The Path to Glory is Paved with Hierarchy: When Hierarchical Differentiation Increases Group Effectiveness

Richard Ronay
Columbia University
r.ronay@columbia.edu

Katharine Greenaway
University of Queensland
k.greenaway@psy.uq.edu.au

Eric Anicich
Columbia University
eanicich16@gsb.columbia.edu

Adam Galinsky
Northwestern University
agalinsky@kellogg.northwestern.edu

Paper Presented at the
25th Annual International Association of Conflict Management Conference
Spier, South Africa
July 12 – 14, 2012

Abstract: Two experiments examined the psychological and biological antecedents of hierarchical differentiation and the resulting consequences for productivity and conflict within small groups. In Experiment 1, which used a priming manipulation, hierarchically differentiated groups (i.e., groups comprising 1 high-power-primed, 1 low-power-primed, and 1 baseline individual) performed better on a procedurally interdependent task than did groups comprising exclusively either all high- power or all low-power individuals. There were no effects of hierarchical differentiation on performance on a procedurally independent task. Experiment 2 used a biological marker of dominance motivation (prenatal testosterone exposure as measured by a digit-length ratio) to manipulate hierarchical differentiation. The pattern of results from Experiment 1 was replicated; mixed-testosterone groups achieved greater productivity than did groups comprising all high-testosterone or all low-testosterone individuals. Furthermore, intragroup conflict mediated the productivity decrements for the high- testosterone but not the low-testosterone groups. This research suggests possible directions for future research and the need to further delineate the conditions and types of hierarchy under which hierarchical differentiation enhances rather than undermines group effectiveness.

Poultry scientists have made a surprising discovery: too many high egg-producing chickens ironically reduce overall egg production. Although breeding for greater egg production works for single-cage birds housed separately, when high egg producers are all placed together in a multiple-bird colony, cage-wide fertility plummets (Muir, 1996). It turns out the best egg producers are also the most competitive birds, and in a group setting they quickly devolve into fighting over food, space, and territory – these intragroup conflicts drive egg production down and bird mortality up. Chicken farmers take note; if you want to maximize group-level productivity you need harmony, and it seems that hierarchy provides the key.

The opposite appears to be the case, however, for humans. Research has found that inequality in groups can impair group functioning and performance. For example, more equal member contributions to group discussion predict collective intelligence at the group level (Woolley, Chabris, Pentland, Hashmi, & Malone, 2010). Similarly, wide disparities in pay increase organizational attrition (Wade, O'Reilly, & Pollock, 2006) and predict worse on-field performance in Major League Baseball (MLB; Bloom, 1999). Across corporations and baseball diamonds, hierarchical differentiation appears to hurt commitment and performance. Consistent with these data, many theories, including political ideologies (e.g. Marx & Engels, 1848), libertarian principles (Hancock et al., 1776), and utopian visions (Bellamy, 1888), have promoted the creation of egalitarian social structures.

Despite these compelling data and various attempts to model societies along egalitarian principles, hierarchy appears a universal default for social organization (Fiske, 2010). Hierarchy forms rapidly in human groups, requiring only minimal social interaction to emerge (Anderson & Kilduff, 2009; Van Vugt, 2006), and once formed perpetuates in a self-reinforcing manner (Magee & Galinsky, 2008). The ubiquity and tenacity of hierarchy as a social structure (Leavitt, 2005) lends support to its social-evolutionary value (King, Johnson, & Van Vugt, 2009; Van Vugt, 2006; 2008), and provides the basis for functional theories of hierarchy. These theories argue that when a group resolves itself into a clear hierarchy, it enhances the lot of all. Their central tenet is that the unequal distribution of power within groups facilitates the coordination of individuals' efforts and ultimately benefits the group as a whole (Halevy, Chou, & Galinsky, 2011; van Vugt, Hogan, & Kaiser, 2008). In the presence of a clear hierarchy, division of labor and patterns of deference reduce conflict, facilitate coordination, and ultimately improve productivity. In the absence of a clear hierarchy, competition and conflict undermine group cohesion and efficacy.

Consistent with the predictions of functional theories, recent work has demonstrated that status conflicts within groups, like those of the all high-producing multiple-bird colonies, can impair team performance. For instance, status disagreements within small work teams redirect energy and effort towards status contestation and away from group productivity (Bendersky & Hays, 2010). In research examining the group-level performance of Wall Street sell-side equity research analysts, the presence of too many high-achieving individuals within a single team had a negative effect on performance (Groysberg, Polzer, & Elfenbein, in press). These studies suggest that, for humans and chickens alike, too many high-status individuals create all-consuming status contests that disrupt the integration of activities essential for group productivity. It is important to point out that all the previous research cited above involved measured levels of hierarchical differentiation. We offer the first studies that have manipulated the overall levels of hierarchal differentiation of groups and then measured its effects on group productivity.

Recent perspectives have suggested that the benefits of hierarchy are most pronounced in situations of procedural interdependence (Halevy et al., 2011). The various group-level processes that contribute to the advantages of hierarchy – enhanced coordination, reduced conflict, and increased cooperation – are most relevant in contexts involving high, rather than low levels of coordination. For instance, although higher levels of pay dispersion *harm* performance when interdependence is low (e.g., professional baseball teams, Bloom, 1999), pay dispersion *benefits* performance in the case of professional basketball teams where procedural interdependence is high (Halevy, Chou, Galinsky, Murnighan, in press).

