less strong than you. If you consider yourself to be stronger than 40 out of 100 men of your age, indicate so by stating 40."

Pre-study design. The following sections describe the experimental design of the pre-study.

*Stage 1***.** Participants went through the following nine steps: 1) Participants showed up at a front desk, placed outside of a closed area made up by moveable walls. They were informed about the purpose and procedure of the study by an experimenter and signed a consent form. 2) One by one participants entered the closed area, were they changed into a pair of shorts (two sizes were available and no participants had any problems fitting into one of them). 3) Full body photographs were taken (Fig. S2). The camera was placed on a tripod to ensure stability during the photo session. The participant was asked to stand at a particular point on the floor (marked with an X) against a white wall. This ensured that all the pictures had the same sizes and ratios. All participants were instructed to keep a neutral face expression and let their arms hang along their body in a relaxed manner. On the wall, a line was set at one meter's height to provide raters in the second part of the study, who assess the full-body pictures, with an indication of relative size. This deviates from previous studies where an experimenter stood right next to the participant[15]. 4) We measured the *chest strength* (Chest). The participants held a dynamometer with inversely set grips and with both hands they pushed the grips together (Fig. S1a). Two trials were given and the best result was recorded. They were encouraged to provide maximum effort and knew that better performance would increase their chances of winning one of the gift cards. This applies to step 6 as well. 5) *Height* and *weight* were measured and recorded in centimetres and kilograms. 6) *Grip strength* (Handgrip) was recorded using the participant's dominant arm (Fig. S1b). The participant was asked to hold his elbow at the side of his body and point the arm downward to achieve maximum tension. He then squeezed as hard as he could. Two trials were given and the best result was recorded. 7) A measurement tape was put around the biceps of the participant and he was asked to flex it. The *flexed bicep circumference* was then measured in centimeters at its widest point. 8) The participant changed back into his own clothes. Then, while still in the private area, he filled out a short questionnaire. 9) Once the questionnaire was filled out, the participant exchanged his questionnaire for a coffee voucher at the desk.

Stage 2. To prepare the ratings of the pictures, one picture of each participant in stage 1 was chosen, gray-scaled to reduce effects of skin colour, and edited into four different versions using Adobe Photoshop CS5: a full-body picture (with covered face), a body silhouette, a full face picture, and a silhouette of the face (Fig. S2). From the full body pictures of the 50 participants in the first stage we hence obtained 200 unique stimuli. The stimuli were divided into four different surveys. Pictures were randomly split into two groups, A and B, for which the mean physical strength level of the persons in the pictures was not significantly different (two-tailed t-test, P=0.2835), that is both sets exposed raters to pictures of participants with comparable levels of physical strength.

Four surveys were constructed, which each contained 50 pictures. Two surveys contained 25 pictures of full bodies of one group (A or B) and 25 pictures of body silhouettes of the other group, so that no participants' picture would be rated twice by the same person. The other two surveys contained 25 pictures of full faces of one group (A or B) and 25 pictures of face silhouettes of the other group. The surveys varied the order of presentation of the full picture and silhouette blocks. Pictures within each set were randomized to overcome potential order effects. Each rater participated in only one of the surveys. He rated pictures both on physical strength and physical attractiveness, one picture at a time. Each dimension was scored on a 7-point Likert scale (1 = very weak; 7 = very strong), following previous methods[14]. It was not possible to return to a rated picture. Just before rating each set, raters were shown a picture reel of the 25 pictures they were about to rate, one second per picture. With 100 raters in each survey, 100 observations were obtained per picture.

Experimental procedures of the main study. Participants were contacted by email and reminded to show up on time. All did so. At arrival, the participants registered at the front-desk and received an envelope containing their player ID and were assigned to isolated cubicles based on a prepared random allocation. The player ID links players to their pictures and was a randomly drawn 6-digits number that participants needed to enter, ensuring that participants would not accidentally be linked with a wrong picture.

Each cubicle contained a computer, a set of instructions and a booklet containing pictures with body silhouettes, printed on thick, high-quality paper (+120 g/m^2 –gloss). Each body silhouette had its own page (A4-size), prepared according to the procedure from the pre-study. As a further safe-guard to ensure anonymity of the participants, the face part of the silhouette was replaced by a black circle to avoid the possibility of identification through e.g. hairstyle. On the upper-right corner a picture number was stated (different from the participant ID). Throughout the entire session sound files with random sequences of mouse-clicking were played through several headphones lying at the non-used computers, adjusted in volume so that the room was filled with the sound of random mouse-clicking that could not be attributed to any particular location in the lab. This was to ensure that there would be no behavioral spill overs from other participants' decisions to click a button on the computer with the mouse during the experiment.

Participants were asked to carefully read the instructions[†] and to answer several questions to check that they had understood the game. Once a participant had completed reading and answered the questions, he would raise a hand and an experimenter would check the answers and privately follow up on any misunderstandings. After reviewing all the participants, the experimenter announced that the experiment was about to begin. Participants were asked to raise their hands if they had any further questions to the experiment. Then the participants were instructed to open the envelope with their participant ID's and were given an entry-code to start the experiment on the computer.

[†] Instructions are available upon request from the corresponding author

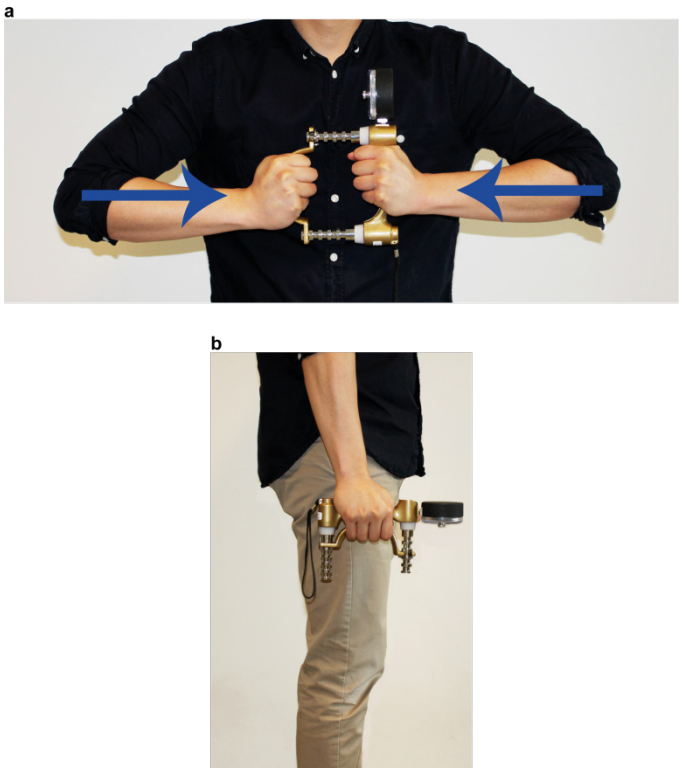


Fig. S1 Illustration of dynamometer measures. a, Assessment of chest strength. **b,** Assessment of grip strength.

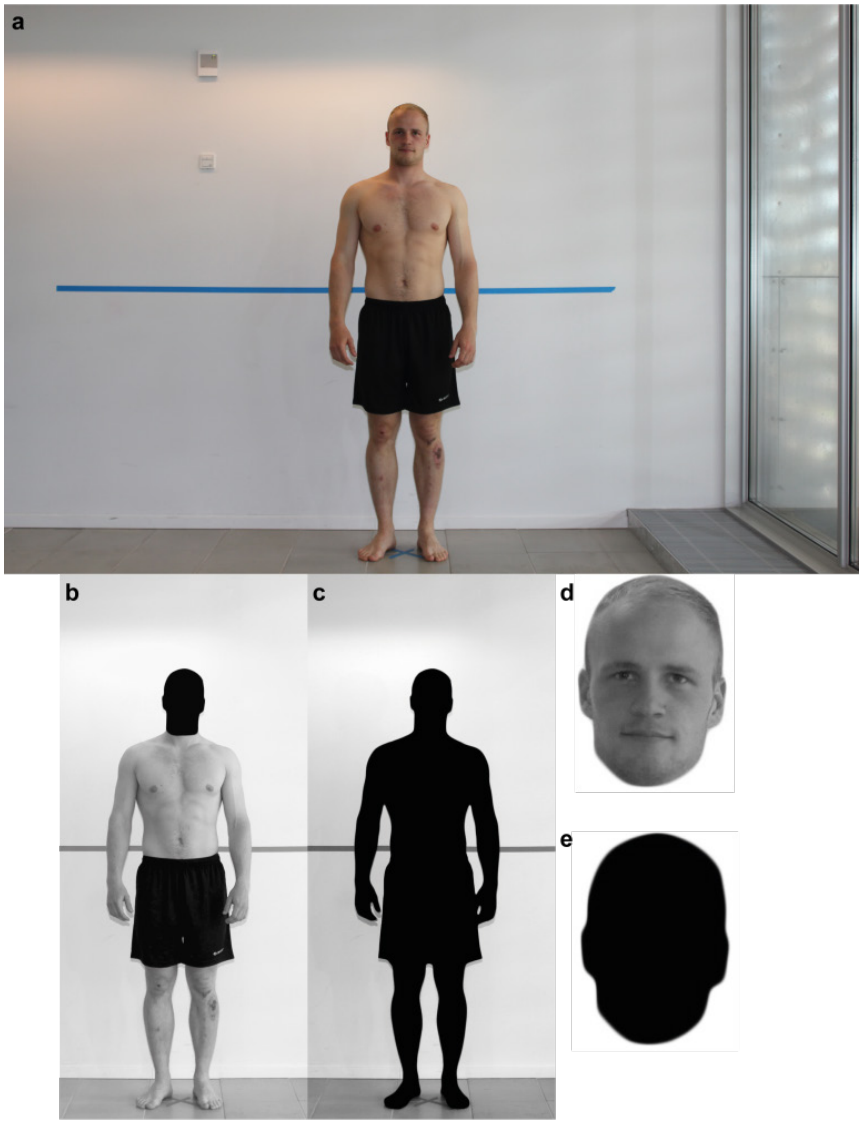


Fig. S2 Examples of the stimuli in the pre-study. a, Original photograph. A horizontal line placed at 1 meter's distance from the floor level provided an indication of relative height. **b,** Full body: A picture of the full body in a neutral stance, with the head kept on the picture but turned into a silhouette **c,** Body silhouette: A picture of the silhouette of the full body in a neutral stance. Some blurring and rounding was used to smooth out edges on the silhouette. **d,** Full face: Faces were cropped out from the full body picture and adjusted to a neutral background. Necks were not shown on the pictures. For some faces the particular light settings of the room made it difficult to distinguish the chin from the neck, complicating separation. As in (Sell et al., 2009: p. 577) head-sizes were standardised to fill a standard box of 400 pixels height. **e,** Face silhouette: A silhouette of the full face, where again edges were smoothed out.

Table S1 Logistic regressions on whether the strongest contestant wins (Mutual Assessment Treatment)

	Model 1	Model 2	Model 3	Model 4
|Difference in strength|	-0.185	1.598**		1.547**
	(0.227)	(0.675)		(0.690)
|Difference in strength|^2		-0.722***		-0.683**
		(0.257)		(0.269)
Risk tolerance of stronger rival			0.163*	0.134
			(0.084)	(0.085)
Risk tolerance of weaker rival			-0.058	-0.054
			(0.083)	(0.084)
Constant	-0.065	-0.694	-1.097	-1.338
	(0.792)	(0.825)	(1.222)	(1.315)
Observations	191	191	191	191

Dependent variable: dummy=1 (0) if strongest (weakest) contestant within a pair wins. 14 contests that ended in a tie were excluded. The predicted probability that the stronger contestant wins is increasing in the absolute difference in strength between the contestants (Model 2). Beta coefficients with robust standard errors in parentheses. All models include session and period fixed effects. * p<0.10, ** p<0.05, *** p<0.01.

Table S2 Negative binomial regressions for the duration of contests (Mutual Assessment Treatment)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model7
|Difference in strength|	-0.444**	-0.356*	-1.006***				-0.927***
	(0.181)	(0.200)	(0.233)				(0.235)
|Difference in height|		-0.302**	-0.901***				-0.869***
		(0.153)	(0.229)				(0.231)
|Diff. strength x Diff. height|			0.623***				0.639***
			(0.191)				(0.198)
Stronger rival strength				-0.497***		-0.373*	
				(0.183)		(0.193)	
Weaker rival strength				0.276		-0.054	
				(0.279)		(0.281)	
Rival with high risk Tolerance					0.065	0.107	0.051
					(0.073)	(0.071)	(0.078)
Rival with low risk Tolerance					0.148**	0.148**	0.100
					(0.070)	(0.073)	(0.075)
|Difference in BMI|							-0.063
							(0.109)
Shorter rival height						0.641***	
						(0.165)	
Heavier rival BMI						0.030	
						(0.110)	
Constant	4.385***	4.439***	5.214***	4.160***	2.629***	1.967**	4.313***
	(0.717)	(0.716)	(0.754)	(0.762)	(0.884)	(0.959)	(1.026)
ln(α)	1.001***	0.989***	0.964***	0.999***	0.995***	0.942***	0.951***
	(0.105)	(0.104)	(0.105)	(0.105)	(0.106)	(0.106)	(0.108)
Observations	205	205	205	205	205	205	205

Dependent variable: duration of contest in seconds. Beta coefficients with robust standard errors in parentheses.
All models include session and period fixed effects. * p<0.10, ** p<0.05, ***p<0.01.

Table S3 Logistic regressions on whether the strongest contestant wins (Self-Assessment Treatment)

	Model 1	Model 2	Model 3	Model 4
|Difference in strength|	-1.394***	-2.391*		-1.851***
	(0.524)	(1.436)		(0.535)
|Difference in strength|^2		0.517		
		(0.707)		
Risk tolerance of stronger rival			0.569***	0.629***
			(0.175)	(0.193)
Risk tolerance of weaker rival			-0.210***	-0.287***
			(0.076)	(0.099)
Constant	2.733**	3.102**	-1.310	1.285
	(1.340)	(1.412)	(2.102)	(1.965)
Observations	102	102	102	102

Dependent variable: dummy=1 (0) if strongest (weakest) contestant within a pair wins. 4 contests that ended in a tie were excluded. Beta coefficients with robust standard errors in parentheses. All models include session and period fixed effects.*p<0.10, **p<0.05, ***<0.01.

Table S4 Negative binomial regressions for the duration of contests (Self-Assessment Treatment)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
|Difference in strength|	0.068	0.008	0.040				0.352
	(0.317)	(0.306)	(0.522)				(0.568)
|Difference in height|		0.298**	0.317				0.642*
		(0.133)	(0.314)				(0.376)
Dif. Strength x Dif. Height			-0.017				-0.190
			(0.234)				(0.249)
Stronger rival strength				-0.541		-0.193	
				(0.500)		(0.530)	
Weaker rival strength				-0.298		-0.316	
				(0.324)		(0.263)	
Rival with high risk tolerance					0.477***	0.242**	0.487***
					(0.107)	(0.110)	(0.113)
Rival with low risk tolerance					-0.053	0.211***	0.053
					(0.045)	(0.070)	(0.070)
|Difference in BMI|							0.111
							(0.186)
Shorter rival height						-0.785***	
						(0.154)	
Heavier rival BMI						-0.246	
						(0.153)	
Constant	2.677***	2.172**	2.130*	2.928***	-0.758	-1.323	-2.703
	(0.989)	(0.983)	(1.173)	(1.026)	(1.031)	(1.169)	(1.856)
ln(α)	0.983***	0.960***	0.960***	0.949***	0.844***	0.643***	0.795***
	(0.157)	(0.156)	(0.156)	(0.156)	(0.158)	(0.169)	(0.155)
Observations	106	106	106	106	106	106	106

Dependent variable: duration of contest in seconds. Beta coefficients with robust standard errors in parentheses. All models include session and period fixed effects. * p<0.10, ** p<0.05, ***<0.01.

Table S5 Logistic regressions on whether the tallest contestant wins (Mutual Assessment Treatment)

	Model 1	Model 2	Model 3	Model 4
|Difference in height|	0.146	-0.261		0.158
	(0.222)	(0.705)		(0.233)
|Difference in height|^2		0.137		
		(0.238)		
Risk tolerance of taller rival			0.101	0.104
			(0.080)	(0.079)
Risk tolerance of shorter rival			-0.125	-0.124
			(0.088)	(0.088)
Constant	-1.358*	-1.110	-0.883	-1.077
	(0.787)	(0.873)	(1.249)	(1.265)
Observations	181	181	181	181

Dependent variable: dummy=1 (0) if tallest (shortest) contestant within a pair wins. 10 contests where rivals had equal height and 14 contests that ended in a tie were excluded. Beta coefficients with robust standard errors in parentheses. All models include session and period fixed effects. * p<0.10, ** p<0.05, ***<0.01.

Table S6 Logistic regressions on whether the leanest contestant wins (Mutual Assessment Treatment)

	Model 1	Model 2	Model 3	Model 4
|Difference in BMI|	0.065	0.494		-0.026
	(0.185)	(0.612)		(0.190)
|Difference in BMI|^2		-0.137		
		(0.189)		
Risk tolerance of leaner rival			0.047	0.047
			(0.087)	(0.087)
Risk tolerance of heavier rival			-0.246***	-0.248***
			(0.082)	(0.084)
Constant	1.320*	1.075	3.029**	3.084**
	(0.763)	(0.836)	(1.289)	(1.355)
Observations	191	191	191	191

Dependent variable: dummy=1 (0) if contestant with lowest BMI (highest BMI) within a pair wins. 14 contests that ended in a tie were excluded. Beta coefficients with robust standard errors in parentheses. All models include session and period fixed effects. * p<0.10, ** p<0.05, ***<0.01.

