ELECTRONIC SUPPLEMENTARY MATERIALS FOR 'PRESTIGE IN A LARGE-SCALE SOCIAL GROUP PREDICTS LONGITUDINAL CHANGES IN TESTOSTERONE'

TABLE OF CONTENTS

PARTICIPANTS AND DESCRIPTIVE INFORMATION

DESCRIPTIVE SUMMARIES

Below, in Table S1, we report the participant observables for each gender and the combined sample.

TABLE S1. DEMOGRAPHIC DESCRIPTIVE SUMMARY.

	Gender		Pooled
	Males	**Females**	
Age			
Mean	19.36	19.55	19.47
SD	1.3	1.75	1.53
Ethnicity			
Caucasian	45	65	110
Hispanic/Latino American	20	16	36
African American	5	3	8
Asian American	6	5	11
Native American	3	4	7
Other	1	1	2
Marching Band Experience (years, including the current)			
Mean	2.23	2.14	2.18
SD	1.22	1.17	1.19
Section Leader (*n*)	8	11	19
N	83	94	177
% of sample	46.89%	53.11%	100%

Tables S2 toS7 present the raw frequency counts of talent, advice, coercion, popularity, and friendship nominations, respectively, for men and women.

TABLE S2. DESCRIPTIVE SUMMARY OF TALENT NOMINATIONS.

	Number of Nominations as the Most Talented	Frequency	Percent	Cumulative Percent
Males				
	0	40	48.19	48.19
	1	18	21.69	69.88
	2	9	10.84	80.72
	3	2	2.41	83.13
	4	3	3.61	86.75
	5	1	1.2	87.95
	7	1	1.2	89.16
	9	2	2.41	91.57
	12	2	2.41	93.98
	16	1	1.2	95.18
	23	1	1.2	96.39
	49	1	1.2	97.59
	58	1	1.2	98.8
	65	1	1.2	100
	Total	83	100	-
Females				
	0	49	52.13	52.13
	1	17	18.09	70.21
	2	10	10.64	80.85
	3	5	5.32	86.17
	4	3	3.19	89.36
	5	5	5.32	94.68
	11	1	1.06	95.74
	20	1	1.06	96.81
	21	1	1.06	97.87
	22	1	1.06	98.94
	30	1	1.06	100
	Total	94	100	-

TABLE S3. DESCRIPTIVE SUMMARY OF ADVICE NOMINATIONS.

	Number of Nominations as an Advisor	Frequency	Percent	Cumulative Percent
Males				
	0	21	25.30	25.30
	1	18	21.69	46.99
	2	18	21.69	68.67
	3	8	9.64	78.31
	4	3	3.61	81.93
	5	2	2.41	84.34
	6	1	1.20	85.54
	7	2	2.41	87.95
	10	1	1.20	89.16
	11	1	1.20	90.36
	12	2	2.41	92.77
	15	2	2.41	95.18
	31	1	1.20	96.39
	41	1	1.20	97.59
	47	1	1.20	98.80
	48	1	1.20	100.00
	Total	83	100	-
Females				
	0	25	26.60	26.60
	1	22	23.40	50.00
	2	14	14.89	64.89
	3	6	6.38	71.28
	4	5	5.32	76.60
	5	4	4.26	80.85
	7	2	2.13	82.98
	8	3	3.19	86.17
	9	1	1.06	87.23
	10	5	5.32	92.55
	11	1	1.06	93.62
	12	2	2.13	95.74
	15	1	1.06	96.81
	18	1	1.06	97.87
	22	1	1.06	98.94
	23	1	1.06	100.00
	Total	94	100	-

120 TABLE S4. DESCRIPTIVE SUMMARY OF COERCION NOMINATIONS

	Number of Nominations as the Most Talented	Frequency	Percent	Cumulative Percent
Males				
	0	51	61.45	61.45
	1	15	18.07	79.52
	2	6	7.23	86.75
	3	2	2.41	89.16
	4	1	1.2	90.36
	5	3	3.61	93.98
	6	1	1.2	95.18
	12	1	1.2	96.39
	15	1	1.2	97.59
	20	1	1.2	98.8
	38	1	1.2	100
	Total	83	100	-
Females				
	0	59	62.77	62.77
	1	16	17.02	79.79
	2	6	6.38	86.17
	3	4	4.26	90.43
	4	1	1.06	91.49
	5	2	2.13	93.62
	6	1	1.06	94.68
	7	2	2.13	96.81
	8	1	1.06	97.87
	9	1	1.06	98.94
	27	1	1.06	100
	Total	94	100	-