Building off functional theories of hierarchy, we propose that hierarchical differentiation within groups will improve performance especially when procedural interdependence is high. A high-functioning team needs both its leaders and its followers (Van Vugt et al., 2008), and too many of either is likely to present problems for productivity. Indeed, a formal test of functional theories of hierarchy requires evidence that productivity goes down both when there are too many *and* too few high-power individuals within a single group.

Although there is correlational evidence in the literature for the negative effects of too many high-power individuals (Groysberg at al., in press), the consequence of having too few high-power individuals remains an open question. We believe the reasons for lower productivity may be different in groups that have too few high-power individuals versus those that have too many. We predict that there will be greater conflict in those groups with all high-power individuals and this conflict will undermine group performance. In contrast, we do not expect intragroup conflict to drive the performance decrements of groups comprised of all low-power individuals. Although we have no specific hypotheses or tests in the current research, it may be that in groups with all low-power individuals there is too little hierarchy to produce coordinated behavior or too little agency to drive the group forward.

The current research manipulated hierarchical differentiation using two different dimensions of hierarchy – priming power and measuring testosterone. Individual differences in testosterone predict a desire for power (e.g. Schultheiss, Dargel, & Rohde, 2003) and status (e.g. Mazur & Booth, 1998) and high-testosterone individuals prefer being in high-power roles (Josephs, Sellers, Newman, & Mehta, 2006). In addition, experimental manipulations of power have been shown to increase circulating testosterone (Carney, Cuddy, & Yap, 2010). Thus, manipulations of power and measures of testosterone feed into and mutually reinforce each other with dominance serving as the basis for the formation of hierarchies in both.

In each experiment, we created three types of groups – all high-power/testosterone, all low-power/testosterone, or a mix of high-power/testosterone, low-power/testosterone, and baseline individuals – and had them work on a task characterized by a high level of procedural interdependence. Further, Experiment 1 included an additional task that required little integration or coordination of individual efforts to demonstrate that hierarchy does not facilitate performance on a procedurally independent task.

The current research makes a number of important contributions. It is the first to experimentally create different levels of hierarchical differentiation and then measure performance in small groups. We also use multiple bases of hierarchy – power and testosterone – to establish the robust advantage of hierarchical differentiation in procedurally interdependent groups. Further, we show no effect of hierarchy on a procedurally independent task. Finally, we establish that both groups of all high- and all low-power individuals perform worse than hierarchically differentiated groups but we show that the same process does not account for both types of groups. Like the chickens discussed at the outset, groups of all high-testosterone members have more conflict than groups of either all low or mixed levels of testosterone and this mediates the lower productivity of the all high-testosterone groups but not the lower productivity of the low-testosterone groups.

Experiment 1

In Experiment 1, we created groups of three individuals. We manipulated hierarchy by priming participants with high power, low power, or a baseline prime, and then placed them into groups of all high-power individuals, all low-power individuals, or a hierarchically differentiated group of one high-power, one low-power, and one baseline participant. Each group then engaged in two tasks. Task 1 was characterized by a high level of procedural interdependence: participants created sentences that required at least one word from each group member. Task 2 was characterized by a low level of procedural interdependence: participants generated novel uses for common household items. We predicted that the hierarchically differentiated groups of one high-power participant, one low-power participant, and one baseline participant would outperform both groups of all high-power and all low-power individuals when procedural interdependence was high. We predicted hierarchy would not enhance performance when procedural interdependence was low.

Method

Participants. 138 (37 male) undergraduate students were randomly assigned to one of three experimental conditions: a high-power condition, a low-power condition, and a baseline condition. Participants were then organized into same-sex triads that varied by condition: (1) three high-power participants, (2) three low-power participants, or (3) one high-power participant, one low-power participant, and one baseline participant. Triads worked together under face-to-face conditions for the remainder of the experiment.

Power manipulation. To manipulate power, we had participants recall and write about an incident in their lives (Galinsky, Gruenfeld, & Magee, 2003). This manipulation has been used in over twenty published papers (see Smith & Galinsky (2010) for a review). Participants in the *high-power condition* recalled a time in which they had power over another individual. Participants in the *low-power condition* recalled a time in which someone had power over them. Those in the baseline condition recalled their last trip to the supermarket (Rucker & Galinsky, 2008).

High Procedural Interdependence: Letter-Word Task. To measure group productivity under conditions of high procedural interdependence, we used a modified version of Crown's (2007) Letter-Word-Sentence game. While seated together at a table, each participant was presented with a unique matrix of 16 letters and instructed to find and record on a separate sheet as many words of three or more adjoining letters as possible. Letter-matrices were matched for number of potential words (M = 142.33, SD = 5.51). Groups then had the shared goal of combining their individual words to create sentences. Each sentence required at least one word from each group member. Thus, to succeed on this task group members were required to coordinate and integrate their individual efforts and unique information to create sentences. Groups were given five minutes to complete the task. Our measure of group productivity was the total number of sentences created by each group.

Low Procedural Interdependence: Creative Generation Task. To measure group productivity under conditions of low procedural interdependence, we used a creative generation task (e.g. Friedman & Forster, 2001; Markman, Lindberg, Kray, & Galinsky, 2007). Groups were asked to generate as many novel uses for three common items (i.e. newspaper, paperclip, brick) as they could. They were given two minutes per item to complete the task. Our measure of productivity was the sum of individuals' suggestions. Thus, unlike the letter-word-sentence game described above, success on this task was not contingent upon the successful coordination and integration of efforts. Indeed, a single group member could successfully complete the task alone if need be.