APPENDIX - INSTRUCTIONS

Instructions[‡]

Welcome to this experiment on human behavior. Read the following instructions carefully. You can earn a considerable amount of money during the experiment. Your earnings depend on your performance as well as the performance from other participants. After the computer session, you will be asked to perform a few tasks in the adjacent room. This will only take a few minutes and you will receive further instructions at the end of the computer experiment. This part will not affect your outcome at the computer session.

During the experiment you earn "Experimental Currency Units" (ECU). At the end you will be paid out in DKK (1 ECU = 0.35 DKK).

Please do not to talk or communicate with the other participants in the room. If you have brought any bags, notes, books, and cell phones into the laboratory, please raise your hand and we collect and store the item during the experiment. Moreover, you are not allowed to start other programs on the computer. If you do not follow the rules we have to exclude you from the experiment and you will receive no payment. If you have any questions, please raise your hand and we will come to your seat.

Overview of the experiment

The experiment consists of a number of rounds. The number of rounds will be shown in the upper-left corner of the screen during the experiment. In each round, you are paired with another randomly selected opponent

[‡] This is an example of the instruction sheets. The graphical layout varies slightly in the original version. Wordings are identical, however. There is no mentioning of any booklet or pictures in the Self-Assessment Treatment.

from this room. You will see a picture of your opponent before the round starts.

Specifically, at the beginning of the experiment you will be asked to enter your ID from the envelope. Once you have entered the ID correctly, you will see a number on the screen. Please look up the picture with that number in the booklet by your desk. The person on the picture will be your opponent in the round that is about to start.

After you have looked up the picture, and you are ready to begin – press the 'OK' button. The first round is then about to start. Please wait for a 10 second countdown. The countdown will start as soon as your opponent is ready to start as well. Please use the countdown to think about your decision.

Every following round will have the same structure:

1. You get to look up in the booklet by your desk the opponent for the round
2. There is a 10 second countdown that gives you time to think about your decision
3. The round starts

Rules and payoffs

In each round you begin with 225 ECU (the endowment).

A round lasts until you or your opponent presses the "STOP"-button (see figure).

Once a round starts every second the round lasts will cost you 1 ECU. The costs are paid no matter how the rest of the round develops. The same goes for your opponent. You can stop a round by pressing the "STOP"-button. If you press this button, neither you nor your opponent will incur any further costs.

The first one who presses the "STOP"-button wins nothing, the other person wins a prize of 100 ECU.

Specifically, here is how your round payoff is determined:

- If you press "STOP" before your opponent, then he wins the prize. Then your earnings are:

  Your payoff
  = 225 - number of seconds until first one pressed "STOP"
  (Endowment) (total costs)

- If your opponent presses the button before you, then you have won the prize. Then your earnings are:

  Your payoff
  = 225 + 100 - number of seconds until first one pressed "STOP"
  (Endowment) (Value of the prize) (total costs)

- In case you both press the button at the same time or if no one has pressed it after 225 seconds, then you share the prize. Then your earnings are:

  Your payoff
  = 225 + 100/2 - number of seconds until first one pressed "STOP"
  (Endowment) (Value of the prize/2) (total costs)

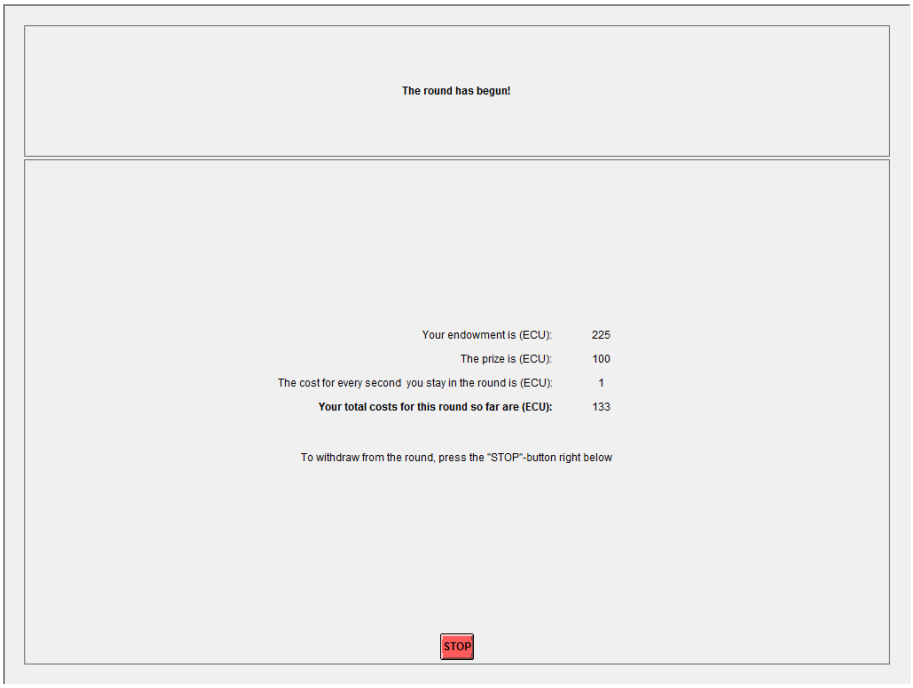


Figure: Example of the screen you will see during the experiment

After every round, a summary of the outcome of a round will be shown. Once you have read the summary and pressed 'Continue', please wait until a new opponent is matched with you. This may take a couple of minutes. While waiting for the next round to start, at random intervals a box will appear on your screen. If you see a box, please click the box as fast as possible. Please repeat if another box shows.

Once all the rounds are over the computer will randomly select one of the rounds as the round that counts. Your earnings are equal to the payoff from that round, converted to DKK at the rate 1 ECU = 0.35kr.

Please note: every round can be selected as the round that counts! So make your decisions in every round as if it was the one that counts.

End of experiment

After all the rounds have been played, a short survey will be shown on the screen. All the information you provide will be treated anonymously.

Please stay seated when the experiment is over. The experimenter will call you up one by one and you will be guided to another room for measurement taking and a final questionnaire. After this, you will receive your payments. Bring your ID-number.

Once you have read these instructions and understood them, please answer the questions on the next page. Please raise your hand once you have finished these, or if you have a question.

Thank you for participating!

Chapter 3

The Role of Physical Attractiveness and Strength in the Resolution of Intrasexual Female Conflicts

Dan Nguyen
Aarhus University

Michael Bang Petersen
Aarhus University

Alexander Koch
Aarhus University

ABSTRACT[†]

Natural selection entails that organisms inevitably face competition from other organisms. Understanding how organisms resolve such competition is a key endeavour within the biological and psychological sciences. In humans, most attention has been paid to competition between males and less attention to competition among females. However, females also compete for mates and resources among each other, making it likely that specialized psychological mechanisms for resolving conflicts in female-female

[†] **Acknowledgments:** We are grateful to Nancy Etcoff, Leonie Gerhards, Emma von Essen and Jim Sidanius for fruitful insights and valuable comments. We would also like to thank Sarah Knudsen and Jakob Fabrin for research assistance.

competitions exist. While there is consistent evidence that male-male competitions are resolved on the basis of relative differences in physical strength, we predict that female-female competitions are resolved on the basis of relative differences in physical attractiveness. Our study utilizes two existing measures of female attractiveness – self-reported attractiveness (SR-attractiveness) and the waist-to-hip ratio (WHR) – and tests if they influence the outcome and duration of competitions between female participants in a series of war-of-attrition games in which contestants compete by means of perseverance to win a monetary prize. We compare the effects of attractiveness to the effects of physical strength, which previous research has shown influences the outcome of competitions involving males. In the main experimental treatment participants received visual information about their opponent in a contest, allowing contestants to mutually assess physical cues related to attractiveness and physical strength. We provide tentative evidence that females rely on visual cues about relative attractiveness when engaging in intra-gender competition, and suggest pathways for future research.

KEYWORDS
evolutionary psychology, female intrasexual competition, attractiveness, decision-making

INTRODUCTION

There are considerable gender differences in how males and females take part and perform in competitive settings (Niederle and Vesterlund 2011; Buser et al., 2014) and competitive behavior within groups seems to be moderated by their gender composition (Niederle and Vesterlund, 2007; Gupta et al., 2013). Studies of sex differences in competitive behavior show that males are the more competitive sex, even when considering single-sex competitions (Gneezy, Niederle and Rustichini 2003, Gneezy and Rustichini 2004), although this observation seems to be dependent on the specific task performed (Apicella & Dreber, 2015). Because of this sexual dimorphism in levels of competiveness, a larger number of studies have focused on understanding the psychology underlying male-male competition and, in the course of this, have tended to neglect female-female competition (Clutton-Brock, 2007; Cross and Campbell, 2011; Stockley and Campbell, 2013). Yet, while sexual dimorphisms might exist in the level of competitiveness, any human organism - whether male or female - faces competition from others. Like males, females engage in intrasexual competition to secure mates with higher mate values and higher shares of resources for their offspring and themselves (Vaillancourt and Sharma, 2011). Furthermore, in order to avoid costly competitions escalating, both males and females need to be able to resolve their conflicts in fitness-enhancing ways. In this manuscript, we focus on intrasexual competition among females and provide empirical evidence for one particular set of strategies that females use to resolve such conflicts: coordination on the basis of physical attractiveness such that relatively less attractive females are more likely to cede in competitions against relatively more attractive females.

An increasing amount of evidence suggests that males utilize cues of relative physical strength to resolve male intrasexual conflicts. In line with the canonical model of conflict-resolution within the animal behavior literature, the asymmetric war of attrition-model (Maynard Smith and Price, 1973; Hammerstein and Parker, 1982; Maynard Smith, 1974; Maynard Smith and Parker, 1976), males with relative larger physical strength are more likely to bargain for better treatment and tend to receive better treatment (Price et al., 2011; Sell et al., 2009). Because relative differences

in physical strength are directly predictive of who would prevail if a conflict turned physical, individuals can avoid costly confrontations by assessing each other's strength and, in the case of the weaker individual, opt out of the conflict before it escalates.

It has been shown that females – like males – are able to assess upper body strength in other males or females from minimal cues (Sell et al., 2009), suggesting that females may also respond to cues of physical strength in conflict situations. Unlike males, however, females rarely engage in direct physical confrontation (Buss and Duntley, 2003, Archer 2004) and, hence, there seems to be less of a basis for resolving conflicts by assessing relative strength (Sell et al., 2009). Furthermore, whereas male intrasexual competition is often about "status, reputation and honor" (Buss, 2012: 309), female intrasexual competition is more directly focused on attracting and retaining valuable mates (Campbell, 2004). As summarized by Buss (2012: 311) "the functions of female aggression are primarily to inflict costs in intrasexual rivals". This makes it likely that the cues that females use to resolve intrasexual conflicts are cues that predict mate value in the eyes of males, such that females with lower mate value should be quicker to opt out of the conflicts (see also Sell et al., 2009). For females, a key cue of mate value is their physical attractiveness (Buss, 1988; Buss, 1989; Fisher, 2004). It is well-documented that males put a premium on female attractiveness in mating contexts (Buss, 1989) and that more attractive people obtain higher rewards in both economic, social, and health related domains (Gupta et al., 2016; Etcoff, 1999; Hamermesh and Biddle, 1994; Langlois et al., 2000). Prior studies have provided some evidence of the role of female attractiveness in conflict behavior and have showed that measures of attractiveness influence self-reported anger-proneness and feelings of entitlement in females (Sell et al., 2009; Price et al., 2011). No studies, however, have assessed whether relative differences in physical attractiveness between female contestants coordinate their actual conflict behavior.

In this manuscript, we assess the role of physical strength and attractiveness in female-female conflicts . We utilize a composite measure of upper-body strength introduced by Sell et al. (2009) and two common

measures of female attractiveness: self-reported attractiveness (SR-attractiveness) and waist-to-hip ratio (WHR). As a multi-dimensional feature (see Gangestad and Scheyd, 2005), attractiveness is often captured through self-reported attractiveness. For females, SR-attractiveness is found to be a robust measure that correlates with others' perceptions of physical attractiveness (Saad and Gill, 2009; Rand and Hall, 1983; Yen-Lin Sim et al., 2015). As a marker for female attractiveness, WHR has received a vast amount of attention (Singh, 1993, 1994; Singh et al., 2010; Streeter and McBurney, 2003). From an evolutionary perspective, WHR is an honest signal that is negatively related to a woman's fecundity (Jasienska and Ellison, 2004) and maternal abilities (Pawlowski and Grabarczyk, 2003) and a lower WHR should thus be attractive from a male's perspective. Crucially for this study, while WHR is a biological signal that we hypothesize to activate a relative assessment of attractiveness when visible, SR-attractiveness might tap into additional dimensions of a person's attractiveness and general self-esteem that play a role for the behavior contests even when no information on the opponent is available.

We tested the role of these different factors in female-female conflict-resolution utilizing the war-of-attrition paradigm – an experimental approach which has previously been successfully applied to male-male conflicts to provide evidence that the relative upper-body strength of contestants is a crucial determinant for intra-male conflict resolution (Nguyen et al., 2016). In this paradigm, two anonymous participants interact in a non-physical, economic conflict game and compete over a prize. The winner is the opponent who persists longer and persisting comes with a cost that rises with the duration of the contest. In our main experimental treatment (named the *Mutual Assessment Treatment*), participants were given an opportunity to mutually assess physical differences before they compete, by observing a picture of a body silhouette of the opponent. In another experimental treatment (the Self-Assessment Treatment), participants competed without prior assessment of mutual physical differences (i.e. they played without any visual stimuli). We test the following set of hypotheses: To the extent that resource conflict implicitly activates strategies for conflict-resolution in the contestants that use relative

physical strength or attractiveness as inputs, a greater – and, hence, more easily detectable – asymmetry in upper-body strength or attractiveness between them should (i) decrease the duration of the contest and (ii) increase the probability that the stronger contestant or the more attractive contestant prevails under conditions of mutual assessment in the war-of-attrition game. Under conditions of pure self-assessment (i.e., where no information about the opponent is available) these effects should be weaker or non-existing.

METHODS

Participants. Ninety-two females (age: M = 23.69, SD = 3.75) were recruited through an online announcement to the subject pool of the Cognition and Behavior lab at Aarhus University. Each participant received a participation fee of 80 Danish kroners (DKK), approx. 12 USD, in addition to the amount she would earn in the experimental contests.

Research design. Participants signed up for a two-part study and were obliged to complete both parts to receive any compensation. In the first part participants had an individual photo shoot time slot, where they signed a consent form, changed into one of three available sizes of hot pants and sports bras, and had a full-body picture taken. Participants stood with their back to the wall on a marked spot with their legs slightly apart and arms hanging along the sides of their body in a relaxed manner. Long hair was removed using a hair band such that the neck would be visible. Pictures were edited to only reveal the outline of the body silhouette (see Figure 1).

Participants returned within three weeks for the second part, which was conducted in a computer lab using the z-Tree experimental software (Fishbacher, 2007). They took part in one of two experimental conditions: the *Mutual Assessment Treatment* (MAT, 50 participants) or the *Self-Assessment Treatment* (SAT, 42 participants). In each treatment participants completed several rounds of a two-person war-of-attrition game according to an absolute stranger-matching protocol such that each of the six to twelve participants in a session played against the same opponent only once, in a

random order. In MAT (210 contests) participants had a booklet with body silhouettes and before each contest were given the page number with the body silhouette of their opponent. SAT (148 contests) differs from MAT only in that participants did not have such a booklet and could not see pictures of the body silhouette of their opponent. Before playing the war-of-attrition game, participants answered control questions to check that they understood the written instructions for the session. Every game started with each participant having 225 Experimental Currency Units (1 ECU=0.35 DKK ≈ 5 US cents) from which 1 ECU got deducted for each second both participants continued the game. By pressing a button, each participant could withdraw from the game at any time with a payoff of [225 – duration of contest in seconds] ECU for him and would leave the other contestant, who would receive a prize of 100 ECU for persisting longest, with a payoff of [100 + 225 – duration of contest in seconds] ECU. In case no one withdrew before 225 seconds elapsed the prize was split, leaving both contestants with a payoff of 50 ECU. After each game, a payoff-screen showed a participant the duration of the contests and whether the participant won or lost (i.e., withdrew from) the contest. The unit of analysis is a single round of the war-of-attrition game between two participants (N = 358).

Figure 1 **Examples of body silhouettes in the Mutual Assessment Treatment**

From left we portray the participants with lowest weight, tallest, heaviest, and shortest. The line on the wall indicates 1 meter.

Once all the rounds were completed, participants filled out a questionnaire on basic background data, self-perceived attractiveness and strength, and on their risk tolerance using a validated (Dohmen et al., 2011) survey measure used to reliably predict actual behavior in lotteries.

While answering questionnaire participants, one by one, were taken to an adjacent room, where we measured participants' hand grip strength (dominant arm) and chest strength using a hand dynamometer, the circumference around their flexed biceps, breasts, hips and waist using a tape measure, and their height and weight (see Table 1). None of the characteristics varied significantly across the two treatments (i.e. the confidence intervals overlap). Following Sell et al. (2009) we combined standardized scores for chest and handgrip strength, bicep circumference, and a self-reported measure of physical strength into a composite measure of physical strength (Cronbach's $\alpha = .63$).