121

122

123

124 TABLE S5. DESCRIPTIVE SUMMARY OF POPULARITY NOMINATIONS.

	Number of Nominations as the Most Socially Popular	Frequency	Percent	Cumulative Percent
Males				
	0	47	56.63	56.63
	1	12	14.46	71.08
	2	6	7.23	78.31
	3	3	3.61	81.93
	4	1	1.20	83.13
	5	1	1.20	84.34
	6	1	1.20	85.54
	7	1	1.20	86.75
	8	1	1.20	87.95
	9	1	1.20	89.16
	10	1	1.20	90.36
	11	1	1.20	91.57
	12	1	1.20	92.77
	14	1	1.20	93.98
	35	1	1.20	95.18
	39	1	1.20	96.39
	77	1	1.20	97.59
	94	1	1.20	98.80
	99	1	1.20	100.00
	Total	83	100	-
Females				
	0	40	42.55	42.55
	1	15	15.96	58.51
	2	18	19.15	77.66
	3	3	3.19	80.85
	4	2	2.13	82.98
	5	3	3.19	86.17
	6	2	2.13	88.30
	7	3	3.19	91.49
	8	1	1.06	92.55
	9	1	1.06	93.62
	17	1	1.06	94.68
	22	1	1.06	95.74
	23	2	2.13	97.87
	28	1	1.06	98.94
	30	1	1.06	100.00
	Total	94	100	-

125

TABLE S6. DESCRIPTIVE SUMMARY OF FRIENDSHIP NOMINATIONS (IN-COMING).

	Number of Close Friend Nominations Received (In-Coming)	Frequency	Percent	Cumulative Percent
Males				
	0	4	4.82	4.82
	1	4	4.82	9.64
	2	8	9.64	19.28
	3	7	8.43	27.71
	4	13	15.66	43.37
	5	6	7.23	50.6
	6	8	9.64	60.24
	7	6	7.23	67.47
	8	7	8.43	75.9
	9	6	7.23	83.13
	10	3	3.61	86.75
	11	3	3.61	90.36
	16	2	2.41	92.77
	18	1	1.2	93.98
	21	3	3.61	97.59
	24	1	1.2	98.8
	33	1	1.2	100
	Total	83	100	-
Females				
	0	1	1.06	1.06
	1	4	4.26	5.32
	2	7	7.45	12.77
	3	9	9.57	22.34
	4	8	8.51	30.85
	5	13	13.83	44.68
	6	7	7.45	52.13
	7	11	11.7	63.83
	8	8	8.51	72.34
	9	6	6.38	78.72
	10	4	4.26	82.98
	11	2	2.13	85.11
	12	3	3.19	88.3
	13	2	2.13	90.43
	14	4	4.26	94.68
	15	1	1.06	95.74
	17	1	1.06	96.81
	18	1	1.06	97.87
	19	1	1.06	98.94
	20	1	1.06	100
	Total	94	100	-