Results

High Procedural Interdependence: Letter-Word Task. As can be seen in Figure 1, there was a significant effect of group composition on productivity, $F(2,43) = 3.46$, $p = .04$, $\eta^2 = 0.14$. Consistent with predictions, the mixed-power groups ($M = 4.50$, $SD = 3.06$) were more efficient compared to the all high ($M = 2.53$, $SD = 1.06$) and all low-power groups ($M = 3.07$, $SD = 1.79$), $t(43) = 2.54$, $p = .02$, $d = 0.75$. Pairwise comparisons revealed that the mixed-power groups were more productive than the all high-power groups $t(43) = 2.53$, $p = .02$, $d = 0.92$, and marginally more productive than the all low-power groups, $t(43) = 1.85$, $p = .07$, $d = 0.59$. There was no difference in performance between the all high and all low-power groups $t(43) = .68$, $p = .50$, $d = 0.32$.

Low Procedural Interdependence: Creative Generation Task. As expected, no differences emerged between the high-power ($M = 16.58$, $SD = 6.42$), mixed-power ($M = 19.33$, $SD = 6.18$), or low-power ($M = 18.07$, $SD = 2.37$) groups on the creative generation task, $F(2,43) = 1.03$, $p=.37$, $\eta^2 = 0.05$. When procedural interdependence was low, there was no effect of hierarchy on productivity.

Experiment 1 provides the first experimental evidence that hierarchical differentiation facilitates greater productivity. Hierarchically differentiated groups – those that had a mix of high power, low power, and baseline participants – were more productive than groups of all high- or all low-power groups. Additionally, Experiment 1 provides support for the hypothesis that hierarchy is most beneficial in environments characterized by high levels of procedural interdependence (Halevy, Chou, & Galinsky, in press) as there was no effect on the procedurally independent creative generation task.

Experiment 2

The next experiment had two goals. First, we wanted to examine the biological foundations of hierarchical differentiation by measuring individual differences in prenatal testosterone exposure, which has organizing effects on the development of the brain and body (Manning, 2002). One marker of in utero testosterone exposure is the ratio between the length of the index finger (2D) and the ring finger (4D), with lower ratios indicating exposure to higher levels of androgens during prenatal development (Manning, 2002). We used differences in prenatal testosterone exposure as measured by 2D:4D ratio to create groups of all high-testosterone, all low-testosterone, or a mix of high, low, and average-testosterone individuals.

There is a large literature on the relationship between testosterone and dominance behaviors in humans (e.g. Mazur & Booth, 1998), non-human primates (e.g. Beehner, Bergman, Cheney, Seyfarth, & Whitten, 2006), and a range of other animals (e.g. Ruiz-de-la-Torre & Manteca, 1999). The overwhelming finding within this literature is that higher levels of testosterone motivate the pursuit and possession of power and dominance (e.g. Schultheiss, Dargel, & Rohde, 2003) and experimental manipulations of power increase testosterone levels (Carney et al., 2010). This hormone-fueled drive for power and dominance results in selective attention to cues of hierarchical threat (van Honk et al., 1999; 2001), and when placed in low-ranking roles, high-testosterone individuals experience elevated emotional and physiological arousal, increased concerns with their current standing, and diminished cognitive function (Josephs, Sellers, Newman, & Mehta, 2006). Importantly, our measure of the organizing effects of testosterone – 2D:4D ratio – has been shown to predict a number of dominance-seeking behaviors including retributional responding following provocation (Ronay & Galinsky, 2011), sporting ability and within-team performance (Manning, 2002), and perceived male dominance (Neave, Laing, Fink, & Manning, 2003). Thus, it seems likely that individual differences in testosterone play a role in the formation of naturally occurring hierarchies. We therefore predicted that limiting within-group variance in testosterone would disrupt the development of a clear hierarchy and thereby reduce group productivity.

Our second goal was to understand the processes that produce lower productivity when hierarchical differentiation is compressed. Because higher levels of testosterone motivate the pursuit of higher rank, we predicted that groups comprised exclusively of high-testosterone individuals would experience elevated competition and conflict as they jostled for dominance, which would impair productivity. On the other hand, although we anticipated that groups comprised entirely of low-testosterone individuals would experience similar productivity decrements due to compressed hierarchical differentiation, we did not expect conflict to emerge within these groups or to drive their lower productivity. Specifically, we expected that intragroup conflict would mediate performance decrements when there were too many high-testosterone individuals, but not when there were too many low-testosterone individuals. In those groups where prenatal testosterone exposure was broadly distributed, we expected to see the benefits of hierarchical differentiation – both higher productivity and less conflict.

Method

Participants-Participants were 109 (88 female) second-year psychology students.

Digit Ratio and Group Formation. Digit ratio was calculated by dividing the length of the 4th digit on the right hand by the length of the 2nd digit on the right hand, $M = 0.97$, $SD = 0.03$ (Manning, 2002). A criterion for "high" prenatal testosterone exposure was set at 1 SD below the mean for digit ratio. A criterion for low prenatal testosterone exposure was set at 1 SD above the mean for digit ratio. Although male ($M = 0.94$, $SD = 0.03$) and female ($M = 0.97$, $SD = 0.03$) digit ratios differed significantly, $F(1,106) = 10.88$, $p<.005$, we did not standardize within sex because the influence of digit ratio on behavior has been found to be consistent across sexes (e.g. Millet & Dewitte, 2009; Ronay & Galinsky, 2011). Groups of all high-testosterone, all low-testosterone, or a composite of high-testosterone, low-testosterone, and average-testosterone individuals were formed based on these criteria. Group sizes ranged from three to five participants ($M = 3.81$, $SD = 0.47$).