Table 1 **Summary statistics of the independent variables**

	Mean [95% confidence interval]		**Min**		**Max**	
	MAT *n = 52*	SAT *n = 40*	MAT	SAT	MAT	SAT
Age (years)	23.75 [22.85 ; 24.64]	23.62 [22.21 ; 25.03]	18	18	38	45
Height (cm)	167.60 [165.64 ; 169.55]	169.78 [167.74 ; 171.82]	152	154	185	184
Weight (kg)	62.36 [59.77 ; 64.94]	63.46 [60.60 ; 66.31]	46.8	44.8	84.5	86.9
Grip strength (kg/F)	26.94 [25.60 ; 28.28]	29.47 [27.81 ; 31.13]	19	22	39	40
Chest strength (kg/F)	22.98 [21.09 ; 24.86]	24.17 [21.84 ; 26.50]	9	10	40	40
Flexed biceps (cm)	28.27 [27.56 ; 28.99]	28.25 [27.32 ; 29.17]	24	23	34.5	33
Self-reported strength (0-100)	53.51 [46.88 ; 60.15]	51.87 [44.78 ; 58.96]	10	10	90	95
Self-reported attractiveness (0-100)	55.46 [50.20 ; 60.71]	60.97 [54.69 ; 67.25]	10	20	90	95
Waist (cm)	73.18 [71.31 ; 75.04]	72.82 [71.02 ; 74.62]	60	63	92.5	86
Hip (cm)	98.89 [97.07 ; 100.70]	98.85 [95.90 ; 101.79]	86	62	115	119
Breast (cm)	88.74 [86.97 ; 90.50]	90.11 [88.27 ; 91.95]	76.5	79	104.5	105
Risk tolerance (0-10)	5.88 [5.29 ; 6.46]	5.72 [5.01 ; 6.43]	1	2	10	9

Note: MAT = Mutual Assessment Treatment; SAT = Self-Assessment Treatment

RESULTS

We first consider the intensity of conflicts. We hypothesized that the duration of contests would decline in the Mutual Assessment Treatment (MAT, duration: $M = 46.17$, $SD = 59.21$) when physical differences in strength or attractiveness become more apparent. As contest duration is measured in seconds and the mean duration is lower than its standard deviation, we apply the negative binomial regression model which is suitable for count data with over-dispersion in the outcome variable (Hilbe, 2011). Table 2 shows the effect of differences in attractiveness (SR-

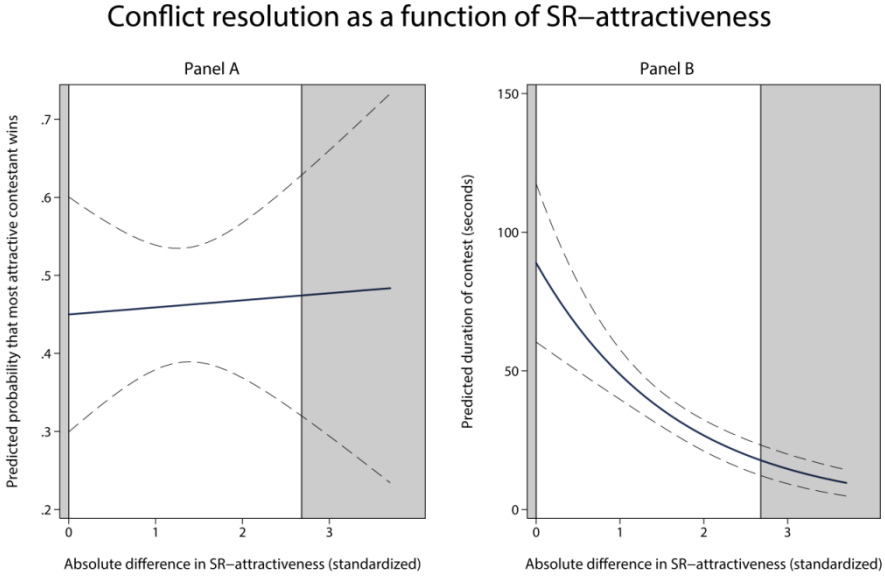
attractiveness and WHR) and strength. To the extent that both relative strength and attractiveness influence conflict intensity, there might be an interaction effect such that in the assessment of physical differences the dimension dominates where differences are most salient. We therefore include interaction effects between relative strength and the attractiveness measures. Additional specifications add differences in height -- which were found to serve as a secondary co-ordination device in intra-male contests next to differences in strength (Nguyen et al. 2016) -- and differences in body mass index (BMI) – which also capture aspects of relative attractiveness. To control for the effect of risk preferences, we add in the full model the highest and lowest risk tolerance in the pair. Contest duration should be increasing in risk tolerance.

In line with our hypothesis, differences in SR-attractiveness and strength have a significant and negative impact on contest duration (Table 2, Models 1, 2, and 8-10). We also find a negative interaction effect (Model 8), consistent with the hypothesis that greater differences in one dimension of physical difference make differences in the other dimension less salient, and hence attenuate their impact on the conflict intensity. Differences in SR-attractiveness and strength however are correlated (Pearson's $r = .22$, $p = .04$) and hence we consider the residuals from a regression of one measure on the other, to capture the effect of SR-attractiveness not explained by strength (Model 3) and the effect of strength not explained by SR-attractiveness (Model 4). The coefficient on SR-attractiveness is virtually unchanged and remains significant, whereas the effect of strength no longer is significant. This suggests that mainly SR-attractiveness impacts conflict duration. Contrary to our hypothesis regarding the alternative attractiveness measure, differences in WHR have no significant effect on conflict intensity (Model 5). Similarly, conflict intensity is not affected by differences in height (Model 6) or BMI (Model 7). Indeed, SR-attractiveness is correlated neither with WHR (Pearson's $r = .01$, $p = .92$) nor height ($r = 0.12$, $p = .24$), but it has the expected negative correlation with BMI ($r = -.30$, $p = .003$).

If intra-female competitions are resolved on the basis of relative differences in physical attractiveness, we should see that SR-attractiveness not only affects contest duration but also that greater physical differences

make it more likely that the more attractive rival wins the contest in the Mutual Assessment Treatment.

Figure 2 **The effect of differences in attractiveness on conflict outcome in the Mutual Assessment Treatment**

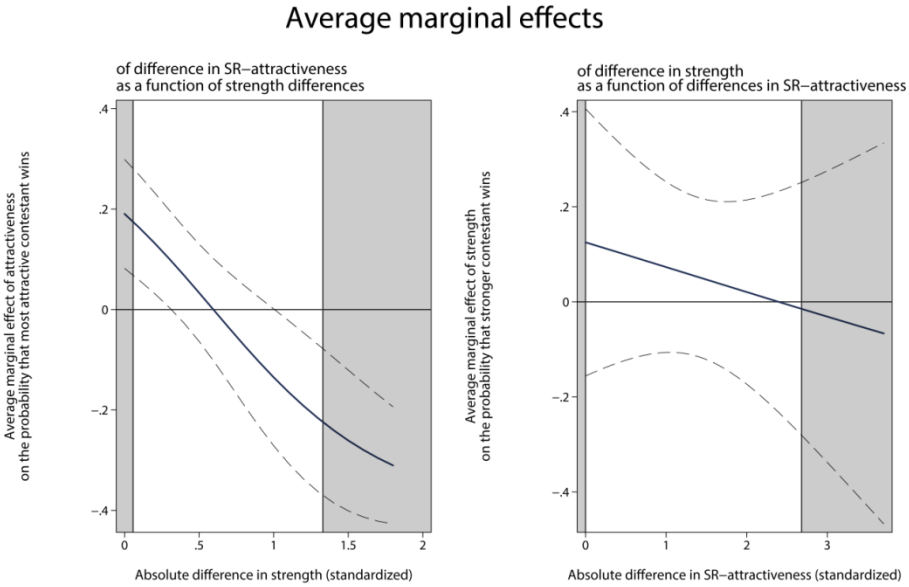


An increased absolute difference in SR-attractiveness between the contestants (a) tends to increase the predicted probability that the winner is the more attractive contestant (Table 3, Model 2 at the mean of the absolute difference in strength), (b) decreases the predicted duration of the contest (Table 2, Model 1). The dashed lines show 95% confidence intervals. The white area shows the 5th to 95th percentile range in the data.

Similarly, greater differences in physical strength should increase the probability that the strongest contestant prevails. We test these hypotheses using logistic regressions with a binary outcome variable, whether the most attractive in the pair (according to SR-attractiveness) won the game or whether the strongest contestant won the game. Table 3 shows that for the respective outcome variables the main effects of the absolute difference in SR-attractiveness (Model 1) and strength are insignificant (Model 5), but Model 2 reveals a significant interaction effect between the two for the probability that the most attractive contestant wins (joint test of linear and interaction terms, $\chi 2(2) = 10.31$, $p = .006$). Figure 2 summarizes the results regarding the impact of whether the most attractive opponent wins (Panel A) and SR-attractiveness on contest duration (Panel B). The interaction is

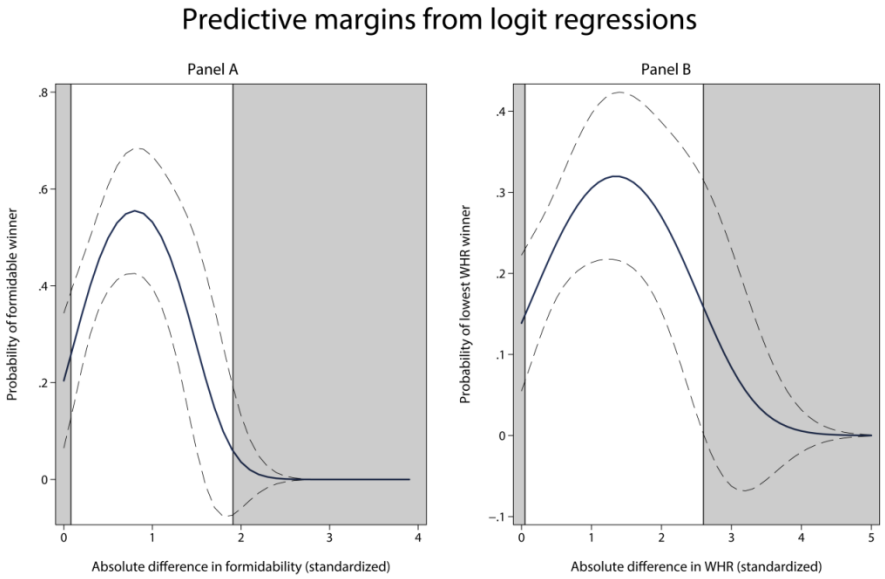
consistent with our hypothesis that the probability of the most attractive contestant prevailing is increasing in the differences in SR-attractiveness and that this effect is weakened when these differences become less salient because greater difference in relative strength detract attention from the attractiveness domain. Figure 3 illustrates this by plotting the average marginal effects in the logit regressions as a function of the variable with which the main effect is interacted.

Figure 3 **Average marginal effects from logit regressions in the Mutual Assessment Treatment**



Panel A plots the average marginal effect of SR-attractiveness differences on the probability of the most SR-attractive in the pair winning as a function of differences in strength (Table 3, Model 2 at mean of differences in SR-attractiveness)., Panel B plots the average marginal effect of strength differences on the probability of the strongest contestant winning as a function of differences in SR-attractiveness (Table 3, Model 6 at mean of strength differences). The dashed lines show 95% confidence intervals. The white area shows the 5th to 95th percentile range in the data.

Figure 4 **Predictive margins from logistic regressions in the Mutual Assessment Treatment**



Panel A plots the predicted probability of the strongest contestant winning as a function of differences strength (Table 3, Model 7). Panel B plots the predicted probability of the contestant with lowest WHR winning as a function of differences WHR (Table A1 in appendix, Model 6). The dashed lines show 95% confidence intervals. The white area shows the 5th to 95th percentile range in the data.

The significant quadratic effect in Model 7 (Table 3) indicates a diminishing marginal effect of relative strength on the probability that the strongest contestant wins (joint test of linear and quadratic terms, $\chi2(2) = 6.97$, $p = .03$). The significant coefficient on differences in waist-to-hip ratios in Model 4 suggests that WHR plays a role for the conflict outcomes, even though WHR was not significant in the context of contest duration. Regressing the binary outcome variable whether the more attractive contestant according to WHR –the contestant with lowest WHR – reveals no significant main effect but a significant quadratic effect (joint test of linear and quadratic terms, $\chi2(2) =5.22$, $p = .07$). Details are in Table A1 in the appendix. Figure 4 plots the predictive margins

Table 2 **Negative binomial regressions for the duration of contests (Mutual Assessment Treatment)**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
|Difference in SR-attractiveness|	-0.601***							-0.826***	-0.462***	-0.715***
	(0.101)							(0.165)	(0.119)	(0.194)
|Difference in strength|		-0.418*						-0.769**	-0.388*	-0.878**
		(0.222)						(0.326)	(0.234)	(0.345)
|Diff. SR-attract.| x |Diff. strength|								0.364**		
								(0.159)		
Residualized |difference in SR-attractiveness|										0.350**
			-0.577***							(0.168)
			(0.107)							
Residualized |difference in strength|				-0.273						
				(0.234)						
|Difference in WHR|					0.015				-0.113	-0.129
					(0.115)				(0.105)	(0.106)
|Difference in height|						0.128			0.211	0.187
						(0.149)			(0.144)	(0.143)
|Difference in BMI|							0.060		0.0597*	0.0668*
							(0.041)		(0.035)	(0.0342)
Highest risk tolerance in the pair									0.0248	0.0360
									(0.051)	(0.0495)
Lowest risk tolerance in the pair									0.123**	0.102*
									(0.059)	(0.0612)
Constant	5.376***	4.612***	4.631***	4.314***	4.224***	4.103***	4.105***	5.905***	4.464***	4.889***
	(0.400)	(0.375)	(0.343)	(0.308)	(0.336)	(0.365)	(0.321)	(0.496)	(0.765)	(0.866)
No. of contests	210	210	210	210	210	210	210	210	210	210

Dependent variable: duration of contest in seconds. Beta coefficients are reported with robust standard errors in parentheses
All models include period and session fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Residualized |Difference in SR-attractiveness|: residuals from regression on |Difference in strength|.
Residualized |Difference in strength|: residuals from regression on |Difference in SR-attractiveness|.

As additional controls, we again add risk tolerance. The probability of the most attractive/strongest contestant prevailing should be increasing (decreasing) in her (the opponent's) risk tolerance. The signs of the two risk tolerance coefficients are consistent with this hypothesis, though the coefficients are only jointly significant in Model 4 ($\chi2(2) = 5.14$, $p = .08$) and not in Model 8 ($\chi2(2) =1.28$, $p = .53$) (Table 3).

In sum, we find tentative support for the hypothesis that intra-female competitions are resolved on the basis of relative differences in physical attractiveness because differences in SR-attractiveness help predict both contest duration and contest outcomes. The effect of relative attractiveness appears to be influenced by relative differences in strength and strongest when physical differences exist mainly in the attractiveness dimension. We reject the hypothesis that relative differences in WHR, height, or BMI affect intra-female conflict resolution because they are not significant or not consistently significant predictors of both contest duration and outcomes.

To further investigate the hypothesis that mutual assessment of relative attractiveness matters for conflict resolution, we apply an analytical approach that is well-known from the animal contest literature that allows to detect a spurious effect of physical differences on contest duration (Taylor and Elwood, 2003). If contestants base their strategies on mutual assessment of difference in attractiveness, contest duration should be decreasing in the SR-attractiveness of the most attractive contestant (as this increases physical differences, all else equal) and decreasing in the SR-attractiveness of the least attractive contestant (as this decreases physical differences, all else equal). Alternatively, if subjects contestants base their strategies purely on how attractive they see themselves – known as a pure self-assessment strategy (Taylor and Elwood, 2003) – or if SR-attractiveness captured traits like confidence, contest duration should be increasing in the SR-attractiveness of each participant. Figure 5 shows that the predicted patterns from the estimations support the mutual assessment hypothesis.

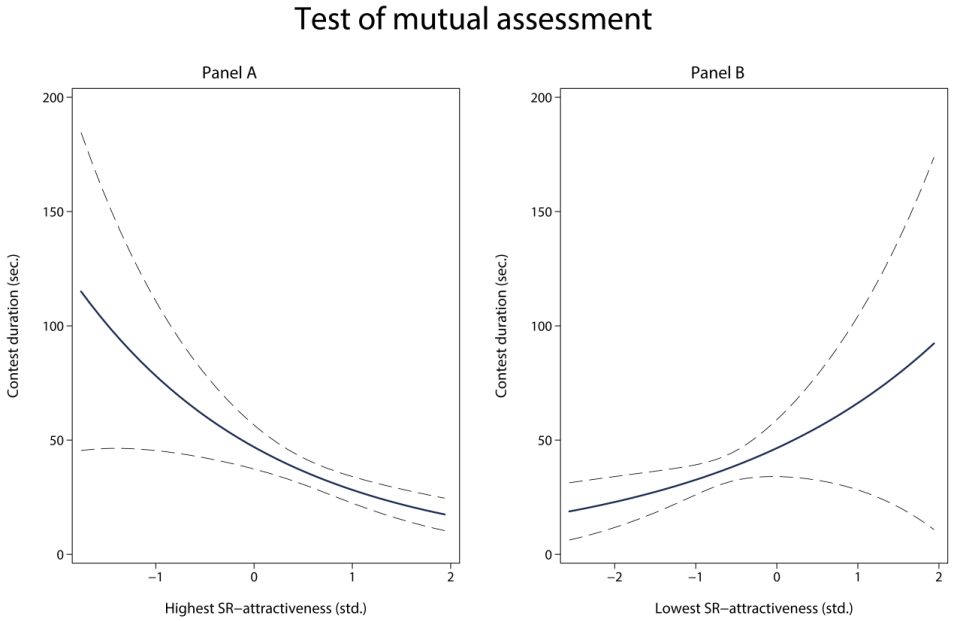
Table 3 **Logistic regressions for winner of contests (Mutual Assessment Treatment)**

Outcome variable	Most SR-attractive contestant wins				Strongest contestant wins			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
|Difference in SR-attractiveness|	0.0344	0.930***	0.989	0.641*		0.023		-0.023
	(0.215)	(0.322)	(0.733)	(0.329)		(0.326)		(0.229)
|Difference in strength|		1.970**		1.559	0.259	0.598	3.341**	3.027**
		(0.959)		(0.964)	(0.412)	(0.692)	(1.325)	(1.302)
|Diff. SR-attractiveness| x |Diff. strength|		-1.562***		-1.430***		-0.250		
		(0.509)		(0.494)		(0.399)		
$\|\text{Difference in SR-attractiveness}\|^2$			-0.306					
			(0.238)					
$\|\text{Difference in strength}\|^2$							-2.243***	-1.991**
							(0.849)	(0.825)
|Difference in WHR|				-0.490**				0.0952
				(0.229)				(0.214)
Risk tolerance of more attractive contestant				0.0352				
				(0.0963)				
Risk tolerance of less attractive contestant				-0.216**				
				(0.0961)				
Risk tolerance of stronger contestant								0.0785
								(0.0943)
Risk tolerance of weaker contestant								-0.0668
								(0.0861)
Constant	-0.207	-1.372	-0.639	1.064	-0.902	-0.958	-1.358*	-1.531
	(0.710)	(0.865)	(0.804)	(1.405)	(0.680)	(0.800)	(0.772)	(1.342)
Joint-test of linear and interaction/quadratic terms for SR-attractiveness		$\chi^2(2) = 10.31$***	$\chi^2(2) = 1.82$			$\chi^2(2) = 0.80$	$\chi^2(2) = 6.97$**	
No. of contests	181	181	181	181	197	197	197	197

Dependent variable: dummy = 1 (0) if most SR-attractive (least SR-attractive) / strongest (weakest) contestant in a pair wins. 13 contests that ended in a tie were excluded, and Models 1-4 exclude 16 contests where contestants had identical SR-attractiveness. Beta coefficients are reported with standard errors in parentheses. All models include session and period fixed effects. * p<0.10, ** p<0.05, *** p<0.01.