129 TABLE S7. DESCRIPTIVE SUMMARY OF FRIENDSHIP NOMINATIONS (OUT-GOING).

	Number of Close Friends Nominated (Out-Going)	Frequency	Percent	Cumulative Percent
Males				
	0	3	3.61	3.61
	1	4	4.82	8.43
	2	4	4.82	13.25
	3	9	10.84	24.1
	4	11	13.25	37.35
	5	9	10.84	48.19
	6	4	4.82	53.01
	7	7	8.43	61.45
	8	7	8.43	69.88
	9	8	9.64	79.52
	10	2	2.41	81.93
	11	2	2.41	84.34
	12	3	3.61	87.95
	13	1	1.2	89.16
	14	4	4.82	93.98
	16	2	2.41	96.39
	17	1	1.2	97.59
	19	2	2.41	100
	Total	83	100	-
Females				
	0	1	1.06	1.06
	1	7	7.45	8.51
	2	13	13.83	22.34
	3	9	9.57	31.91
	4	8	8.51	40.43
	5	9	9.57	50
	6	9	9.57	59.57
	7	6	6.38	65.96
	8	6	6.38	72.34
	9	2	2.13	74.47
	10	3	3.19	77.66
	11	2	2.13	79.79
	12	1	1.06	80.85
	13	4	4.26	85.11
	14	2	2.13	87.23
	15	5	5.32	92.55
	16	1	1.06	93.62
	17	3	3.19	96.81
	18	1	1.06	97.87
	19	2	2.13	100
	Total	94	100	-

130

TABLE S8. CORRELATION MATRIX FOR OUR KEY VARIABLES. CORRELATIONS ABOVE DIAGONAL ARE FOR MEN, AND BELOW DIAGONAL ARE FOR WOMEN. NOMINATION VARIABLES WERE TRANSFORMED USING THE NATURAL LOGARITHM FUNCTION TO REDUCE SKEW.

	Talent Nominations	Advice Nominations	Coercion Nominations	Popularity Nominations	Friendship Nominations (In-Coming)	Friendship Nominations (Out-Going)	Testosterone at Time 1	Testosterone at Time 2	Residual Change in Testosterone	Raw Change in Testosterone
Talent Nominations	1	0.5114*	0.3799*	0.3632*	0.3440*	0.2801*	0.1293	0.3669*	0.3232*	0.2776*
Advice Nominations	0.5040*	1	0.2617*	0.3808*	0.3676*	0.2819*	0.1580	0.3761*	0.2984*	0.2442*
Coercion Nominations	0.4553*	0.3347*	1	0.5400*	0.2607*	0.2280*	0.0539	0.1442	0.1236	0.1049
Popularity Nominations	0.5645*	0.3047*	0.3593*	1	0.1925	0.1083	0.0422	0.1131	0.0527	0.0280
Friendship Nominations (In-Coming)	0.3338*	0.4133*	0.2246*	0.3529*	1	0.0555	0.0602	0.1252	0.1243	0.1119
Friendship Nominations (Out-Going)	0.2233*	0.2792*	0.1751	0.2559*	0.8986*	1	0.1353	0.1679	0.0830	0.0474
Testosterone at Time 1	-0.1055	0.0513	-0.2354*	-0.0097	-0.0144	-0.0478	1	0.4437*	-0.2346*	-0.4314*
Testosterone at Time 2	-0.0565	0.1096	-0.1693	-0.0165	-0.0672	-0.0897	0.6374*	1	0.7671*	0.6170*
Residual Change in Testosterone	0.0062	-0.0398	0.0242	-0.0106	-0.0836	-0.0671	-0.2664*	0.5728*	1	0.9782*
Raw Change in Testosterone	0.0268	-0.0851	0.0884	-0.0069	-0.0733	-0.0464	-0.5339*	0.3112*	0.9572*	1

* $p < 0.05$

SUPPLEMENTAL ANALYSES

REGRESSION MODELS WITH CONTROLS

To explore the robustness of our findings, we examined alternative specifications based on residual change in T (as the dependent variable). In Tables S9 and S10, Model 1 shows the baseline model reported in the main text for talent nominations and advice nominations, respectively. Models 2 to 9 address the possibility that our findings may be driven by dominance, social popularity, friendship ties, age, ethnicity, prior marching band experience, or section leader status by including these variables as controls. As can be seen in Table S9, coefficients on gender × talent interaction remain negative and sizable across models, and range from -1.53 to -1.80 (though marginally significant in Models 8 and 9). Moreover, as shown in Table S10, coefficients on the gender × advice interaction remain negative and significant in all specifications, and range from -1.93 to -2.29. These checks reveal no qualitative divergences from our baseline findings, suggesting that the prestige

effect, as measured using either index, is not driven by between-subject differences in these controls.