Group-Coordination Task. Group productivity was measured using the same modified Letter-Word-Sentence game as in Experiment 1, except there were two rounds and groups were given 10 minutes per round to complete the task. To control for variance in group size, we divided each group's total number of sentences in rounds one ($M = 23.41$, $SD = 9.48$) and two ($M = 27.97$, $SD = 13.80$) by the number of participants within each group. Because productivity was not affected by the interaction between round and group composition $F(1,23) = .16$, $p = .70$. ($M = 6.72$, $SD = 2.58$), we collapsed across the two rounds and used group-size-adjusted productivity outcomes as our primary dependent variable ($M = 6.72$, $SD = 2.58$).

Intragroup conflict. To determine the level of conflict present within each group, participants responded to seven items ($a = .93$; *e.g. There was conflict within our group*, (1) *Very True*, and (7) *Very Untrue).* A list of the scale items along with descriptive statistics and intercorrelations can be seen in Table 1. Although the items were intended to capture a breadth of conflict domains (i.e. process conflict, status conflict, relationship conflict, task conflict; see Bendersky & Hays, 2010) the obtained reliability value indicates a consistent, unifying construct, and so all analyses were conducted on the average of all item responses. We then created a group-level measure of conflict by averaging across individuals' responses within triads.

Results

Group Outcomes. As can be seen in Figure 2, there was a significant effect of group composition on productivity, $F(2,23) = 3.88$, $p = .04$, $\eta^2 = 0.25$. Mixed-testosterone groups ($M = 8.07$, $SD = 2.80$) were more productive than the all high ($M = 5.01$, $SD = 2.64$), and the all low ($M = 5.61$, $SD = 1.32$) testosterone groups $t(23) = 2.79$, $p = .01$, $d = 1.12$. Pairwise comparisons revealed that mixed testosterone groups were more productive than both high-testosterone, $t(23) = 2.32$, $p = .03$, $d = 1.13$, and low-testosterone groups, $t(23) = 2.16$, $p = .04$, $d = 1.19$. There was no difference between high and low-testosterone groups, $t(23) = 0.36$, $p = .72$, $d = 0.30$.

Intragroup conflict. As can be seen in Figure 3, a significant effect of group composition also emerged for conflict $F(2.23) = 3.43$, $p = .05$, $\eta^2 = 0.30$. Consistent with expectations, the all high-testosterone groups ($M = 2.29$, $SD = 0.40$) experienced more intragroup conflict than the mixed-testosterone groups ($M = 1.77$, $SD = 0.36$) and the all low-testosterone groups ($M = 1.84$, $SD = .43$), $t(23) = 2.49$, $p = .02$, $d = 1.27$. Pairwise comparisons revealed that the all high-testosterone groups experienced more intragroup conflict than the mixed-testosterone groups, $t(23) = 2.59$, $d = 1.37$, $p = .02$, and marginally more conflict than the all low-testosterone groups, $t(23) = 1.98$, $p = .06$, $d = 1.08$. There was no difference in reported conflict between the low-testosterone and mixed-testosterone groups $t(23) = .42$, $p = .68$, $d = 0.12$.

Mediation by Conflict. We next tested our hypothesis that conflict would mediate the effect of group composition on productivity for the all high-testosterone groups, but not for the all low-testosterone groups (i.e. moderated mediation; Preacher, Rucker, & Hayes, 2007). As can be seen in Figure 4, when compared with the mixed-testosterone groups, productivity decrements in high-testosterone groups were mediated by intragroup conflict. A bootstrapping procedure (Preacher & Hayes, 2004) with 10000 resamples confirmed that conflict significantly mediated the relationship between high-testosterone groups and reduced productivity (indirect effect = -1.08, SE = 0.79, 95% bias corrected confidence interval = -2.99, -0.01). However, intragroup conflict did not mediate group productivity for the all low- testosterone groups (indirect effect = 0.22, SE = 0.41, 95% bias corrected confidence interval = -0.47, 1.23; see Figure 5).

General Discussion

The current research experimentally tested, for the first time, the central prediction of functional theories of hierarchy: when power is distributed, intragroup conflicts go down and productivity goes up. We tested this prediction by manipulating hierarchy in two different ways – by priming power in Experiment 1 and by measuring a biological marker of individual differences in prenatal testosterone exposure in Experiment 2. In Experiment 1, hierarchically differentiated groups – those with a broader distribution of high, low, and baseline power – outperformed groups comprised of all high-power *and* groups comprised of all low-power individuals. Experiment 1 also demonstrated that the functional benefits of hierarchy are most pronounced under conditions of high procedural interdependence – when group productivity was simply the sum of participants' efforts, no between-group differences emerged. Consistent with the findings of Experiment 1 is research showing that higher levels of pay dispersion impair performance when interdependence is low (e.g., professional baseball teams, Bloom, 1999), but facilitate performance when procedural interdependence is high (e.g., professional basketball; Halevy, et al., in press).