In addition, we can compare results with those from the *Self-Assessment Treatment* (duration: $M = 38.23$, $SD = 56.35$) treatment, which is identical to our main treatment except that contestants saw no pictures of their opponent. If contest behavior is determined by processes within each participant, qualitatively similar data patterns as in our main treatment should emerge. In contrast, if the above results reflect behavior that relies on simultaneous information about both the self and the opponent, then the effects should disappear in the Self-Assessment Treatment. Consistent with the latter hypothesis, (i) relative attractiveness is not a significant predictor of contest duration in the SAT treatment (Table 4, Model 2) and the individual contestants' SR-attractiveness are jointly insignificant (Table 4, Model 4), and (ii) in the logit regressions whether the most attractive contestant was most likely to win, the coefficients for relative attractiveness and its interactions are jointly insignificant in the SAT treatment (Table 4, Model 6).

Figure 5 **Evidence for mutual assessment**



The mutual assessment hypothesis predicts that contest duration is determined by the difference in attractiveness, which increases (decreases) with the attractiveness of the most attractive (least attractive) contestant. Under the alternative pure self-assessment hypothesis contest duration is determined by a contestant's own attractiveness and should be increasing in attractiveness for both contestants. The predicted duration of the contests derived from the negative binomial regressions for the Mutual Assessment Treatment (Table 4, Model 3). Panel A (B) holds the SR-attractiveness of the least attractive (most attractive) contestant as well as the other covariates fixed at the average observed values. The dashed lines show 95% confidence intervals.

Table 4 **Mutual Assessment Treatment vs. Self-Asessment Treatment**

Outcome variable	Duration		Duration		Most SR-attractive contestant wins	
	Model 1 *MAT*	Model 2 *SAT*	Model 3 *MAT*	Model 4 *SAT*	Model 5 *MAT*	Model 6 *SAT*
|Difference in strength|	-0.388*	0.0839			1.559	0.982
	(0.234)	(0.269)			(0.964)	(0.941)
|Difference in SR-attractiveness|	-0.462***	-0.200			0.641*	-0.146
	(0.119)	(0.127)			(0.329)	(0.468)
|Difference in WHR|	-0.113	-0.152	-0.110	-0.175*	-0.490**	0.0695
	(0.105)	(0.0981)	(0.106)	(0.0964)	(0.229)	(0.190)
|Difference in height|	0.211	0.229	0.220	0.280		
	(0.144)	(0.163)	(0.144)	(0.179)		
|Difference in BMI|	0.0597*	0.0495	0.0551	0.0299		
	(0.0347)	(0.0530)	(0.0358)	(0.0538)		
Highest risk tolerance in the pair	0.0248	0.00677	0.0243	0.00886		
	(0.0513)	(0.0832)	(0.0503)	(0.0904)		
Lowest risk tolerance in the pair	0.123**	0.229***	0.160**	0.253***		
	(0.0590)	(0.0645)	(0.0674)	(0.0668)		
Highest SR-attractiveness in the pair			-0.509***	-0.321*		
			(0.132)	(0.174)		
Lowest SR-attractiveness in the pair			0.354**	0.134		
			(0.172)	(0.139)		
Strongest in the pair			-0.375	0.227		
			(0.243)	(0.327)		
Weakest in the pair			0.118	-0.0267		
			(0.339)	(0.310)		
|Difference in SR-att.| x |Difference in strength|					-1.430***	-0.157
					(0.494)	(0.545)
Risk tolerance of most attractive rival					0.0352	0.0144
					(0.0963)	(0.109)
Risk tolerance of least attractive rival					-0.216**	-0.120
					(0.0961)	(0.0875)
Constant	4.464***	2.739***	4.083***	2.740***	1.064	0.559
	(0.765)	(0.771)	(0.810)	(0.857)	(1.405)	(1.266)
Test of joint significance of coefficients			Most attractive and least attractive		Attractiveness and strength linear and interaction terms	
			$\chi^2(2)$ =15.63***	$\chi^2(2)$ =3.49	$\chi^2(3)$ =12.87***	$\chi^2(3)$ =2.67
N	210	148	210	148	181	132

DISCUSSION

The findings presented above provide tentative evidence that attractiveness plays an important role in competitions between females, and that female intrasexual competitions are resolved by assessing relative differences in physical attractiveness. When participants were able to visually assess the attractiveness of their opponent, the difference in attractiveness affected the outcome and duration of a series of war-of-attrition games. We find robust evidence that contest duration is decreasing in the differences in attractiveness, and that this effect is driven by mutual assessment rather than being an artefact of internal processes that rely on self-assessment only. While the evidence also suggests that less attractive females where quicker at opting out of the competition, the greater the differences in attractiveness, these results are less clear cut and more tentative.

Our measure of relative attractiveness is based on self-reported assessments of attractiveness by the participants, which is likely to be a noisy proxy for the way participants in fact perceive relative attractiveness when assessing the body silhouette of their opponent. To avoid experimenter demand effects we did not ask participants about their perception of their opponent and instead infer physical differences in attractiveness from self-reported attractiveness measures collected in a questionnaire after all contests were over. In future research, external raters could be presented with pictures of the body silhouettes of the players and asked to rate physical attractiveness. Our SR-attractiveness measure could then be replaced with the average or median attractiveness rating, differences in attractiveness be computed from the z-scored measure, and our above analysis could be redone with the more appropriate proxy measure. The standard deviation of ratings could provide an indication about how noisy this measure is as a proxy for the way contestants in our experiment perceived relative attractiveness. With this additional information, the analysis could be performed on a subsample of contests for which the proxy measure is likely to be more precise – because raters tend to agree in their assessment of both opponents.

Both SR-attractiveness and WHR are conventional proxies for attractiveness in the literature and we had no a prior reason to expect that WHR would only influence the outcome and not the duration of contests. Future research will need to explore this finding further.

Overall, this paper provides the first behavioral and experimental evidence that females spontaneously resolve their conflicts in a coordinated fashion on the basis of cues of physical attractiveness. Furthermore, it provides evidence of the existence of sex-specific conflict-resolution strategies that are narrowly tied to the sex-specific nature of competitions. To contrast with previous findings for intra-male conflicts (Nguyen et al. 2016), physical attractiveness rather than physical strength affects the outcome and duration of contests among females. Conversely, there is no effect of SR-attractiveness on conflict resolution in intra-male contests. As the hypotheses related to SR-attractiveness were not tested Nguyen et al. (2016), we provide the related regressions in the appendix (Table A2, Model 1 and 4). WHR was not recorded in the experiments with intra-male contests. This adds credibility to the notion that males and females use distinct, evolutionary evolved psychological mechanisms when engaging in and resolving competitions.

REFERENCES

Apicella, C. L., & Dreber, A. (2015). Sex differences in competitiveness: hunter-gatherer women and girls compete less in gender-neutral and male-centric tasks. *Adaptive Human Behavior and Physiology*, *1*(3), 247-269.

Archer, J. (2004) Sex differences in aggression in real-world settings: A meta-analytic review. *Review of general Psychology* 4 291–322.Bates, E. A., Graham-Kevan, N., & Archer, J. (2014). Testing predictions from the male control theory of men's partner violence. *Aggressive Behavior*, *40*(1), 42-55.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness and career choices. *The Quarterly Journal of Economics*, *129*(3), 1409-1447

Buss, D. M. (1988). The evolution of human intrasexual competition: tactics of mate attraction. *Journal of personality and social psychology*, *54*(4), 616.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and brain sciences*, *12*(01), 1-14.

Buss, D. M. (1989). Conflict between the sexes: strategic interference and the evocation of anger and upset. *Journal of personality and social psychology*,*56*(5), 735.

Buss, D. (2012). *Evolutionary psychology: The new science of the mind*. Psychology Press.

Buss, D. M., & Duntley, J. D. (2011). The evolution of intimate partner violence. *Aggression and Violent Behavior*, *16*(5), 411-419.

Campbell, A. (2004). Female competition: Causes, constraints, content, and contexts. *Journal of sex research*, *41*(1), 16-26.

Clutton-Brock, T. (2007). Sexual selection in males and females. *Science*,*318*(5858), 1882-1885.

Cross, C. P., & Campbell, A. (2011). Women's aggression. *Aggression and Violent Behavior*, *16*(5), 390-398.

Etcoff, N. (1999). Survival ofthe prettiest. *The Science of Beauty. London: Little, Brown And Company*.

Fisher, M. L. (2004). Female intrasexual competition decreases female facial attractiveness. *Proceedings of the Royal Society of London B: Biological Sciences*, *271*(Suppl 5), S283-S285

Gangestad, S. W., & Scheyd, G. J. (2005). The evolution of human physical attractiveness. *Annual Review of Anthropology*, *34*, 523-548.

Gneezy, U, Niederle, M, & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics* 118,1049–1074.

Gneezy, U. & Rustichini, A. (2004). Gender and competition at a young age. American Economic Review 94,377–381

Gupta, N. D., Poulsen, A., & Villeval, M. C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, *51*(1), 816-835

Gupta, N. D., Etcoff, N. L., & Jaeger, M. M. (2016). Beauty in mind: The effects of physical attractiveness on psychological well-being and distress.*Journal of Happiness Studies*, *17*(3), 1313-1325.

Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the Labor Market. *The American Economic Review*, *84*(5), 1174-1194.

Hammerstein, P., & Parker, G. A. (1982). The asymmetric war of attrition. *Journal of Theoretical Biology* 96(4), 647-682.

Hoffman, M., & Gneezy, U. (2010). Left Handed Women are More Competitive than Right Handed Men: On the Biological Basis of Gender Differences in Competitiveness. *Work. Pap., Univ. Calif., San Diego*

Jasienska, G., & Ellison, P. T. (2004). Energetic factors and seasonal changes in ovarian function in women from rural Poland. *American Journal of Human Biology*, *16*(5), 563-580

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*(7042), 673-676

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological bulletin*, *126*(3), 390.

Maynard Smith, J. & Price, G.R. (1973). The logic of animal conflict. *Nature* 246, 15-18.

Maynard Smith, J.. (1974). The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* 47, 209-221.

Maynard Smith, J. & Parker, G.A. (1976). The logic of asymmetric contests. *Animal Behaviour* 24, 159-4175.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics 122*, 1067-1101

Niederle, M., & Vesterlund, L. (2011). Gender and Competition, Annual Review in Economics, 3, 601–30.

Nguyen, D., Petersen, M. B., Nafziger, J., & Koch, A. (2016). Upper-body strength and conflict resolution in human males. *Working paper*

Pawłowski, B., & Grabarczyk, M. (2003). Center of body mass and the evolution of female body shape. *American Journal of Human Biology*, *15*(2), 144-150

Price, M. E., Kang, J., Dunn, J., & Hopkins, S. (2011). Muscularity and attractiveness as predictors of human egalitarianism. *Personality and Individual Differences*, *50*(5), 636-640.

Rand, C. S., & Hall, J. A. (1983). Sex differences in the accuracy of self-perceived attractiveness. *Social Psychology Quarterly*, 359-363

Saad, G., & Gill, T. (2009). Self-ratings of physical attractiveness in a competitive context: When males are more sensitive to self-perceptions than females. *The Journal of social psychology*, *149*(5), 585-599

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.

Sim, S. Y. L., Saperia, J., Brown, J. A., & Bernieri, F. J. (2015). Judging attractiveness: Biases due to raters' own attractiveness and intelligence.*Cogent* Psychology, *2*(1), 996316

Singh, D. (1993). Adaptive significance of female physical attractiveness: role of waist-to-hip ratio. *Journal of personality and social psychology*, *65*(2), 293

Singh, D. (1994). Is thin really beautiful and good? Relationship between waist-to-hip ratio (WHR) and female attractiveness. *Personality and Individual Differences*, *16*(1), 123-132

Singh, D., Dixson, B. J., Jessop, T. S., Morgan, B., & Dixson, A. F. (2010). Cross-cultural consensus for waist–hip ratio and women's attractiveness.*Evolution and Human Behavior*, *31*(3), 176-181

Streeter, S. A., & McBurney, D. H. (2003). Waist–hip ratio and attractiveness: New evidence and a critique of "a critical test". *Evolution and Human Behavior*,*24*(2), 88-98

Stockley, P., & Campbell, A. (2013). Female competition and aggression: interdisciplinary perspectives. *Phil. Trans. R. Soc. B*, *368*(1631), 20130073.

Taylor, P. W., & Elwood, R. W. (2003). The mismeasure of animal contests.*Animal Behaviour*, *65*(6), 1195-1202

Vaillancourt, T., & Sharma, A. (2011). Intolerance of sexy peers: Intrasexual competition among women. *Aggressive behavior*, *37*(6), 569-577.

SUPPLEMENTARY MATERIALS

Table A1 **Logistic regressions for the winner with lowest WHR of contests (Mutual Assessment Treatment)**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
|Difference in WHR|	0.139			0.239	-0.0338	1.786**	1.559**
	(0.221)			(0.355)	(0.332)	(0.781)	(0.786)
|Difference in SR-attractiveness|		-0.210		-0.0540			-0.194
		(0.252)		(0.361)			(0.269)
|Difference in strength|			-0.0793		-0.333		0.0334
			(0.484)		(0.636)		(0.479)
|Difference in SR-attractiveness| x |Difference in WHR|				-0.151			
				(0.323)			
|Difference in strength| x |Difference in WHR|					0.359		
					(0.464)		
|Difference in WHR|^2						-0.662**	-0.638*
						(0.309)	(0.332)
Risk tolerance of competitor with lower WHR							0.117
							(0.109)
Risk tolerance of competitor with higher WHR							-0.155
							(0.106)
Constant	-2.810***	-2.324***	-2.541***	-2.654***	-2.680**	-3.348***	-2.769*
	(0.862)	(0.837)	(0.972)	(1.014)	(1.100)	(0.874)	(1.564)
N	195	195	195	195	195	195	195

Table A2 **The effect of differences in attractiveness on intra-male conflict outcomes in the Mutual Assessment Treatment of Nguyen et. al (2016)**

Outcome variable	Contest duration (sec.)			Most SR-attractive contestant wins		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
|Difference in SR-attractiveness|	0.0819		0.241	0.188	0.491*	0.912*
	(0.105)		(0.156)	(0.170)	(0.262)	(0.540)
|Difference in strength|		-0.444**	-0.268		0.0300	
		(0.181)	(0.254)		(0.363)	
|Diff. SR-attract.| x |Diff. strength|			-0.127		-0.227	
			(0.131)		(0.154)	
|Difference in SR-attractiveness|^2						-0.208
						(0.146)
Constant	3.562***	4.385***	3.915***	-1.073	-1.038	-1.300
	(0.650)	(0.717)	(0.758)	(0.853)	(0.947)	(0.897)
Joint-test of linear and interaction/quadratic terms for SR-attractiveness			$\chi^2(2) = 2.49$		$\chi^2(2) = 3.52$	$\chi^2(2) = 3.23$
No. of contests	205	205	205	176	176	176

Dependent variable: mdels 1-3: duration of contest in seconds ; models 4-6: dummy = 1 (0) if most SR-attractive (least SR-attractive) contestant in a pair wins. In models 4-5, 14 contests that ended in a tie and 15 contests where contestants had identical SR-attractiveness were excluded. Beta coefficients are reported with standard errors in parentheses. All models include session and period fixed effects. * p<0.10, ** p<0.05, *** p<0.01.