TABLE S9. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) ON TALENT NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Talent Nomination	1.7383***	1.7085***	1.8263***	1.8091***	1.8360**	1.7990**	1.7601**	1.9097**	1.8941**
	(0.0002)	(0.0006)	(0.0004)	(0.0008)	(0.0011)	(0.0015)	(0.0022)	(0.0011)	(0.0014)
Gender (1 = Female)	-19.1177***	-19.2662**	-18.5660**	-16.0253*	-16.4383+	-14.2069+	-13.1237	-13.7552	-13.7805
	(0.0002)	(0.0019)	(0.0030)	(0.0441)	(0.0527)	(0.0965)	(0.1301)	(0.1143)	(0.1150)
Gender × Talent Nomination	-1.7223**	-1.7230*	-1.8032*	-1.7408*	-1.7564*	-1.7559*	-1.6352*	-1.5310+	-1.5396+
	(0.0068)	(0.0132)	(0.0180)	(0.0266)	(0.0293)	(0.0302)	(0.0499)	(0.0742)	(0.0739)
Coercion Nomination		0.0917	0.3802	0.3697	0.3899	0.4494	0.5916	0.5590	0.5421
		(0.8574)	(0.5157)	(0.5343)	(0.5214)	(0.4726)	(0.3529)	(0.3808)	(0.4003)
Gender × Coercion Nomination		-0.0205	-0.2973	-0.2699	-0.2944	-0.3089	-0.4005	-0.1836	-0.1779
		(0.9771)	(0.7012)	(0.7309)	(0.7120)	(0.7025)	(0.6251)	(0.8265)	(0.8324)
Popularity Nomination			-0.5654	-0.5638	-0.5730	-0.6467	-0.7136	-0.6130	-0.6267
			(0.3078)	(0.3120)	(0.3091)	(0.2588)	(0.2178)	(0.2949)	(0.2886)
Gender × Popularity Nomination			0.4936	0.5498	0.5646	0.7698	0.9049	0.6697	0.6726
			(0.5249)	(0.4855)	(0.4780)	(0.3400)	(0.2701)	(0.4291)	(0.4287)
Friendship Nomination (In-Coming)				0.1375	0.1142	0.1550	0.2316	0.3081	0.3102
				(0.9073)	(0.9239)	(0.8967)	(0.8497)	(0.8019)	(0.8013)
Gender × Friendship Nomination (In-Coming)				-1.1576	-1.7171	-1.3890	-1.4063	-1.3098	-1.2514
				(0.6038)	(0.7039)	(0.7582)	(0.7595)	(0.7767)	(0.7875)
Friendship Nomination (Out-Going)					-0.2237	-0.2729	-0.3384	-0.4561	-0.4600
					(0.8490)	(0.8160)	(0.7766)	(0.7025)	(0.7011)
Gender × Friendship Nomination (Out-Going)					0.8285	0.0426	-0.2660	0.3113	0.2306
					(0.8448)	(0.9920)	(0.9512)	(0.9435)	(0.9585)
Age						-1.5565	-1.6963	0.0517	0.0329
						(0.2672)	(0.2444)	(0.9763)	(0.9850)
Ethnicity: Other							-1.0534	4.0375	4.3495
							(0.9555)	(0.8319)	(0.8203)
Ethnicity: Native American							17.1804	18.3860	18.5463
							(0.1282)	(0.1040)	(0.1029)
Ethnicity: Asian American							2.9364	2.7687	2.7760
							(0.7396)	(0.7540)	(0.7542)
Ethnicity: African American							-6.5226	-2.0897	-1.6817
							(0.5285)	(0.8434)	(0.8760)
Ethnicity: Hispanic American							3.3668	3.8328	3.9267
							(0.5298)	(0.4774)	(0.4696)
Marching Band Experience								-4.5860+	-4.6618+
								(0.0565)	(0.0559)
Section Leader (1 = Leader)									1.7628
									(0.8247)
R^2	0.121	0.121	0.127	0.129	0.129	0.127	0.147	0.177	0.177
adj. R^2	0.104	0.093	0.088	0.078	0.066	0.057	0.046	0.069	0.063
AIC	1517.1007	1521.0459	1523.9305	1527.6052	1531.5420	1512.8421	1519.1854	1482.0279	1483.9715
BIC	1529.4757	1539.6084	1548.6805	1558.5427	1568.6670	1552.9003	1574.6507	1540.0966	1545.0964
N	163	163	163	163	163	161	161	157	157

p-values in parentheses.