In Experiment 2 we replicated this pattern of results using a biological marker of power-motivation (2D:4D) to manipulate the degree of hierarchical differentiation. Hierarchically differentiated groups with broadly distributed testosterone exposure outperformed groups comprised of all high-testosterone *and* groups comprised of all low-testosterone individuals.

Experiment 2 also demonstrated that the processes that produced lower productivity were different in the all high versus the all low-testosterone groups. Groups comprised exclusively of high-testosterone individuals experienced higher levels of intragroup conflict compared to both the mixed and the all low-testosterone groups. Furthermore, intragroup conflict mediated the performance decrements for the all high-testosterone groups, but not for the all low-testosterone groups. This is a critical insight and demonstrates that a different process accounts for the effects for the all high-testosterone groups versus the all low-testosterone groups. Future research should establish the precise reasons for poorer performance in the groups of all low-testosterone individuals. We suspect that a lack of agency may be at play. Just as too much conflict inhibited the ability of high-testosterone groups to coordinate their efforts, a lack of clear direction and agency may explain poorer performance in groups of all low-testosterone individuals.

It is worth noting that the measure of conflict employed here generalized across conflict domains. Future research might explore whether different types of all high-power groups generate different types of conflict (e.g., task, process, relationship, status) (Bendersky & Hays, 2010).

When Hierarchical Differentiation Enhances versus Impairs Group Success

The present research has practical implications for the composition and distribution of power and status within groups. Despite the widespread intuition that teams of high-performers will outperform their competition, the present data contribute to a growing body of literature (Groysberg et al., 2010; Halevy et al., 2011) that suggests this is not always the case. Our findings indicate that such teams are likely to experience elevated levels of conflict, reduced role differentiation, less coordination and integration, and poorer productivity than teams with a broader distribution of power and status.

Finally, as the present research focused on power and dominance-motives as the foundation for hierarchy, future research might seek to examine whether all forms of hierarchy are similarly functional. Recent theories have proposed that hierarchies can be conceptualized either as based in prestige or in dominance (Cheng, Tracy, & Henrich, 2010; Henrich & Gil-White, 2001). Prestige represents influence via respect and reverence and is based in the belief that one possesses socially desirable skills or expertise. In contrast, “dominance is typically seen in individuals who control access to resources…” (Cheng et al., 2010, p. 335), or who enter every situation “expecting to be in charge or to compete for control” (Fiske, 2010, p. 942). Future research should explore whether and in what ways prestige hierarchies carry the same functional form as dominance hierarchies.

Conclusion

Despite the overt appeal of egalitarian social structures, there remains an enduring implicit preference for hierarchy (Gruenfeld & Tiedens, 2010). The present data suggest that this preference may have its roots in the utilitarian value of distributed power. Pecking orders, it seems, are not just for the birds.

References

Anderson, C., & Kilduff, G.J. (2009). Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *Journal of Personality and Social Psychology*, 96, 491–503.

Beehner, J.C., Bergman, T.J., Cheney, D.L., Seyfarth, R.M., & Whitten, P.L. (2006). Testosterone predicts future dominance rank and mating activity among male chacma baboons. *Behavioral Ecology and Sociobiology, 59*, 469-479.

Bellamy, E. (1888) *Looking Backward, 2000-1887*. Ticknor and Company, Boston, MA.

Bendersky, C., & Hays, N. A. (2010). Status conflict in groups. Organization Science, ePub ahead of print August 4: http://orgsci.journal.informs.org/cgi/content/abstract/orsc.1100.0544v1

Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal, 42*, 25–40.

Cheng, J. T., Tracy, J. L., & Henrich, J. (2010). Pride, Personality, and the Evolutionary Foundations of Human Social Status. *Evolution and Human Behavior*. *31*, 334-347.

Crown, D.F. (2007). Effects of structurally competitive multilevel goals for an interdependent task. *Small Group Research, 38*, 265-288.

Fiske, S.T. (2010). Interpersonal stratification: Status, power, and subordination. In S.T. Fiske, D.T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 941–982). New York, NY: Wiley.

Friedman, R.S., & Forster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology*, *87*, 1001-1013.

Galinsky, A.D., Gruenfeld, D.H., & Magee, J.C. (2003). From power to action. *Journal of Personality & Social Psychology, 85*, 453–466.

Groysberg, B., Polzer, J., & Elfenbein, H. (forthcoming). Too many cooks spoil the broth: How high status individuals decrease group effectiveness. *Organization Science*.

Gruenfeld, D.H., & Tiedens, L.Z. (2010). Organizational preferences and their consequences. In S.T. Fiske, D.T. Gilbert, & G. Lindsay (Eds.), *Handbook of social psychology* (5th ed., pp.1252–1287). New York, NY: Wiley.

Halevy, N., Chou, E.Y., & Galinsky, A. (2011). A functional model of hierarchy: Why, how, and when vertical differentiation enhances group performance, *Organizational Psychological Review, 1*, 32-52.

Halevy, N., Chou, E., Galinsky, A.D., Murnighan, K. (In press). When Hierarchy Wins: Evidence from the National Basketball Association. Manuscript submitted for publication.

Hancock, et al. (1776). The United States Declaration of Independence. Retrieved June 24, 2011, from USHistory.org

Heinrich, J., Gil-White, F.J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*, 165-196.