Chapter 4

Institutional Norms of Fairness and Support for Taxation

Dan Nguyen
Aarhus University

ABSTRACT[†]

It is often argued that *how* the welfare state operates matters for citizens' willingness to pay tax. Using a lab experiment (N=135), I investigate this notion by testing the effect of institutional rules on preferred tax rates. Building on Rothstein (1998; 2001), I study variations in institutional rules that manipulate citizens' perception of *substantial fairness* (that only those deserving of help get help), *procedural fairness* (that help is given according to transparent rules), and *fair burden* (that the costs of distributional policies are fairly distributed between citizens). These institutional rules mirror real world differences between universal and selectivist welfare states. I find that perceptions of both *substantial fairness* and *procedural fairness* affect the participants' preferred tax rate. This partially corroborates the notion that *how* the welfare state operates matters for citizens' willingness to pay tax.

KEYWORDS
taxation, institutions, experimental, fairness norms

[†] **Acknowledgments:** I am grateful for valuable comments and suggestions from Mette Trier Damgaard, Emma von Essen, Alexander Koch, Daniel Wajsman Mahler, Carsten Jensen, Michael Bang Petersen and participants at the Sidanius Lab seminar. Thanks to the Department of Economics and Business Economics at Aarhus University for providing the necessary funding.

INTRODUCTION

The top income tax rate in Denmark in 2012 was 60% with an income threshold at $54.900 a year, while in the United States the top tax rate was only 44% at $400.300 (OECD, 2012). Why do citizens in some countries accept a higher tax rate than citizens in other countries? This question has vexed researchers for a long time. Rothstein (1998; 2001) presents one of the most widely cited and interesting explanations. He argues that the institutional organization of the welfare state is an important component: if the institutional rules of the welfare state are perceived as fair, citizens are willing to pay the required taxes. Conversely, if a variation in the institutional rules is perceived as making them less fair, citizens' willingness to pay taxes will decrease.

Rothstein specifies three fairness norms that are particularly important to citizens when deciding their preferred tax level. These are *substantial fairness* (that only those deserving of help get help), *procedural fairness* (that help is given according to transparent rules), and *fair burden* (that the costs of the welfare state are shared between citizens in a fair manner). These are all norms that directly relate to how the welfare state operates. According to Rothstein, citizens of so-called *universal* welfare states – such as Denmark or Sweden – feel that all three fairness norms are satisfied because of the institutional rules, whereas *selectivist* welfare states – such as the United States and the United Kingdom – create the opposite feeling. This, in turn, helps explain the big cross-country variation observed in both public preferences for taxation and welfare state generosity (Larsen, 2008).

Although Rothstein's argument is widely known, to my knowledge the idea that a variation in the institutional rules of the welfare states can enhance citizens' willingness to pay taxes has not been thoroughly tested. While it is reasonable to assume that individuals vary in their preferences for taxation because of different expectations of the governance structure of the tax money, Rothstein provides little explanation for how the institutions themselves came about. The support for such institutions must at a fundamental level be reflected by preferences about redistribution. My goal is to separate the possible explanation that cultural and historical antecedents provide on preferences on taxation to isolate the effects of the

institutional features. Of course, a problem with field data, e.g. from country comparisons and survey research, is the lack of insight it brings in understanding causality (Gerber and Green, 2002).

This paper therefore tests Rothstein's argument in a controlled laboratory experiment, where participants can earn money through a real effort task and afterwards vote for a tax rate where tax proceeds are redistributed lump sum. The experimental approach allows me to isolate the hypothesized link from institutions to preferences for taxation excluding alternative explanations. Through this systematic variation in the treatments and institutional setting, I test how participants' votes for redistribution are affected when either one or more of the fairness norms proposed by Rothstein are violated. I find that both *substantial fairness* and *procedural fairness* affect the participants' preferred tax rate. This is in line with Rothstein's notion that *how* the welfare state operates matters for citizens' willingness to pay tax, but also suggests that not all three proposed norms are equally important.

EXISTING LITERATURE ON TAX PREFERENCES

A review of the vast literature in political science and economics on the welfare state is beyond the scope of this article. In this section, I selectively review the most relevant literature, focusing on experimental studies of voters' preferences for taxation. The main take-away point is that none of these studies integrates the institutional rules or resource allocation; i.e. neglect how the welfare state operates. More broadly, this literature is not well-equipped to explain cross-country variation in tax preferences because it does not mimic real-world cross-country differences in institutional rules of welfare states. Rather, the studies implicitly assume that human behavior will be identical across national contexts. However, theories about universal human traits will inevitably have difficulty explaining the particulars of individual countries without knowledge of the social environment.

I structure the review by the way in which income has been generated. Tyran and Sausgruber (2006) conduct an experiment where voters' endowments are randomly drawn. The study aims at classifying and testing the Fehr-Schmidt (1999) model of inequity aversion in the context of voting

for redistribution and it identifies three broad types of voters: About 25% make decisions in a self-interested manner, for 37.5% of the voters weak inequality aversion explains behavior, whereas about 30% of the voters display strong inequality aversion. The last 7.5% of the subjects do not correspond to any of the types. The study is relevant for how fairness preferences alter redistributional preferences in a democratic voting setup.

Similarly, Ackert et al. (2007) use random assignment of endowments and let participants choose between a lump sum tax and a progressive tax system. 15% of the participants endowed with high income voted for the progressive tax scheme and thus rejected the payoff maximizing option (the lump sum tax). This result is similar to the findings of Engelmann and Strobel (2004).

However, as highlighted by Cherry et al. (2002), the artificial experimental setting where income is randomly assigned might induce altruistic behavior that is not present in real life where income must be earned. Introducing production phases with quizzes or real effort tasks, however makes the notion of fairness more ambiguous. For example, Cappelen et al. (2007) show that, depending on context, people differ in how much they care about fairness norms and in their fairness ideals. In addition, people apply fairness norms conditionally on whether differences in outcomes are based on luck or the result of explicit choices. For example, Cappelen et al (2013) found that third-party spectators were much more likely to redistribute payoffs between lucky and unlucky participants who were exposed to a random income shock compared to when participants' income was the result of a deliberate choice between a risky bet (going for the "jackpot") and playing it safe.

In the experiment by Krawcyk (2010), income was generated through luck or effort. In addition, he varied two additional parameters: the loss in collecting and redistributing taxes (distortionary effects), and the probability with which a particular player would win a fixed monetary prize (the player's chances to "climb the social ladder" veil of ignorance). He finds that more redistribution is favored if winning came from a random draw rather than performance in a task but was not related to probability itself of winning.

In the experiment of Esaray et al. (2012) participants generate income in a real effort task (spelling test) over several periods. Participants vote for taxation level after five periods that are then followed by additional production phases for which income generated is subject to the implemented tax rate. The authors test whether the origins of unequal endowments (meritocracy or luck) affect preferences for redistribution. They find that support for redistribution is linked to the relative earnings position of the voters rather than to how the inequality was generated.

In the experiment by Kataria and Montinari (2012), subjects generated income through three different mechanisms. In a between-subjects design, the first treatment allows participants to earn income by competing in a quiz and rewarded them based on their ranking. In the second treatment, participants are randomly divided into a rich (above the median) and poor (median and below) group. The participants then only compete within their assigned group and can thus never advance from being poor to rich (or rich to poor) no matter how well they perform in their groups. The authors also vary whether participants have knowledge about which group they are assigned to. The third treatment assigns income randomly. After the income generating phase, the participants vote for a tax rate between 0-100% using the median voting rule in which the preferred median tax rate is implemented. Overall, the authors find that a majority of the rich participants chose a non-zero tax rate rejecting the hypothesis that preferences for taxation are purely driven by self-interest. The effect was stronger when role uncertainty (being rich or poor) was introduced.

Durante et al. (2014) examine how preferences for taxation vary over various levels of costs of taxation, efficiency losses and whether income is uncertain. Overall, the authors find that self-interest motives among the worst performers, insurance concerns, and the concern for others' well-being are important factors in voting for a higher taxation.

Compared to these studies, the focus of my study is to see how differences in fairness norms of the institution itself affect distributional choices. In addition, my experiment contributes to the literature by examining how entitlements to pre-tax income affect redistribution concerns when inequality is kept fixed across sessions and treatments. And contrary

to the study by Cappelen et al. (2013), the participants are informed about the redistribution before they make a choice on the level of the redistribution. My experiment also varies significantly from the Durante et al. (2014) study on both purpose and features. While they both look at shifts in voters' taxation preferences, their study primarily draws attention to the income generating process and the economic efficiency of the tax collected. As Höchtl, Sausgruber and Tyran (2012) stress, such structural factors matter too when people vote for redistribution. However, my study examines a closer-to-reality framework by focusing on the manner in which taxes are collected and redistributed instead of describing tax money in terms of an abstract notion of monetary efficiency.

THEORETICAL FRAMEWORK: THE ROTHSTEIN ARGUMENT

There are big differences in the willingness to pay (high) taxes between nations. Why? Only little research has so far tried to answer this question head-on. One exception, however, emphasizes how moral norms get institutionalized in public policies can explain why some countries, such as the Nordic countries, stand out in terms of the taxes that citizens are willing to pay (Rothstein 1998; 2001). If citizens can expect the institutions of a society to ensure efficient and fair allocations, the willingness to pay should be higher. Rothstein proposes that redistributive policies in the Nordic countries are characterized by three features that align with three universal moral norms, thereby creating a feeling among citizens that the redistributive policies and, hence, the associated taxes are morally just. Redistributive policies are supported by the public if they are defined by:

Substantive justice. According to Rothstein, individuals with fewer resources are seen as deserving of help if they are perceived as being without control over their own situation. Hence, support for redistributive policies is lower when people who benefit are perceived as having high control over their situation and therefore being responsible for needing help.

Procedural justice. Support for redistributive policies is lower when the decisions about granting benefits appear non-transparent or arbitrary. In contrast, when entitlements are granted without any form of means-test or

assessment of worthiness, support is higher, because benefits then are distributed in a non-arbitrary and transparent way.

Fair share of burden. Support for redistributive policies is also lower if individuals believe that all citizens are not paying their fair share because this creates a sentiment that other people are trying to free-ride (Kahan, 2005; Roosma, van Oorschot and Gelissen, 2015).

EXPERIMENTAL DESIGN AND PROCEDURES

The experiment is a two-way mixed factorial design in which participants first encountered one treatment in which *substantive justice* was manipulated (High Control (HC) vs. Low Control (LC) treatment) followed by either a *procedural injustice treatment* (PI) or *unfair share of burden* treatment (UB), while features of the first treatment still being enforced. Intuitively it can be thought of adding additional "unfairness" to the institution that the participants encounter. The experiment can thus be thought of as having four different treatments in which the substantive justice treatment serves as benchmark:

1) High Control + Unfair Share of Burden
2) Low Control + Unfair Share of Burden
3) High Control + Procedural Injustice
4) Low Control + Procedural Injustice

The institutions were designed to mimic fairness norms that separate selective welfare state regimes from universal welfare regimes. The treatments will be explained more thoroughly below.

Each session began with participants being randomly allocated to a computer among two rows of cubicles. Participants were given a set of instructions (see Appendix B), which they were told to read carefully, and then instructed to answer the follow-up questions to control that they had understood the rules of the experiment. At each session the participants were divided into groups of five. This is a direct application of Rothstein's (1998) own paradigm in his explanation of the lump sum mechanism. The experiment consisted of two periods each consisting of two stages.

In stage 1, the participants accrued initial earnings (i.e. pre-tax earnings) by solving sliders in a slider task (Gill and Prowse, 2009). For each correctly solved slider they earned 0.25 DKK. They were given 10 minutes to accomplish as many sliders as they could and wanted to. Once the time had passed, I elicited the participants' beliefs about their ranking within the group [*]. The elicitation of these beliefs was incentivized such that a participant would earn an additional 5DKK for a correct guess.

In stage 2, the participants were handed out instructions (Announcements, see Appendix C-G) for the voting procedure. The specific details about the institution were revealed, and once everyone had a chance to read and understand them, each participant cast their vote. The tax rate was chosen by the median voting rule; that is, the tax rate implemented was the one exactly at the median of all votes. Such a rule is strategy-proof and in theory should induce the participants to reveal their actual preference (Esarey, 2012; Saporiti and Tohmé, 2003). The tax rate is applied to pre-tax earnings and the total proceeds are redistributed lump sum to arrive at the post-tax earnings. More formally, denote by y_i the post-tax income for a participant, y_i^0 the pretax income and let $t \in [0,1]$ denote the tax rate. Post-tax income of participant *i* can then be written as:

$$y_i = (1-t)y_i^0 + \frac{t}{n}\sum_{j=1}^{n} y_j^0 = y_i^0 + t\left(\frac{1}{n}\sum_{j=1}^{n} y_j^0 - y_i^0\right) = y_i^0 + t(\bar{y} - y_i^0) \quad (1)$$

where $\bar{y} = \frac{1}{n}\sum_{j=1}^{n} y_j^0$ is the average income of the participants. Taking the derivative of (1) with respect to the tax rate gives,

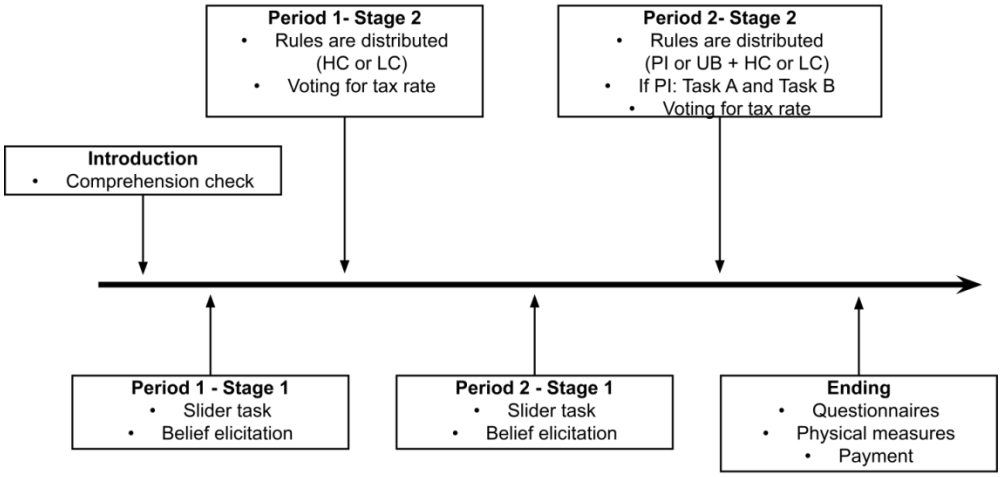
$$\frac{\partial y_i}{\partial t} < 0 \; if \; y_i^0 > \bar{y} \quad \text{and} \; \frac{\partial y_i}{\partial t} > 0 \; if \; y_i^0 < \bar{y}.$$

[*] We asked them the following question: "What do you believe your rank is compared to the others' in your group? (e.g. if you think you did 2nd best then type 2)

The post-tax income is thus decreasing (increasing) in the tax rate when the individual's income is above (below) the average income. As such, the benchmark model which abstracts from risk and social preferences predicts corner solutions where individuals either choose $t = 0$ or $t = 1$ (for a more exhaustive analysis including social preferences and risks, see Durante et al. (2014)). To avoid carry-over effects, the participants were not told what their final earnings were for period 1 and they had no information about the other participants' votes. This concluded period 1. Period 2 would follow right after and proceed in the same way. Earnings in both periods counted toward the final payment.

After both periods were finalized, a survey was shown on the monitor and another questionnaire was distributed to be completed in paper form (see Appendix). While the participants completed these forms, they were individually taken out to an adjacent lab where their physical measurements were taken. For explorative purposes, and in line with recent studies on preference formation, bargaining and formidability (Petersen et al., 2013; Nguyen et al., 2016), we include formidability (upper-body strength). The physical measures include height, weight, the circumference of the flexed bicep, chest and handgrip strength (dominant arm). Finally, the survey included items on various background information and ideology (see Esarey et al., 2012) (1 = left-leaning ideology and 7 = right-leaning ideology). The timing of the study is presented in Figure 1.

Figure 1 **Overview of the experimental design**



SUBSTANTIVE JUSTICE, PROCEDURAL INJUSTICE AND UNFAIR SHARE OF BURDEN

Substantive justice treatment: In a universal welfare system the goal is to not discriminate between "the needy" and "the poor". In a selective welfare system the debate revolves around what to do with "the others". As Rothstein (2001, p. 224) puts it:

"Under a universal system, in which the state furnishes all citizens with *basic capabilities*, the moral logic is altogether different. Since the universal welfare policy embraces all citizens, the debate assumes quite another character: social policy is now thought to concern the entire community, and the question becomes what, *from a general standpoint*, is a fair manner in which to organize social policy".

This "general standpoint" has recently been coined the deservingness heuristic (Petersen et al. 2011). According to this heuristic, individuals are seen as deserving of help if they are perceived as being without control over their own situation. This is better known as the luck vs. effort hypothesis. To the extent that income is a function of effort, low income earners should be perceived to be undeserving. In this study, this justice norm is

approached in the experimental design by having two conditions reflecting low versus high control (cf. substantive justice):

Low Control (LC): One group of participants is informed that 2 out of 5 in the group by *random selection* will lose all their pre-tax income.

High Control (HC): The other group is informed that the 2 participants who have *earned least* will lose their pre-tax earnings.

In both conditions the tax money that is collected is distributed evenly across all the group members.

***Procedural injustice treatment**:* Procedural injustice occurs if the mechanism by which welfare benefits are distributed is perceived as arbitrary. In a selective welfare system entitlements are granted conditionally on an individual assessment of the claimant, often based on seemingly arbitrary dimensions (e.g. quasi-judicial hearings on disability benefits in the U.S. or allowance of food stamps). In a universal system, however, individuals are entitled by some broad general criteria (e.g. being a citizen or being above age 65). In this experiment, the default is the universal system in which just being a member of the group entitles the individual to an (equal) share of the tax money collected.