All nomination variables used in these models have been transformed using the natural logarithm.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

TABLE S10. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) ON ADVICE NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Advice Nomination	1.8600***	1.7768**	1.9657***	1.9516**	1.9963**	2.0593**	2.1929**	2.2805***	2.2579***
	(0.0007)	(0.0016)	(0.0009)	(0.0017)	(0.0022)	(0.0016)	(0.0011)	(0.0008)	(0.0010)
Gender (1 = Female)	-14.6670***	-15.9843**	-14.9564*	-12.8446+	-13.3594	-11.8622	-11.7740	-11.7106	-11.7852
	(0.0006)	(0.0076)	(0.0145)	(0.1000)	(0.1069)	(0.1564)	(0.1628)	(0.1679)	(0.1666)
Gender × Advice Nomination	-1.9779**	-1.9323*	-2.1136**	-2.0073*	-2.0379*	-2.0985*	-2.2855*	-2.0555*	-2.0629*
	(0.0075)	(0.0122)	(0.0083)	(0.0182)	(0.0211)	(0.0173)	(0.0118)	(0.0265)	(0.0265)
Coercion Nomination		0.3232	0.6645	0.6560	0.6876	0.8496	1.0061	0.9688	0.9317
		(0.5150)	(0.2515)	(0.2692)	(0.2610)	(0.1759)	(0.1145)	(0.1298)	(0.1498)
Gender × Coercion Nomination		-0.2149	-0.5455	-0.5299	-0.5650	-0.6930	-0.7856	-0.5654	-0.5556
		(0.7540)	(0.4762)	(0.4961)	(0.4773)	(0.3885)	(0.3320)	(0.4947)	(0.5037)

Popularity Nomination			-0.6520	-0.6499	-0.6653	-0.8563	-0.9715	-0.8785	-0.9053
			(0.2488)	(0.2537)	(0.2486)	(0.1473)	(0.1032)	(0.1474)	(0.1387)
Gender × Popularity Nomination			0.6203	0.6725	0.6948	0.9969	1.2772	1.1514	1.1379
			(0.4038)	(0.3746)	(0.3652)	(0.2013)	(0.1084)	(0.1596)	(0.1661)
Friendship Nomination (In-Coming)				0.0897	0.0544	0.0012	-0.1203	0.0349	0.0354
				(0.9403)	(0.9643)	(0.9992)	(0.9230)	(0.9777)	(0.9774)
Gender × Friendship Nomination (In-Coming)				-1.0098	-1.4417	-1.0875	-0.4712	-0.8943	-0.7682
				(0.6653)	(0.7603)	(0.8170)	(0.9214)	(0.8522)	(0.8734)
Friendship Nomination (Out-Going)					-0.2879	-0.4142	-0.5209	-0.5604	-0.5747
					(0.8090)	(0.7270)	(0.6638)	(0.6406)	(0.6333)
Gender × Friendship Nomination (Out-Going)					0.7598	0.1182	-0.5474	0.1174	-0.0244
					(0.8600)	(0.9782)	(0.9006)	(0.9789)	(0.9956)
Age						-1.4976	-1.7487	-0.0702	-0.0990
						(0.2849)	(0.2274)	(0.9677)	(0.9547)
Ethnicity: Other							-8.8551	-4.6200	-3.9457
							(0.6360)	(0.8070)	(0.8358)
Ethnicity: Native American							20.6802^{+}	21.0468^{+}	21.4110^{+}
							(0.0702)	(0.0660)	(0.0629)
Ethnicity: Asian American							-1.5592	-2.6439	-2.5096
							(0.8618)	(0.7690)	(0.7812)
Ethnicity: African American							-7.2872	-3.6132	-2.7452
							(0.4798)	(0.7316)	(0.7985)
Ethnicity: Hispanic American							2.6042	2.6579	2.9054
							(0.6198)	(0.6156)	(0.5864)
Marching Band Experience								-4.6011^{+}	-4.7492^{+}
								(0.0604)	(0.0557)
Section Leader (1 = Leader)									3.3899
									(0.6654)
R^2	0.109	0.112	0.120	0.121	0.121	0.126	0.154	0.178	0.179
adj. R^2	0.093	0.084	0.080	0.069	0.057	0.056	0.053	0.071	0.066
AIC	1519.2009	1522.7049	1525.2974	1529.0660	1532.9887	1513.0276	1517.8734	1481.7431	1483.5279
BIC	1531.5759	1541.2674	1550.0474	1560.0035	1570.1137	1553.0858	1573.3387	1539.8118	1544.6528
N	163	163	163	163	163	161	161	157	157