Josephs, R. A., Sellers, J. G., Newman, M. L., & Mehta, P. H. (2006). The mismatch effect: When testosterone and status are at odds. *Journal of Personality and Social Psychology*, *90*, 999-1013.

King, A.J., Johnson, D.D.P., Van Vugt, M. (2009). The Origins and Evolution of Leadership. *Current Biology, 19*, R-911-R916.

Leavitt, H.J. (2005). Top down: Why hierarchies are here to stay and how to manage them more effectively. Boston, MA: Harvard Business School Press.

Magee, J.C., & Galinsky, A.D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals, 2*, 351–398.

Manning, J.T. (2002). *Digit ratio: A pointer to fertility, behavior, and health*. Rutgers University Press: New Brunswick, New Jersey.

Markman, K. D., Lindberg, M. J., Kray, L. J., & Galinsky, A. D. (2007). Implications of counterfactual structure for creative generation and analytical problem-solving. *Personality and Social Psychology Bulletin, 33*, 312-324.

Marx, K. & Engels, F. (1848, 1948) *The Commnunist Manifesto*. International Publishers.

Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Brain and Behavioral Sciences, 21*, 353-397.

Millet, K., & Dewitte, S. (2009). The presence of aggression cues inverts the relation between digit ratio (2D:4D) and pro-social behaviour in a dictator game. *British Journal of Psychology, 100*, 151−162.

Muir, W.M. (1996) Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science,* 75, 447-458.

Neave, N., Laing, S., Fink, B. & Manning, J.T. (2003). Second to fourth digit ratio, testosterone, and perceived male dominance. *Proceedings of the Royal Society of London B: Biological Sciences, 270*, 2167-2172.

Preacher, K.J., & Hayes, A.F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717–731.

Preacher, K.J., Rucker, D.D., & Hayes, A.F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*, 185-227.

Ronay, R., & Galinsky, A.D. (2011). Lex Talionis: Testosterone and the law of retaliation. *Journal of Experimental Social Psychology, 47,* 702–705.

Rucker, D.D., & Galinsky, A.D. (2008). Desire to acquire: Powerlessness and compensatory consumption. *Journal of Consumer Research, 35,* 257-267.

Ruiz-de-la-Torre, J., & Manteca, X. (1999). Effects of testosterone on aggressive behavior after social mixing in male lambs. *Physiology and Behavior, 68*, 109-113.

Schultheiss, O. C., Dargel, A., & Rohde, W. (2003). Implicit motives and gonadal steroid hormones: Effects of menstrual cycle phase, oral contraceptive use, and relationship status. *Hormones & Behavior, 43,* 293– 301.

Smith, P.K., & Galinsky, A.D. (2010). The nonconscious nature of power: Cues and consequences. *Social and Personality Psychology Compass, 4*, 918-938.

van Honk, J., Tuiten, A., Hermans, E., Putman, P., Koppeschaar, H., Thijssen, J., Verbaten, R., & van Doornen, L. (2001). A single administration of testosterone induces cardiac accelerative responses to angry faces in healthy young women. *Behavioral Neuroscience, 115,* 238 –242.

van Honk, J., Tuiten, A., Verbaten, R., van den Hout, M., Koppeschaar, H., Thijseen, J., & de Haan, E. (1999). Correlations among salivary testosterone, mood, and selective attention to threat in humans. *Hormones & Behavior, 36,* 17–24.

Van Vugt, M. (2006). Evolutionary origins of leadership and followership. *Personality and Social Psychology Review, 10*, 354–372.

Van Vugt, M., Hogan, R., & Kaiser, R.B. (2008). Leadership, followership, and evolution: Some lessons from the past. *American Psychologist, 63*, 182–196.

Wade, J.B., O'Reilly, C.A.I., & Pollock, T.G. (2006). Overpaid CEOs and underpaid managers: Fairness and executive compensation. *Organization Science, 17*, 527–544.

Woolley, Chabris, Pentland, Hashmi, & Malone, (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science, 330,* 686-688.

FIGURE 1. Mean number of sentences created by group composition.