In the *procedural injustice treatment,* participants are told that the experimenter will make a decision on disbursement of tax revenues based on two parameters: 1) Does the participant like Kandinsky or Klee paintings? This mechanism is used to assign group identity in psychology (Tajfel et al., 1971) and also implemented in economic experiments (Chen and Li, 2009; Ku and Salmon, 2013). In the present context, the participant is uncertain of whether she obtains the in- or outgroup status. Because there is no cue as to which category is the in-group, the procedure is expected to be seen as unfair. 2) After the slider task and voting, participants are asked to write an argument why they should receive tax money in case they lose their pre-tax earnings. Participants are told that both criteria may influence the experimenter's decision whether to disburse payments.

The experiment thus proceeds as following:

1) Earned income is revealed privately
2) Each participant makes a guess for his/her rank.
3) Each participant is shown 5 pairs of Kandinsky and Klee paintings and they make a choice between painting A or B for each pair.
4) Each participant writes up arguments for why, in case he loses his earned income, should be entitled to receive benefits that are collected through the tax money.
5) Two participants lose their earned income, either randomly (LC) or through their effort choice (HC)
6) Tax money is collected from all the participants under the chosen tax rate.
7) Between the two participants who each lost their income the experimenter now makes a decision on how to redistribute between the two of them the 2/n share of the tax revenue that the two of them collectively would receive if tax revenue was distributed lump sum. The experimenter has full discretion to allocate the amount in whichever he or she wants to including allocating everything to one person, but must spend the entire share (i.e. 2/n*T). The three remaining participants each receive (1/n*T).

Unfair share of burden treatment: In the last treatments, the *unfair burden treatment*, the person with the highest score in the slider task is exempted from paying taxes. He or she will still receive the same share of the collected tax money as the others though. Anecdotal evidence from the US suggests that the debate is not only about how much that is paid but also who pays them. For example, in 2011 around 57% of the respondents agreed to the statement that "What bothers you most… Feel wealthy people don't pay fair share" increasing the dissatisfaction from 51% in 2003 (Pew Research Center, 2011). The participant chosen can coincide with being the one who also lost his/her earned income.

The experiment in this treatment thus proceeds as following:

1) Earned income is revealed
2) Each participant makes a guess for his/her rank.
3) Two participants lose their earned income (either randomly or through their effort).
4) One person is randomly drawn to be exempted from paying any tax.
5) Tax money is collected from all the participants under the chosen tax rate.
6) All tax money is redistributed evenly (1/n of T).

The experimental sessions were conducted at Cognition and Behavior lab at Aarhus University and the main part of the experiment was programmed using z-Tree (Fishbacher, 2007). For each session I recruited between 5-20 participants and the experiment lasted about 60 mins in total. The study was posted online on the lab's website and advised that participants would receive 40 kroner as an appearance fee plus whatever they could earn during the experiment[†]. In total 135 individuals participated over 16 sessions during June -September 2016. Only 48% of the subjects were identified as Danish. This is noteworthy because it minimizes the risk that results are driven by respondents' being socialized into thinking that one set of institutional rules (mimicking the universal welfare state) is more appropriate or natural than the alternative set of institutional rules.

Table 1 **Overview of experimental sessions**

	HC+UB	**LC+UB**	**HC+PI**	**LC+PI**
Participants	45	35	30	30

[†] The show-up fee was increased in a few sessions (up to 60DKK) to compensate delays by experimenter

FINDINGS

This section presents the results from the experiments. Figure 2 displays the mean tax rate in the first period, where participants either participated in the high or low control treatment. It is evident that participants in the high control treatment were willing to pay much less than participants in the low control treatment: 7.3 percentage points to be exact. This is in accordance with the expectations in the substantive justice hypothesis. It is interesting to note that the participants not only responded to the treatment in a substantial – and substantially meaningful – way, but also the choice of tax rate is neither extremely low nor extremely high. Both these scenarios would have cast doubt on the intuition behind the argument of Rothstein.

Figure 2 **Mean tax rate in low and high control treatment**

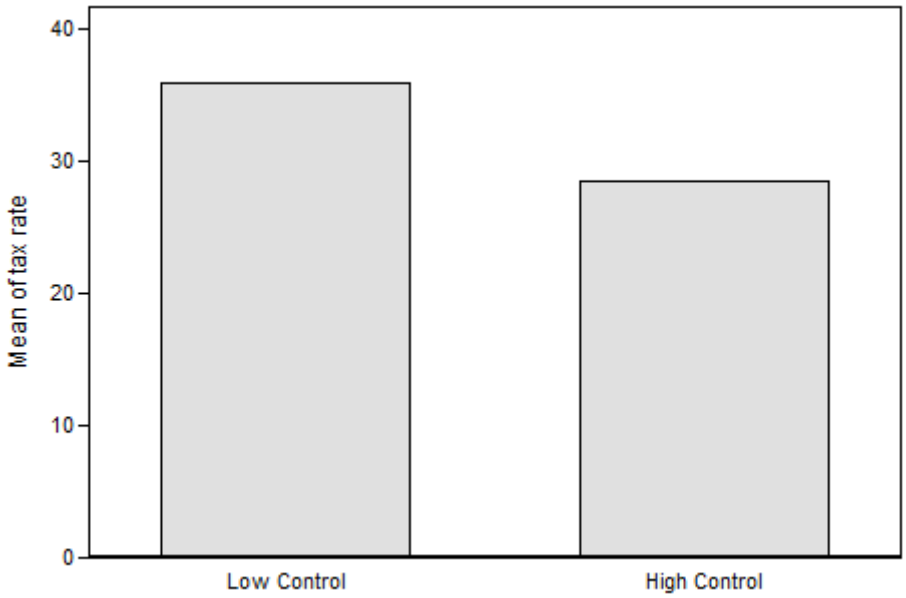


Table 2 presents a formal test of the high-low control treatment effect on the tax rate in the first period. In line with Durante et al. (2014), I employ tobit models for the estimation since my dependent variable is restricted between 0 and 100. It is, of course, very likely that neither ordinary citizens nor my participants would want a tax rate below 0 or above 100, but the tobit estimation nevertheless allows me to account for the eventuality. I use robust clustered standard errors (clustered around each session). It is worth

stressing, however, that the results I present in the following in both Table 2 and 3 are robust to different estimation choices.

I include four control variables in the estimation. The first control is physical formidability, which I previously have shown in the dissertation to matter for economic behavior. The second is the participant's belief about his or her own ranking in the game. The assumption is that the participants who feel they performed well will think they have less need for the insurance that a high tax rate provides. Third, political ideology should matter, with right-leaning participants being less willing to pay high tax compared to left-leaning participants. Fourth and finally, I control for risk preference using a well-documented risk measure that correlates with a wide range of behavioral and financial decisions (Dohmen et al., 2010).

The estimation reported in Table 2 supports the conclusion from Figure 1 (Model 1-2 are estimated using normal standard errors, while Model 3-4 are estimated using robust clustered standard errors at the session level). Indeed, in this setup the mean tax rate in the high control treatment is about 9 percentage point lower than the low control treatment. The four controls all predict tax rates. The more formidable a participant is, the higher the tax rate he or she prefers. Auxiliary analyses not reported here show that the effect of formidability primarily is located among the left-leaning participants. This is in line with Petersen et al. (2013) who show how formidability intensifies existing political attitudes. The worse a participant believes he or she performed, the higher the preferred tax rate. Finally, the more right-leaning and risk-averse the participant is, the lower the preferred tax rate is.

Table 2 **Results for the first period**

	Model 1	Model 2	Model 3	Model 4
High Control	-9.638	-9.034	-9.638	-9.034
	[5.145]	[5.367]	[4.468]	[5.278]
	(0.063)	(0.095)	(0.033)	(0.090)
Formidability (z-scored)		9.001		9.001
		[3.483]		[3.877]
		(0.011)		(0.022)
Belief (1-5)		9.217		9.217
		[2.697]		[3.425]
		(0.001)		(0.008)
Ideology		-5.695		-5.695
		[1.897]		[2.721]
		(0.003)		(0.039)
Risk		2.050		2.050
		[1.128]		[0.718]
		(0.072)		(0.005)
Constant	36.171	21.093	36.171	21.093
	[3.689]	[13.674]	[3.450]	[18.428]
	(0.000)	(0.126)	(0.000)	(0.255)
Observations	135	108	135	108
Uncensored observations	114	91	114	91
Left-censored observations	13	11	13	11
Right-censored observations	8	6	8	6
Log-likelihood	-573.221	-444.684	-573.221	-444.684
Pseudo-R^2	0.003	0.032	0.003	0.025

Note: Dependent variable is the tax rate (0-100). Two-censored tobit regressions, Marginal effects are reported. For Model 1 and 2, standard errors are reported in brackets. In Model 3 and 4, robust clustered standard errors at session level are reported in the brackets. Two-tailed p-values are reported in parenthesis.

Figure 3 **Mean tax rate across the four treatments in Period 2**

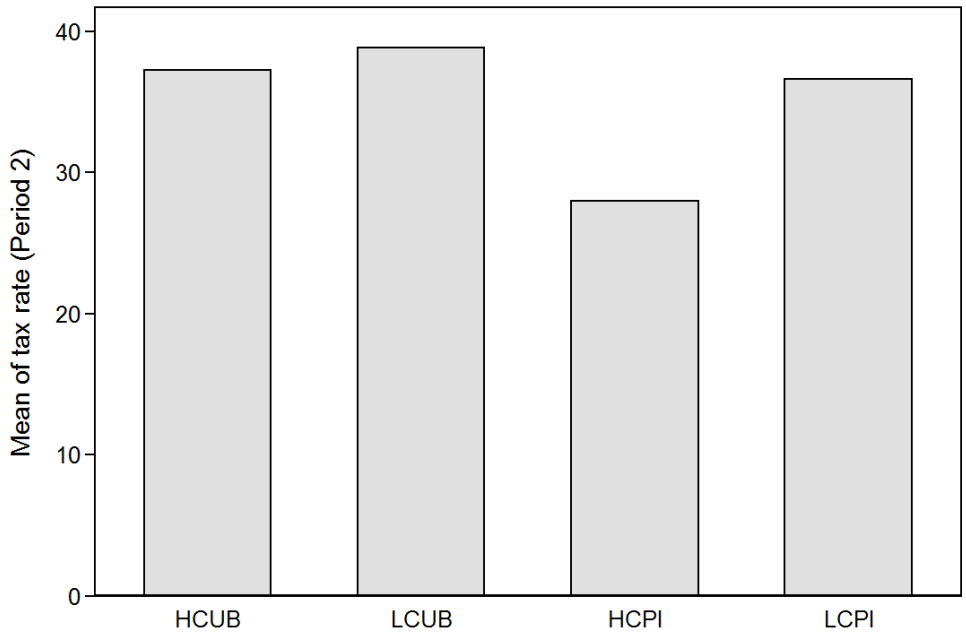


Figure 3 presents the mean tax rate across the four treatments. It is evident from Figure 3 that the High Control + Procedural Injustice Treatment yields substantially lower tax rates. We test this formally in Table 3. A dummy for the procedural injustice treatment captures the treatment effect of the unfair share of burden treatment versus the procedural injustice treatment. Furthermore, in Table 3, the analysis is split such that Model 1 and 3 are estimated for those participants who were in the High Control treatment while the Model 2 and 4 are for the Low Control. I once again employ a tobit regression with clustered robust standard errors. For benchmark, the same models were estimated using OLS-regression yielding substantially the same results (see Appendix A).

Table 3 **Results for the second period**

	Model 1 High Control	Model 2 Low Control	Model 3 High Control	Model 4 Low Control
Procedural Injustice Treatment	-8.887	-5.367	-13.595	-8.639
	[4.983]	[5.707]	[5.251]	[5.689]
	(0.077)	(0.349)	(0.030)	(0.132)
Formidability (z-scored)			4.711	13.011
			[3.663]	[3.882]
			(0.201)	(0.001)
Belief (1-5)			3.601	6.812
			[2.952]	[2.340]
			(0.225)	(0.004)
Ideology			-4.530	-8.394
			[2.041]	[2.153]
			(0.029)	(0.000)
Risk			2.315	2.401
			[1.361]	[1.133]
			(0.092)	(0.036)
Constant	33.538	39.673	30.416	40.044
	[3.262]	[4.048]	[13.795]	[14.310]
	(0.001)	(0.001)	(0.030)	(0.006)
Observations	140	130	102	114
Uncensored observations	119	110	86	96
Left-censored observations	15	7	12	7
Right-censored observations	6	13	4	11
Log-likelihood	-594.01	-563.60	-427.12	-475.84
Pseudo-R^2	0.003	0.001	0.015	0.033

Note: Dependent variable is the tax rate (0-100). Two-censored tobit regressions. Marginal effects are reported. Standard errors are reported in brackets. Two-tailed p-values are reported in parenthesis.

Within the high control condition adding procedural injustice significantly lowers the preferred tax rates by 8.9 percentage points (Model 1). There is no corresponding significant effect for low control (Model 2). Adding the controls in Model 3 and 4 does not qualitatively alter this conclusion although a larger treatment effect is found when control variables are added. The findings, firstly, suggest that the high-low control treatments are so powerful that they structure subsequent responses to new treatments. The participants who got the high control treatment – meant to make them believe the rules of redistribution were substantially unfair – were also susceptible to the procedural injustice treatment that effectively re-emphasized the distaste for redistributive policies. In contrast, if the participants had been made to believe that the rules were substantially fair, the procedural injustice treatments did not make a difference. Second, the

unfair burden treatment is clearly less important than the procedural injustice as the coefficient of the procedural injustice dummy is always negative (and in two instances, in Model 1 and 3, significantly so). All in all, this hints that substantial justice is more important for preferred tax rate than procedural justice, which, in turn, is more important than an unfair burden. It ought to be stressed, however, that this conclusion is likely to be sensitive to the exact experimental design including the sequence by which participants are exposed to the three rules.

As noted, the results are robust even after controlling for individual differences on formidability, belief, ideology and risk preferences. In the low control condition, all the included controls are significant predictors. Participants who were stronger vote for more redistribution while right-leaning individuals vote for less. Beliefs about the individual's ranking also influenced their tax choices such that higher confidence in one's ranking led to less positive attitudes towards redistribution. Finally, being more risk prone led to higher tax vote.[‡] The findings suggest that the high control treatment was so powerful that it cancelled-out the effect of formidability and ideology. This aligns with Petersen et al. (2011) who show that the perception that a welfare claimant is undeserving of help (equivalent to my high control treatment), crowds out the effect of personal values. Formidability and ideology structures participants' attitudes in general, but this general predisposition is trumped when informed that the tax revenue will be allocated in a substantially unjust way.

DISCUSSION

The study potentially helps policy-makers in understanding how to create policies that can be expected to be supported by the general public. The findings reported above broadly provide support to Rothstein's proposition that the organization of the welfare state matters for citizens' willingness to pay taxes. Rothstein located three fairness norms that welfare states can either appease or conflict with. In this paper, I test Rothstein's argument

[‡] This is a surprising effect, but additional analyses show that risk preferences interact with effort (i.e. ability), so that higher effort reduces tax vote, but only for the risk adverse.

using a lab experiment with 135 participants, allowing me to isolate the effect of individual rules. Substantial and procedural fairness mattered for the tax rates that people voted for, whereas the unfair burden, at least in my setup, did not matter much.

There are, however, clear limits to the conclusions that can be drawn based on the experiment I have reported here. I want to highlight two limitations in particular. First, while I have tried to mimic the rules outlined by Rothstein (1998; 2001) an experiment will never be able to fully mirror real-world welfare states. Additional experiments will need to be conducted to study how the exact formulation of rules affects tax votes. Second, the experiment was designed as between-subjects design. It is not suited to study with-in subject preference changes. While I believe that my experiment is a valid and valuable starting-point, it is obvious that a within-subject design would be a relevant supplement in future research. The challenge is of course to avoid experimenter demand effects that arise from exposing the same person to two treatments that manipulate a particular norm. Third, and as hinted above, to isolate the precise effect of individual rules probably requires much more elaborate testing of the effect of the sequence by which participants are exposed to specific rules (that is, so-called order effects).

REFERENCES

Ackert, L. F., Martinez-Vazquez, J., & Rider, M. (2007). Social preferences and tax policy design: some experimental evidence. *Economic Inquiry*, *45*(3), 487-501.

Alesina, A., & Angeletos, G. M. (2005). Fairness and redistribution. *The American Economic Review*, *95*(4), 960-980.

Alesina, A., Glaeser, E., & Sacerdote, B. (2001). Why Doesn't the United States Have a European-Style Welfare State? *Brookings Papers on Economic Activity*, *2001*(2), 187-254.

Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., & Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *The American Economic Review*, *97*(3), 818-827.

Cappelen, A. W., Konow, J., Sørensen, E. Ø., & Tungodden, B. (2013). Just luck: An experimental study of risk-taking and fairness. *The American Economic Review*, *103*(4), 1398-1413.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *The American Economic Review*, *99*(1), 431-457.

Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *The American Economic Review*, *92*(4), 1218-1221.

Durante, R., Putterman, L., & Weele, J. (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association*, *12*(4), 1059-1086.

Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *The American Economic Review*, *94*(4), 857-869.

Esarey, J., Salmon, T. C., & Barrilleaux, C. (2012). What Motivates Political Preferences? Self-Interest, Ideology, and Fairness in a Laboratory Democracy. *Economic Inquiry*, *50*(3), 604-624.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The journal of economic perspectives*, *14*(3), 159-181.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, *10*(2), 171-178.

Green, D. P., & Gerber, A. S. (2002). The downstream benefits of experimentation. *Political Analysis*, *10*(4), 394-402.