p-values in parentheses.

All nomination variables used in these models have been transformed using the natural logarithm.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

SIMPLE EFFECTS FOR MEN AND WOMEN

To investigate the relationship between prestige and T change for each gender separately, we examined the simple slopes for each of the 9 specifications in Tables S9 and S10. These simple slopes are displayed below in Table S11 for talent and Table S12 for advice. As can be seen, across all specifications, talent and advice are both associated with a relatively greater positive change in T over time in men. By contrast, no significant associations emerged for women. These results indicate that our main findings are robust across the board.

TABLE S11. SIMPLE EFFECTS OF TALENT NOMINATIONS ON TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) IN MEN AND WOMEN

	b	*SE*	*t*	*p-value*	.95 CI	
Model 1						
Males	1.7383	0.4591	3.79	< .001	0.8315	2.6451
Females	0.0160	0.4286	0.04	0.970	-0.8304	0.8625
Model 2						
Males	1.7085	0.4907	3.48	0.001	0.7392	2.6778
Females	-0.0144	0.4811	-0.03	0.976	-0.9648	0.9359
Model 3						
Males	1.8263	0.5055	3.61	< .001	0.8278	2.8249
Females	0.0231	0.5599	0.04	0.967	-1.0830	1.1292
Model 4						
Males	1.8091	0.5294	3.42	0.001	0.7633	2.8549
Females	0.0683	0.5692	0.12	0.905	-1.0562	1.1928
Model 5						
Males	1.8360	0.5511	3.33	0.001	0.7472	2.9248
Females	0.0795	0.5778	0.14	0.891	-1.0621	1.2212
Model 6						
Males	1.7990	0.5555	3.24	0.001	0.7012	2.8969
Females	0.0431	0.5770	0.07	0.941	-1.0971	1.1834
Model 7						
Males	1.7601	0.5634	3.12	0.002	0.6464	2.8738
Females	0.1249	0.5961	0.21	0.834	-1.0534	1.3032
Model 8						
Males	1.9097	0.5734	3.33	0.001	0.7758	3.0435
Females	0.3787	0.6184	0.61	0.541	-0.8440	1.6014
Model 9						
Males	1.8941	0.5797	3.27	0.001	0.7478	3.0404
Females	0.3544	0.6300	0.56	0.575	-0.8914	1.6003

TABLE S12. SIMPLE EFFECTS OF ADVICE NOMINATIONS ON TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) IN MEN AND WOMEN

	b	*SE*	*t*	*p-value*	.95 CI	
Model 1						
Males	1.8600	0.5356	3.47	0.001	0.8022	2.9178
Females	-0.1179	0.4961	-0.24	0.812	-1.0978	0.8620
Model 2						
Males	1.7768	0.5531	3.21	0.002	0.6844	2.8692
Females	-0.1555	0.5249	-0.30	0.767	-1.1922	0.8811
Model 3						
Males	1.9657	0.5778	3.40	0.001	0.8244	3.1070
Females	-0.1478	0.5388	-0.27	0.784	-1.2122	0.9166
Model 4						