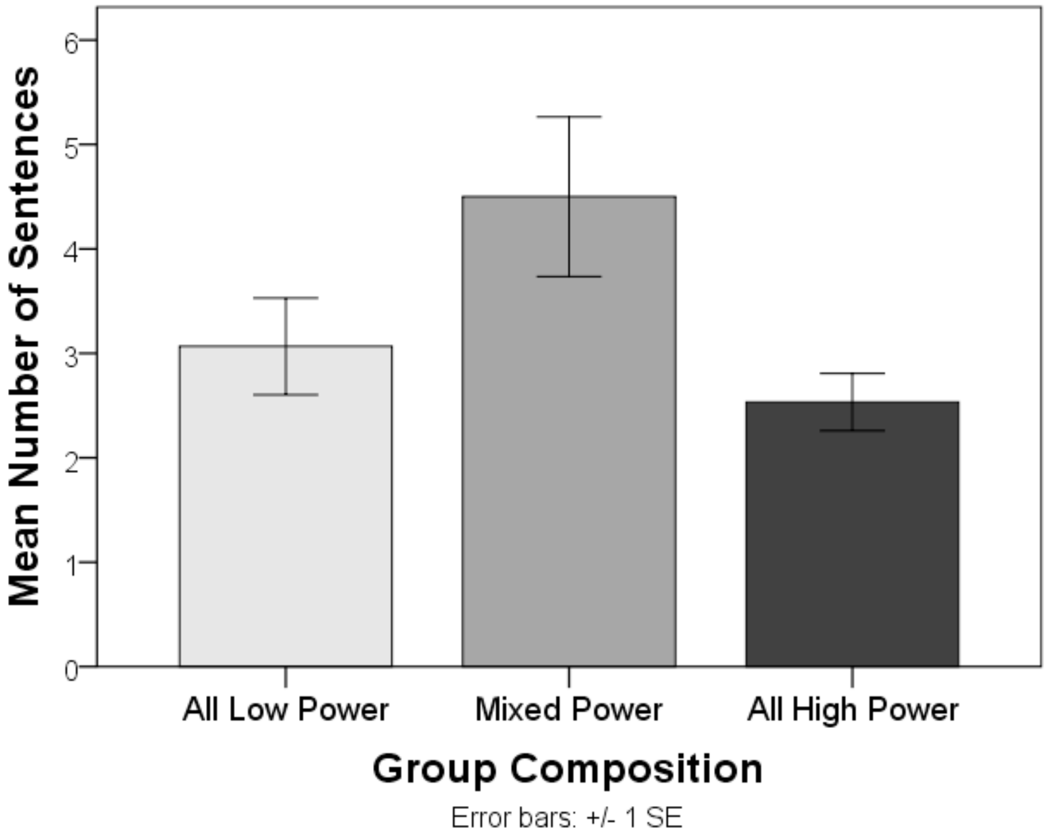


FIGURE 2. Mean number of sentences created by group composition.

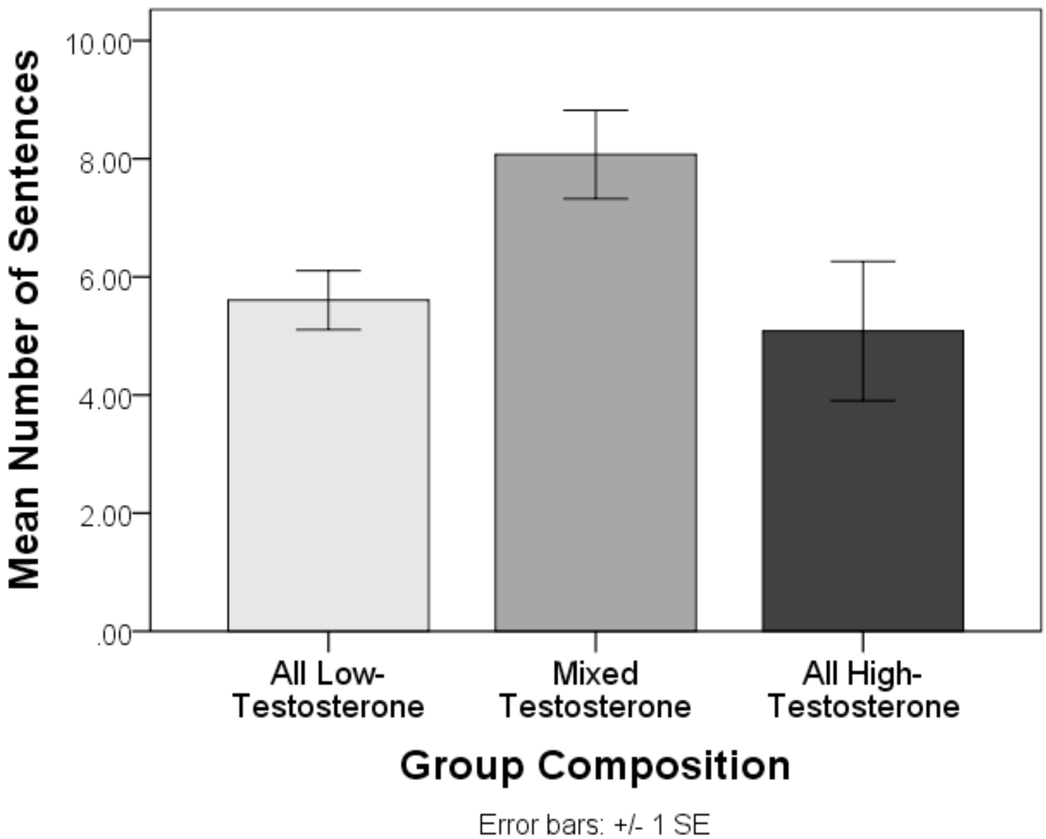


FIGURE 3. Mean level of conflict reported within groups by group composition.