Höchtl, W., Sausgruber, R., & Tyran, J. R. (2012). Inequality aversion and voting on redistribution. *European economic review*, *56*(7), 1406-1421.

Kahan, D. M. (2005). The logic of reciprocity: Trust, collective action and law. In H. Gintis, S. Bowles, R. Boyd, & E. Fehr (Eds.), *Moral sentiments and material interests: The foundation of cooperation in economic life*. Cambridge, MA: The MIT Press.

Kataria, M., & Montinari, N. (2012). Risk, Entitlements and Fairness Bias: Explaining Preferences for Redistribution in Multi-person Setting. *Jena Economic Research Papers*, *2012*, 061.

Krawczyk, M. (2010). A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics*, *94*(1), 131-141.

Ku, H., & Salmon, T. C. (2013). Procedural fairness and the tolerance for income inequality. *European Economic Review*, *64*, 111-128.

Larsen, C. A. (2008). The institutional logic of welfare attitudes: How welfare regimes influence public support. *Comparative political studies*, *41*, 145-168

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a Sword: Self-governance Is Possible. *American political science Review*, *86*(02), 404-417.

Petersen, M. B., Slothuus, R., Stubager, R., & Togeby, L. (2011). Deservingness versus values in public opinion on welfare: The automaticity of the deservingness heuristic. *European Journal of Political Research*, *50*(1), 24-52.

Petersen, M. B., Sznycer, D., Sell, A., Cosmides, L., & Tooby, J. (2013). The ancestral logic of politics upper-body strength regulates men's assertion of self-interest over economic redistribution. *Psychological Science*, *24*(7), 1098-1103.

Pew Research Center (2011). Tax System Seen as Unfair, in Need of Overhaul. Accessed online: http://www.people-press.org/2011/12/20/tax-system-seen-as-unfair-in-need-of-overhaul/ on Sep 25th 2016.

Prowse, V., & Gill, D. (2009). *A Novel Computerized Real Effort Task Based on Sliders* (No. 435).

Roosma, F., Van Oorschot, W., & Gelissen, J. (2015). A Just Distribution of Burdens? Attitudes Toward the Social Distribution of Taxes in 26 Welfare States. *International Journal of Public Opinion Research*, edv020.

Rothstein, B. (1998). *Just institutions matter: the moral and political logic of the universal welfare state*. Cambridge University Press.

Rothstein, B. (2001). The universal welfare state as a social dilemma. *Rationality and society*, *13*(2), 213-233.

Saporiti, A., & Tohmé, F. (2006). Single-crossing, strategic voting and the median choice rule. *Social Choice and Welfare*, *26*(2), 363-383.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European journal of social psychology*,*1*(2), 149-178.

Tyran, J. R., & Sausgruber, R. (2006). A little fairness may induce a lot of redistribution in democracy. *European Economic Review*, *50*(2), 469-485.

APPENDIX A

	Model 1 High Control	Model 2 Low Control	Model 3 High Control	Model 4 Low Control
Procedural Injustice Treatment	-7.708	-4.498	-11.748	-6.620
	4.312	4.898	5.437	4.953
	(0.076)	(0.360)	(0.033)	(0.184)
Formidability (z-scored)			4.537	10.550
			3.203	3.338
			(0.160)	(0.002)
Belief (1-5)			2.767	5.787
			2.604	2.028
			(0.291)	(0.005)
Ideology			-3.420	-6.736
			1.776	1.833
			(0.057)	(0.000)
Risk			1.840	2.251
			1.199	0.996
			(0.128)	(0.026)
Constant	34.225	38.914	31.877	35.990
	2.823	3.328	12.120	12.389
	(0.000)	(0.000)	(0.010)	(0.004)
N	140	130	102	114

Note: Dependent variable is the tax rate (0-100). Beta coefficients are reported with standard errors in brackets.
Two-tailed p-values are reported in parenthesis.

APPENDIX B - INSTRUCTIONS

Welcome! You are about to participate in an experiment funded by Aarhus University. Please switch off your mobile phone and remain quiet. It is strictly forbidden to talk to the other participants. Whenever you have a question, please raise your hand and one of the experimenters will come to your aid.

You will receive 40 kr. for showing up on time and participating. Besides this, you can earn more. The show-up fee and any additional amounts of money you may earn will be paid to you through your NemKonto. We will therefore ask for your CPR-number later. Payments are carried out privately, i.e., the others will not see your earnings.

During the experiment we shall speak of ECUs (Experimental Currency Unit) rather than Kroner.
The conversion rate between them is 100 ECUs = 25 kr. This means that for each ECU you earn you will receive 0.25 kr.

We expect the duration of the experiment to last approximately 50 minutes. In case the experiment runs a bit late, we kindly ask for your understanding. Your participation is very valuable to us.

Detailed instructions

The experiment has two periods, followed by a questionnaire at the end of the experiment. You will be randomly grouped with four other participants and the five of you will remain a group for the rest of the experiment. The choices of the other group members may affect what you earn today. What happens in the other groups does not affect your earnings.

Each of the two periods of the experiment consists of two stages. In stage 1 you perform some tasks to earn initial earnings. In stage 2 all group members vote for a tax rate and final earnings are determined. At the end of the experiment we ask you to fill out a questionnaire while waiting to do a few short tasks in an adjacent room.

Your total earnings for today will be the sum of the earnings in period 1 and the earnings in period 2 as well as the 30 kr. show-up fee.

We now describe how the stages in each period unfold.

STAGE 1: Accruing initial earnings:

In stage 1 you have the opportunity to earn money by performing a task that requires you to move sliders to a certain position. Each slider has integer values from 0 to 100 and will have a start position at 0. Your task is to position each slider at exactly at the 50 position by dragging the slider bar with your mouse. See figure 1 for an example where the slider is currently positioned at 21.

Figure 1

You will have 10 minutes to solve as many sliders as you want to. Your initial earnings increase by 1 ECU for each slider that is located at position 50 at the end of the 10 minutes.

Specifically, you will have access to four screens with sliders. You can move back and forth between each screen via two buttons at the top of the screen. The sliders can be adjusted and readjusted an unlimited number of times and the current position is displayed to the right of each slider.

At the end of stage 1 you assess how well you did compared to the other four members of your group. We ask you to guess what your performance rank is in

your group. If you think you are the one who solved most sliders in this period and therefore are ranked 1 in your group, you type in "1" and press the "OK"-button. If you think you are the worst performer in your group, you type in "5" and press "OK" and similarly for the other ranks. If you correctly guess your actual rank in this period, we will add an additional 5 kr. to your earnings in the experiment today. In the unlikely case that several participants should solve the exact same amount of sliders in this period, the actual ranking between these participants is decided by a random draw.

STAGE 2: Voting on the tax rate

Two participants in each group will experience an "earnings shock" that resets to zero their initial earnings from stage 1 (the details are explained later in the experiment). In stage 2, you and the other members of your group vote for a tax rate. Below we describe how the tax rate that applies in your group is determined based on these votes. Participants pay taxes out of their initial earnings according to this tax rate. The total amount of taxes collected goes to the group account, from which all 5 group members will receive a share. That is, your earnings in the period are your initial earnings minus the tax payment (your initial earnings * tax rate), plus your share of the group account. To summarize:

Two members in each group suffer an "earnings shock" that resets their initial earnings to zero

Your earnings for the period= your initial earnings – tax payment + your share of the group account

If you suffered an earnings **shock, your earnings for the period= your share of the group account**

HOW THE TAX RATE IS DETERMINED:

The tax rate for your group is chosen by voting. You and the other group members each vote on the tax rate he or she wants to apply to the group by choosing a tax rate between 0% and 100%. For example 0% means that no taxes should be collected and 100% means that all initial earnings should flow into the group account and be distributed among the group members.

The tax rate that applies for your group is the median vote in your group. That is, the tax rate that applies is the tax rate chosen by the voter in the exact middle of a ranking of the five group members. **This simply means that you should vote for whatever tax rate you believe should be applied to the whole group.**

Example: Imagine that these are the tax rates chosen by each participant:

- Participant A = 37%
- Participant B = 69%
- Participant C = 57%
- Participant D = 27%
- Participant E = 10%

To find the median voter we order the chosen tax rates from lowest to highest:

Participant E = 10%

Participant D = 27%

Participant A = 37% ← middle ranked: i.e. exactly the same number of votes above and below

Participant C = 57%

Participant B = 69%

The median vote is then Participant A (37%) since there are two above and two below that specific tax rate. The tax rate that applies to the group in this case is 37%. That is, each group member pays a tax of 0.37*initial earnings into the group account.

Suppose now that Participant D had voted differently, and cast a vote for the tax rate to be 42%. Then his vote would be the median vote and thus the tax rate that applies to the group.

Participant E = 10%

Participant A = 37%

Participant D = 42% ← middle ranked: i.e. exactly the same number of votes above and below

Participant C = 57%

Participant B = 69%

Summary

Here is a summary of the experiment.

Period 1
Stage 1:
- Accrue initial earnings by performing the slider task
- Performance guess

Stage 2:
- Two participants are hit by an "earnings shock" that rests to 0 their initial earnings
- The precise rules for tax collection and the distribution of the group account are announced
- Every participant casts a vote for a tax rate between 0% and 100%The median vote in a group determines what tax rate is applied for the group
- Final earnings for the period are calculated

Period 2: (has the same structure as period 1)
Stage 1
- Accrue initial earnings by performing the slider task
- Performance guess

Stage 2:
- Two participants are hit by an "earnings shock" that rests to 0 their initial earnings
- The precise rules for tax collection and the distribution of the group account are announced
- Every participant casts a vote for a tax rate between 0% and 100%The median vote in a group determines what tax rate is applied for the group
- Final earnings for the period are calculated

At the end of the experiment:
- Leave the room and fill out questionnaire in the hallway. Do not communicate with the other participants
- Wait until your participant ID is called and then perform a few simple tasks with the experimenter
- Return to your computer and fill in a few questions and your CPR number
- Receive a receipt at the counter

We will now ask you a few check-up questions to check that you have understood the rules. Once you finished answering the questions, please raise your hand so an experimenter can come by. Once everybody has finished answering the questions we will start period 1 of the experiment.

Thank you for your contribution to science!

Check-up questions:

Question 1: Below you see a table of votes for a tax rate. Please determine the tax rate that would apply in this group.

- Participant A = 40%
- Participant B = 25%
- Participant C = 11%
- Participant D = 44%
- Participant E = 60%

Answer: The tax rate that applies in this group is:

_____

Question 2: Suppose that the tax rate in your group is 10% and that your initial earnings are 60. Two other group members earned a total of 200 (100 each). The final two group members were hit by an "earnings shock" and had their initial earnings reduced to zero. Further, suppose that each participant receives an equal share of the group account. That is, you receive 1/5 of the group account.

Please calculate the following:

Your tax payment is:

The total tax payments (the amount in the group account) are:

The amount you receive out of the group account is:

Your final earnings for the period:

Once you have finished the questions please raise your hand and an experimenter will check your answers!

APPENDIX C – HIGH CONTROL

Announcement: Stage 2 Period 1

Two participants in each group will experience an "earnings shock" that resets to 0 their initial earnings from stage 1. The earnings shock will hit the two participants who solved the lowest number of sliders in their group. We will call them the income losers. That is, ***you will lose all of your initial earnings if you are among the two worst performers on the slider task in stage 1 of this period.***

You now vote for a tax rate. The final tax rate that is applied in your group follows the rules explained earlier. That is, each participant votes for a tax rate and the median vote is the tax rate that is applied. The total amount of taxes collected goes into the group account from which all 5 group members will receive an equal share.
This means:

- All group members pay taxes based on their initial income: All pay initial earnings*tax rate into the group account
- The initial earnings by the two income losers are reset to 0. Hence, their contribution to the group account is also 0
- All group members, including the income losers, receive an **equal share** of the group account. That is, each group member gets 1/5 of the group account balance

Once everyone has cast their vote, the earnings for this period are calculated for all participants and the experiment continues.

You will not receive any feedback about your performance in the slider task, the tax rate, and your earnings until the very end of the experiment.

Example: Suppose that you have initial earnings of 30 ECU. Further assume that the initial earnings of the other group members in stage 1 are:

- Participant B: 25 ECU
- Participant C: 50 ECU

- Participant D: 60 ECU
- Participant E: 65 ECU

Since you and Participant B have the lowest scores in stage 1, both of you lose your initial earnings and thus have only 0 ECU.

The chosen tax rate is 25%. Hence the total amount transferred to the group account is:
(0 + 0 + 50 + 60 + 65) * 0.25 = 43.75 ECU.
Each participant is entitled to an equal share of the group account. That is, each group member gets 1/5*43.75= 8.75 ECU.

Hence, your final earnings for this period are:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
8.75 ECU = 0 – 0 + 43.75/5

Example 2: Suppose now that you were Participant C in the example above. Then you pay 50*0.25=12.5 ECU taxes into the group account and receive 8.75 ECU back from the group account. That is your final earnings for this period would be:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
46.25 ECU = 50 – 12.5 + 43.75/5

APPENDIX D – LOW CONTROL

Announcement: Stage 2 Period 1

Two participants in each group will experience an "earnings shock" that resets to 0 their initial earnings from stage 1. The earnings shock will hit two participants based on a random draw by the computer. That is, there is 1/5 chance that you will be hit by an "earnings shock" and lose all you initial earnings. We will call them the income losers. That is, ***you will lose all of their generated income if you are among the two participants randomly chosen by the computer.***

You now vote for a tax rate. The final tax rate that is implemented follows the rules explained earlier. That is, each participant vote for a tax rate and the median vote is the tax rate that is applied. The total amount of taxes collected goes to the group account from which all 5 group members will receive an equal share.
This means:

- All group members pay taxes based on their initial income: All pay initial earnings*tax rate into the group account
- The initial earnings by the two income losers are reset to 0. Hence, their contribution to the group account is also 0
- All group members, including the income losers, receive an **equal share** of the group account. That is, each group member gets 1/5 of the group account balance

Once everyone has cast their vote, the earnings for this period are calculated for all participants and the experiment continues.

You will not receive any feedback about your performance in the slider task, the tax rate, and your earnings until the very end of the experiment.

Example: Suppose that you have initial earnings of 30 ECU. Further assume that the initial earnings of the other group members in stage 1 are:

- Participant B: 25 ECU
- Participant C: 50 ECU

- Participant D: 60 ECU
- Participant E: 65 ECU

Based on a random draw by the computer you and Participant B lose your initial earnings and thus have only 0 ECU.

The chosen tax rate is 25%. Hence the total amount transferred to the group account is:
(0 + 0 + 50 + 60 + 65) * 0.25 = 43.75 ECU.
Each participant is entitled to an equal share of the group account. That is, each group member gets 1/5*43.75= 8.75 ECU.

Hence, your final earnings for this period are:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
8.75 ECU = 0 – 0 + 43.75/5

Example 2: Suppose now that you were Participant C in the example above. Then you pay 50*0.25=12.5 ECU taxes into the group account and receive 8.75 ECU back from the group account. That is your final earnings for this period would be:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
46.25 ECU = 50 – 12.5 + 43.75/5

APPENDIX D – HIGH CONTROL + PROCEDURAL INJUSTICE

Announcement: Stage 2 Period 2

Two participants in each group will experience an "earnings shock" that resets to 0 their initial earnings from stage 1. The earnings shock will hit the two participants who solved the lowest number of sliders in their group. We will call them the income losers. That is, ***you will lose all of your initial earnings if you are among the two worst performers on the slider task in stage 1 of this period.***

You now vote for a tax rate. The final tax rate that is applied in your group follows the rules explained earlier. That is, each participant votes for a tax rate and the median vote is the tax rate that is applied. The total amount of taxes collected goes into the group account from which all 5 group members will receive an equal share. This is identical to the previous period.

However, the procedure now is different for distributing among the group members the amount of taxes collected in the group account. Each group member is entitled to an equal share of the taxes collected in the group account. However, the experimenter has discretion how to distribute the shares of the two group members who suffered an earnings shock and he may base this decision on some information that you provide now (see below).

That is:

- For those group members who did not suffer an income shock everything is as before. They each receive 1/5 of the taxes collected in the group account.
- For the two group members who suffered an income shock the rules are different. They will together be entitled to 2/5 of the amount in the group account. The experimenter has discretion how to distribute this share among the two group members.

Your answers in the two following tasks may influence whether you will receive any of your entitled share of the group account in case you suffer an earnings shock and your initial income is reset to zero.

Task A: You will be presented with five pairs of paintings from two modern painters, Klee and Kandinsky. Each pair has one painting represented by each painter. You will be asked to choose your preferred painting. Based on your stated preferences you will be classified into two groups. Only the experimenter sees your choices.

Task B: You have the option to provide some arguments for why you should, in case you lose your income, be entitled to receive compensation. You will have 3 minutes to write this statement. Your answer is only read by the experimenter.

The experimenter will review the answers and make a decision on whether an income loser will receive any compensation from the group account. The experimenter can freely dispose over both income losers' share of the group account. Since each participant is potentially distributed 1/5 of the group account, the experimenter can freely distribute their overall entitlement to 2/5 of the group account between them. As an example, if the group account has a balance of 200, the two income losers are jointly entitled to 2/5*200=40 ECU. The experimenter can then distribute 40 ECUs between the two participants in whatever way he wants. For example, if you are one of the income losers, he may give 0 to you and 40 to the other; or you both may get 20; or you may get 34 and the other 6. The total amount distributed must sum to 2/5 of the group account.

There will be no opportunity to learn about why and how the experimenter decided the allocation.