Males	1.9516	0.6106	3.20	0.002	0.7453	3.1580
Females	-0.0556	0.5778	-0.10	0.923	-1.1972	1.0859
Model 5						
Males	1.9963	0.6416	3.11	0.002	0.7287	3.2639
Females	-0.0416	0.5943	-0.07	0.944	-1.2159	1.1327
Model 6						
Males	2.0593	0.6414	3.21	0.002	0.7917	3.3269
Females	-0.0392	0.5903	-0.07	0.947	-1.2057	1.1273
Model 7						
Males	2.1929	0.6595	3.33	0.001	0.8893	3.4965
Females	-0.0925	0.6047	-0.15	0.879	-1.2878	1.1027
Model 8						
Males	2.2805	0.6671	3.42	0.001	0.9615	3.5994
Females	0.2249	0.6289	0.36	0.721	-1.0185	1.4684
Model 9						
Males	2.2579	0.6710	3.36	0.001	0.9310	3.5849
Females	0.1950	0.6345	0.31	0.759	-1.0596	1.4497

ROBUST REGRESSIONS

Next, we check the robustness of our findings to robust regressions. Ordinary least square estimations, which treat extreme and non-extreme scores equally, are vulnerable to outliers. In contrast, by assigning lower weight to outliers (i.e., observations with large residuals), robust methods are robust against the presence of outliers (Barnett & Lewis, 1994; Rousseeuw & Leroy, 2003). We reran our baseline model (Model 1, from Tables S9 and S10 above, again based on residual change in T) using a version of robust regression that uses a combination of Huber weights and biweights (Hamilton, 1992). Reconfirming our findings based on least squares regressions above, these results show that relative T change is significantly predicted by the interaction between gender and talent [b = -1.28, $t(159)$ = -2.10, p = .038, .95 CI(-2.49, -.07]), and between gender and advice [b = -1.82, $t(159)$ = -2.65, p = .009, .95 CI(-3.17, -.46)]. Simple effects shown below in Table S13, which remain highly similar in magnitude to those estimated by least square regressions, once

again indicate a significant positive predictive effect of advice on relatively greater T change in men but a null effect in women. These alternative estimates indicate that our primary results are robust to outliers.

TABLE S13. SIMPLE EFFECTS OF TALENT AND ADVICE NOMINATIONS ON TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) IN MEN AND WOMEN (ESTIMATED USING ROBUST REGRESSIONS)

	b	*SE*	*t*	*p-value*	.95 CI	
Talent Nominations						
Males	1.2806	0.4470	2.87	0.004	0.4046	2.1567
Females	-0.0011	0.4172	0.00	0.998	-0.8189	0.8166
Advice Nominations						
Males	1.7184	0.5039	3.41	0.001	0.7308	2.7059
Females	-0.0990	0.4668	-0.21	0.832	-1.0138	0.8159

SCATTERPLOTS OF TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) AS A FUNCTION OF PRESTIGE-RELATED NOMINATIONS

Figures S1 and S2 below present a scatterplot of intra-individual relative change in T as a function of the number of talent and advice nominations received, separately in men and women. The x-axis displays the raw number of nominations received. Zero nomination corresponds to the bottom 25th percentile for talent nominations, and the bottom 50th percentile for advice nomination. Five and ten nominations correspond to the 90th percentile for talent and advice nominations, respectively. Change in T was indexed using the unstandardized residuals of Time 2 T regressed on Time 1 T. In each, the line of best fit (in gray; 95% confidence intervals in yellow dashed line) shows the positive association between each measure of prestige and a higher than expected increase in T over time in men, and a null association in women. This trend is consistent with the nonparametric lowess curves (in light blue; Cleveland, 1979) shown. Together, these plots confirm our conclusion that prestige in men is associated with a rising T profile over time.

FIGURE S1. SCATTERPLOT OF CHANGE IN TESTOSTERONE (INDEXED USING RESIDUAL CHANGE SCORES) AS A FUNCTION OF THE NUMBER OF TALENT NOMINATIONS RECEIVED FOR MEN AND WOMEN. THE LINE OF BEST FIT (IN GRAY), 95% CONFIDENCE INTERVAL (IN YELLOW DASH), AND LOWESS CURVE (IN LIGHT BLUE) FOR EACH PANEL ARE SHOWN.