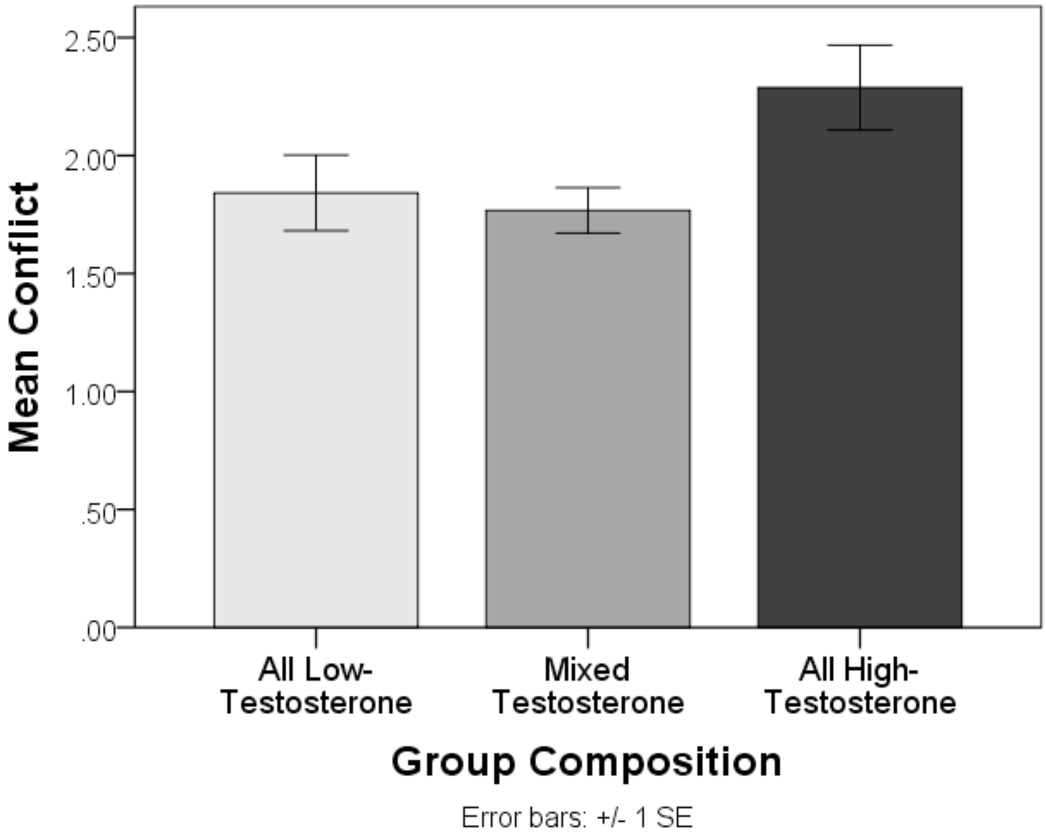
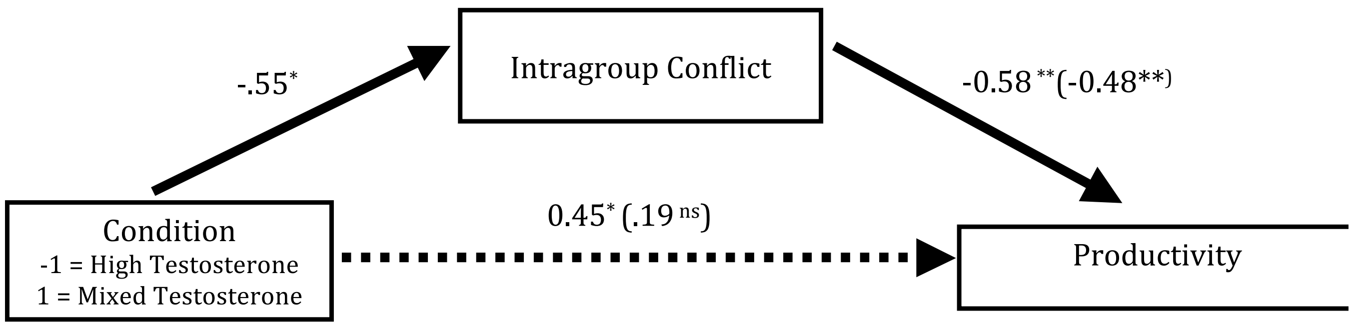
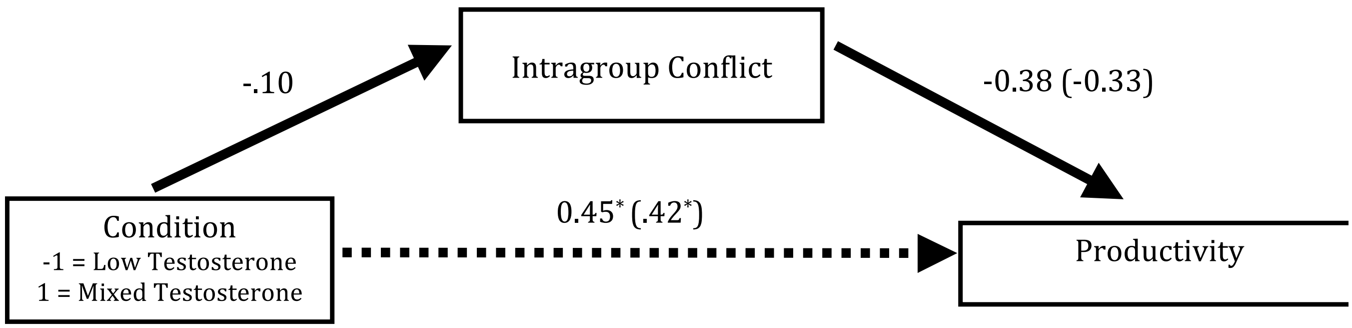


FIGURE 4. Mediation of the effect of group composition on productivity via conflict. This analysis compares the all high-testosterone groups against mixed testosterone groups. Numbers represent standardized regression coefficients; numbers in parentheses represent simultaneous regression coefficients



$^{*}p < .05$, $^{**}p < .01$

FIGURE 5. Mediation of the effect of group composition on productivity via conflict. This analysis compares the all low-testosterone groups against the mixed testosterone groups. Numbers represent standardized regression coefficients; numbers in parentheses represent simultaneous regression coefficients



$^{*}p < .05$, $^{**}p < .01$