This means:

- All group members pay taxes based on their initial earnings: All pay initial earnings*tax rate into the group account
- The initial earnings by the two income losers are reset to 0. Hence, their contribution to the group account is also 0
- All group members who did not suffer an income shock receive an **equal share** of the group account. That is, each group member gets 1/5 of the group account balance.
- The two income losers receive a share of the group account based on the experimenter's discretion.

Once everyone has cast their vote, the earnings for this period are calculated for all participants and the experiment continues.

You will not receive any feedback about your performance in the slider task, the tax rate, and your earnings until the very end of the experiment.

APPENDIX E – HIGH CONTROL + UNFAIR SHARE OF BURDEN

Announcement: Stage 2 Period 2

Two participants in each group will experience an "earnings shock" that resets to 0 their initial earnings from stage 1. The earnings shock will hit the two participants who solved the lowest number of sliders in their group. We will call them the income losers. That is, ***you will lose all of your initial earnings if you are among the two worst performers on the slider task in stage 1 of this period.***

You now vote for a tax rate. The final tax rate that is applied in your group follows the rules explained earlier. That is, each participant votes for a tax rate and the median vote is the tax rate that is applied. The total amount of taxes collected goes into the group account from which all 5 group members will receive an equal share. This is identical to the previous period.

However, the person with the highest score in stage 1 in this period is now exempted from paying any tax. The participant who is exempted from paying tax still receives his/her share of the group account.

This means:

- The initial earning by the income losers is reset to 0. Hence, their contribution to the group account is also 0.
- All group members except the highest performer pay taxes based on their initial income: All pay initial earnings*tax rate into the group account. Hence, the participant with the highest score in stage 1 in the current period does not pay any tax of his initial earnings.
- All group members, including the income losers, receive an **equal share** of the group account. That is, each group member gets 1/5 of the group account balance

Once everyone has cast their vote, the earnings for this period are calculated for all participants and the experiment continues.

You will not receive any feedback about your performance in the slider task, the tax rate, and your earnings until the very end of the experiment.

Example: Suppose that you have initial earnings of 30 ECU. Further assume that the initial earnings of the other group members in stage 1 are:

- Participant B: 25 ECU
- Participant C: 50 ECU
- Participant D: 60 ECU
- Participant E: 65 ECU

Since you and Participant B have the lowest scores in stage 1, both of you lose your initial earnings and thus have only 0 ECU.
The chosen tax rate is 25%. Hence the total amount transferred to the group account is:
(0 + 0 + 50 + 60 + 0) * 0.25 = 27.50 ECU.
Each participant is entitled to an equal share of the group account. That is, each group member gets 1/5*27.50= 5.50 ECU.

Hence, your final earnings for this period are:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
5.50 ECU = 0 – 0 + 27.5/5

Example 2: Suppose now that you were Participant E in the example above. Because you are the highest performer, then you pay 0*0.25=0 ECU taxes into the group account and receive 5.50 ECU back from the group account. That is your final earnings for this period would be:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
70.5 ECU = 65 – 0 + 27.5/5

APPENDIX F – LOW CONTROL + PROCEDURAL INJUSTICE

Announcement: Stage 2 Period 2

Two participants in each group will experience an "earnings shock" that resets to 0 their initial earnings from stage 1. The earnings shock will hit two participants based on a random draw by the computer. That is, there is 1/5 chance that you will be hit by an "earnings shock" and lose all you initial earnings. We will call them the income losers. That is, ***you will lose all of their generated income if you are among the two participants randomly chosen by the computer.***

You now vote for a tax rate. The final tax rate that is implemented follows the rules explained earlier. That is, each participant vote for a tax rate and the median vote is the tax rate that is applied. The total amount of taxes collected goes to the group account from which all 5 group members will receive an equal share.This is identical to the previous period.

However, the procedure now is different for distributing among the group members the amount of taxes collected in the group account. Each group member is entitled to an equal share of the taxes collected in the group account. However, the experimenter has discretion how to distribute the shares of the two group members who suffered an income shock and he may base this decision on some information you provide now (see below).

That is:

- For those group members who did not suffer an income shock everything is as before. They each receive 1/5 of the taxes collected in the group account.
- For the two group members who suffered an income shock the rules are different. They will together be entitled to 2/5 of the amount in the group

account. The experimenter has discretion how to distribute this share among the two group members.

Your answers in the two following tasks may influence whether you will receive any your entitled share of the group account in case you suffer an income shock and your initial income is reset to zero.

Task A: You will be presented by five pairs of paintings from two modern painters, Klee and Kandinsky. Each pair has one painting represented by each painter. You will be asked to choose your preferred painting. Based on your stated preferences you will be classified into two groups. Only the experimenter sees your choices.
Task B: You have the option to provide some arguments for why you should, in case you lose your income, be entitled to receive compensation. You will have 3 minutes to write this statement. Your answer is only read by the experimenter.

The experimenter will review the answers and make a decision on whether an income loser will receive any compensation from the group account. The experimenter can freely dispose over both income losers' share of the group account. Since each participant is potentially distributed 1/5 of the group account, the experimenter can freely distribute their overall entitlement to 2/5 of the group account between them. As an example, if the group account has a balance of 200, the two income losers are jointly entitled to 2/5*200=40 ECU. The experimenter can then distribute 40 ECUs between the two participants in whatever way he wants. For example, if you are one of the income losers, he may give 0 to you and 40 to the other; or you both may get 20; or you may get 34 and the other 6. The total amount distributed must sum to 2/5 of the group account.

There will be no opportunity to learn about how the experimenter decided the allocation.

This means:

- All group members pay taxes based on their initial earnings: All pay initial earnings*tax rate into the group account
- The initial earnings by the two income losers are reset to 0. Hence, their contribution to the group account is also 0
- All group members who did not suffer an income shock receive an **equal share** of the group account. That is, each group member gets 1/5 of the group account balance.
- The two income losers receive a share of the group account based on the experimenter's discretion.

Once everyone has cast their vote, the earnings for this period are calculated for all participants and the experiment continues.

You will not receive any feedback about your performance in the slider task, the tax rate, and your earnings until the very end of the experiment.

APPENDIX G – LOW CONTROL + UNFAIR SHARE OF BURDEN

Announcement: Stage 2 Period 2

Two participants in each group will experience an "earnings shock" that resets to 0 their initial earnings from stage 1. The earnings shock will hit two participants based on a random draw by the computer. That is, there is 1/5 chance that you will be hit by an "earnings shock" and lose all you initial earnings. We will call them the income losers. That is, ***you will lose all of their generated income if you are among the two participants randomly chosen by the computer.***

You now vote for a tax rate. The final tax rate that is implemented follows the rules explained earlier. That is, each participant vote for a tax rate and the median vote is the tax rate that is applied. The total amount of taxes collected goes to the group account from which all 5 group members will receive an equal share. This is identical to the previous period.

However, the person with the highest score in stage 1 in this period is now exempted from paying any tax. The participant who is exempted from paying tax still receives his/her share of the group account.

This means:

- The initial earning by the income losers is reset to 0. Hence, their contribution to the group account is also 0.
- All group members except the highest performer pay taxes based on their initial income: All pay initial earnings*tax rate into the group account. Hence, the participant with the highest score in stage 1 in the current period does not pay any tax of his initial earnings.
- All group members, including the income losers, receive an **equal share** of the group account. That is, each group member gets 1/5 of the group account balance

Once everyone has cast their vote, the earnings for this period are calculated for all participants and the experiment continues.

You will not receive any feedback about your performance in the slider task, the tax rate, and your earnings until the very end of the experiment.

Example: Suppose that you have initial earnings of 30 ECU. Further assume that the initial earnings of the other group members in stage 1 are:

- Participant B: 25 ECU
- Participant C: 50 ECU
- Participant D: 60 ECU
- Participant E: 65 ECU

Based on a random draw by the computer you and Participant B lose your initial earnings and thus have only 0 ECU.

The chosen tax rate is 25%. Hence the total amount transferred to the group account is:
(0 + 0 + 50 + 60 + 0) * 0.25 = 27.50 ECU.
Each participant is entitled to an equal share of the group account. That is, each group member gets 1/5*27.50= 5.50 ECU.

Hence, your final earnings for this period are:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
5.50 ECU = 0 – 0 + 27.5/5

Example 2: Suppose now that you were Participant E in the example above. Because you are the highest performer, then you pay 0*0.25=0 ECU taxes into the group account and receive 5.50 ECU back from the group account. That is your final earnings for this period would be:

Your earnings for the period = your initial earnings – tax payment + equal share of the group account
70.5 ECU = 65 – 0 + 27.5/5

DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS

AARHUS UNIVERSITY
SCHOOL OF BUSINESS AND SOCIAL SCIENCES
www.econ.au.dk

PhD dissertations since 1 July 2011

2011-4 Anders Bredahl Kock: Forecasting and Oracle Efficient Econometrics

2011-5 Christian Bach: The Game of Risk

2011-6 Stefan Holst Bache: Quantile Regression: Three Econometric Studies

2011:12 Bisheng Du: Essays on Advance Demand Information, Prioritization and Real Options in Inventory Management

2011:13 Christian Gormsen Schmidt: Exploring the Barriers to Globalization

2011:16 Dewi Fitriasari: Analyses of Social and Environmental Reporting as a Practice of Accountability to Stakeholders

2011:22 Sanne Hiller: Essays on International Trade and Migration: Firm Behavior, Networks and Barriers to Trade

2012-1 Johannes Tang Kristensen: From Determinants of Low Birthweight to Factor-Based Macroeconomic Forecasting

2012-2 Karina Hjortshøj Kjeldsen: Routing and Scheduling in Liner Shipping

2012-3 Soheil Abginehchi: Essays on Inventory Control in Presence of Multiple Sourcing

2012-4 Zhenjiang Qin: Essays on Heterogeneous Beliefs, Public Information, and Asset Pricing

2012-5 Lasse Frisgaard Gunnersen: Income Redistribution Policies

2012-6 Miriam Wüst: Essays on early investments in child health

2012-7 Yukai Yang: Modelling Nonlinear Vector Economic Time Series

2012-8 Lene Kjærsgaard: Empirical Essays of Active Labor Market Policy on Employment

2012-9 Henrik Nørholm: Structured Retail Products and Return Predictability

2012-10 Signe Frederiksen: Empirical Essays on Placements in Outside Home Care

2012-11 Mateusz P. Dziubinski: Essays on Financial Econometrics and Derivatives Pricing

2012-12 Jens Riis Andersen: Option Games under Incomplete Information

2012-13 Margit Malmmose: The Role of Management Accounting in New Public Management Reforms: Implications in a Socio-Political Health Care Context

2012-14 Laurent Callot: Large Panels and High-dimensional VAR

2012-15 Christian Rix-Nielsen: Strategic Investment

2013-1 Kenneth Lykke Sørensen: Essays on Wage Determination

2013-2 Tue Rauff Lind Christensen: Network Design Problems with Piecewise Linear Cost Functions

2013-3 Dominyka Sakalauskaite: A Challenge for Experts: Auditors, Forensic Specialists and the Detection of Fraud

2013-4 Rune Bysted: Essays on Innovative Work Behavior

2013-5 Mikkel Nørlem Hermansen: Longer Human Lifespan and the Retirement Decision

2013-6 Jannie H.G. Kristoffersen: Empirical Essays on Economics of Education

2013-7 Mark Strøm Kristoffersen: Essays on Economic Policies over the Business Cycle

2013-8 Philipp Meinen: Essays on Firms in International Trade

2013-9 Cédric Gorinas: Essays on Marginalization and Integration of Immigrants and Young Criminals – A Labour Economics Perspective

2013-10 Ina Charlotte Jäkel: Product Quality, Trade Policy, and Voter Preferences: Essays on International Trade

2013-11 Anna Gerstrøm: World Disruption - How Bankers Reconstruct the Financial Crisis: Essays on Interpretation

2013-12 Paola Andrea Barrientos Quiroga: Essays on Development Economics

2013-13 Peter Bodnar: Essays on Warehouse Operations

2013-14 Rune Vammen Lesner: Essays on Determinants of Inequality

2013-15 Peter Arendorf Bache: Firms and International Trade

2013-16 Anders Laugesen: On Complementarities, Heterogeneous Firms, and International Trade

2013-17	Anders Bruun Jonassen: Regression Discontinuity Analyses of the Disincentive Effects of Increasing Social Assistance
2014-1	David Sloth Pedersen: A Journey into the Dark Arts of Quantitative Finance
2014-2	Martin Schultz-Nielsen: Optimal Corporate Investments and Capital Structure
2014-3	Lukas Bach: Routing and Scheduling Problems - Optimization using Exact and Heuristic Methods
2014-4	Tanja Groth: Regulatory impacts in relation to a renewable fuel CHP technology: A financial and socioeconomic analysis
2014-5	Niels Strange Hansen: Forecasting Based on Unobserved Variables
2014-6	Ritwik Banerjee: Economics of Misbehavior
2014-7	Christina Annette Gravert: Giving and Taking – Essays in Experimental Economics
2014-8	Astrid Hanghøj: Papers in purchasing and supply management: A capability-based perspective
2014-9	Nima Nonejad: Essays in Applied Bayesian Particle and Markov Chain Monte Carlo Techniques in Time Series Econometrics
2014-10	Tine L. Mundbjerg Eriksen: Essays on Bullying: an Economist’s Perspective
2014-11	Sashka Dimova: Essays on Job Search Assistance
2014-12	Rasmus Tangsgaard Varneskov: Econometric Analysis of Volatility in Financial Additive Noise Models
2015-1	Anne Floor Brix: Estimation of Continuous Time Models Driven by Lévy Processes
2015-2	Kasper Vinther Olesen: Realizing Conditional Distributions and Coherence Across Financial Asset Classes
2015-3	Manuel Sebastian Lukas: Estimation and Model Specification for Econometric Forecasting
2015-4	Sofie Theilade Nyland Brodersen: Essays on Job Search Assistance and Labor Market Outcomes
2015-5	Jesper Nydam Wulff: Empirical Research in Foreign Market Entry Mode

2015-6 Sanni Nørgaard Breining: The Sibling Relationship Dynamics and Spillovers

2015-7 Marie Herly: Empirical Studies of Earnings Quality

2015-8 Stine Ludvig Bech: The Relationship between Caseworkers and Unemployed Workers

2015-9 Kaleb Girma Abreha: Empirical Essays on Heterogeneous Firms and International Trade

2015-10 Jeanne Andersen: Modelling and Optimisation of Renewable Energy Systems

2015-11 Rasmus Landersø: Essays in the Economics of Crime

2015-12 Juan Carlos Parra-Alvarez: Solution Methods and Inference in Continuous-Time Dynamic Equilibrium Economies (with Applications in Asset Pricing and Income Fluctuation Models)

2015-13 Sakshi Girdhar: The Internationalization of Big Accounting Firms and the Implications on their Practices and Structures: An Institutional Analysis

2015-14 Wenjing Wang: Corporate Innovation, R&D Personnel and External Knowledge Utilization

2015-15 Lene Gilje Justesen: Empirical Banking

2015-16 Jonas Maibom: Structural and Empirical Analysis of the Labour Market

2015-17 Sylvanus Kwaku Afesorgbor: Essays on International Economics and Development

2015-18 Orimar Sauri: Lévy Semistationary Models with Applications in Energy Markets

2015-19 Kristine Vasiljeva: Essays on Immigration in a Generous Welfare State

2015-20 Jonas Nygaard Eriksen: Business Cycles and Expected Returns

2015-21 Simon Juul Hviid: Dynamic Models of the Housing Market

2016-1 Silvia Migali: Essays on International Migration: Institutions, Skill Recognition, and the Welfare State

2016-2 Lorenzo Boldrini: Essays on Forecasting with Linear State-Space Systems

2016-3 Palle Sørensen: Financial Frictions, Price Rigidities, and the Business Cycle

2016-4 Camilla Pisani: Volatility and Correlation in Financial Markets: Theoretical Developments and Numerical Analysis

2016-5 Anders Kronborg: Methods and Applications to DSGE Models

2016-6 Morten Visby Krægpøth: Empirical Studies in Economics of Education

2016-7 Anne Odile Peschel: Essays on Implicit Information Processing at the Point of Sale: Evidence from Experiments and Scanner Data Analysis

2016-8 Girum Dagnachew Abate: Essays in Spatial Econometrics

2016-9 Kai Rehwald: Essays in Public Policy Evaluation

2016-10 Reza Pourmoayed: Optimization Methods in a Stochastic Production Environment

2016-11 Sune Lauth Gadegaard: Discrete Location Problems – Theory, Algorithms, and Extensions to Multiple Objectives

2016-12 Lisbeth Palmhøj Nielsen: Empirical Essays on Child Achievement, Maternal Employment, Parental Leave, and Geographic Mobility

2016-13 Louise Voldby Beuchert-Pedersen: School Resources and Student Achievement: Evidence From Social and Natural Experiments

2016-14 Mette Trier Damgaard: Essays in Applied Behavioral Economics

2016-15 Andrea Barletta: Consistent Modeling and Efficient Pricing of Volatility Derivatives

2016-16 Thorvardur Tjörvi Ólafsson: Macrofinancial Linkages and Crises in Small Open Economies

2016-17 Carlos Vladimir Rodríguez Caballero: On Factor Analysis with Long-Range Dependence

2016-18 J. Eduardo Vera-Valdés: Essays in Long Memory

2016-19 Magnus Sander: Returns, Dividends, and Optimal Portfolios

2016-20 Ioana Daniela Neamtu: Wind Power Effects and Price Elasticity of Demand for the Nordic Electricity Markets

2016-21 Anne Brink Nandrup: Determinants of Student Achievement and Education Choice

2016-22 Jakob Guldbæk Mikkelsen: Time-Varying Loadings in Factor Models: Theory and Applications

2016-23 Dan Nguyen: Formidability and Human Behavior: An Interdisciplinary Approach