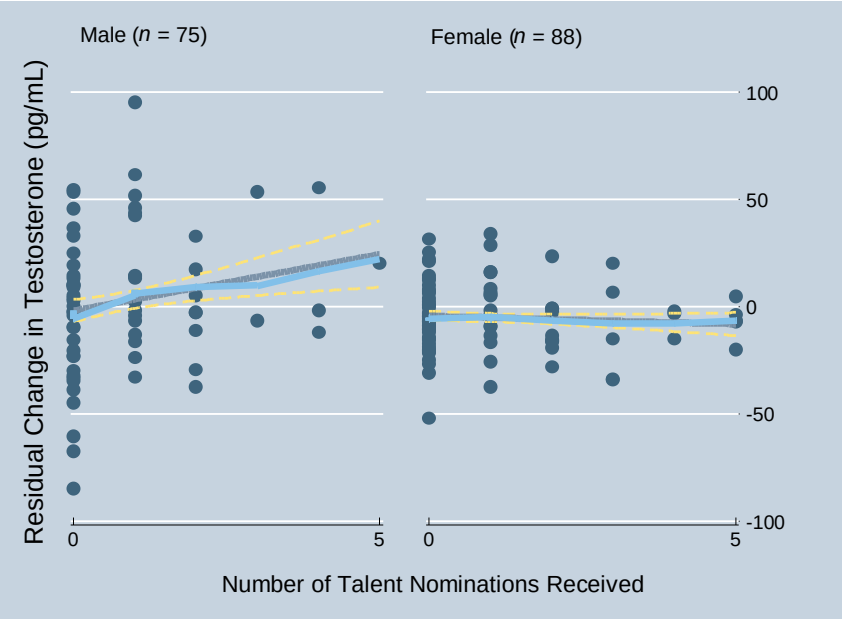
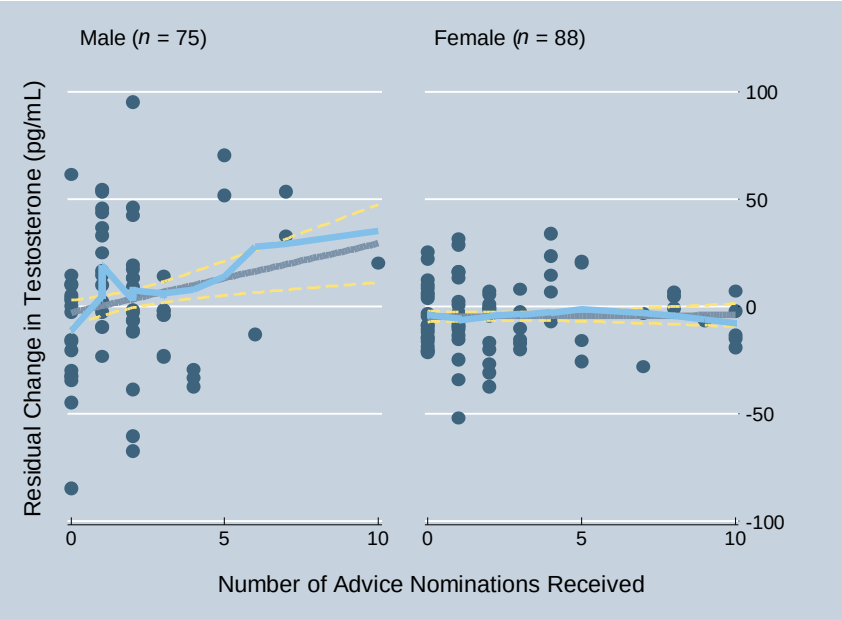


FIGURE S2. SCATTERPLOT OF CHANGE IN TESTOSTERONE (INDEXED USING RESIDUAL CHANGE SCORES) AS A FUNCTION OF THE NUMBER OF ADVICE NOMINATIONS RECEIVED FOR MEN AND WOMEN. THE LINE OF BEST FIT (IN GRAY), 95% CONFIDENCE INTERVAL (IN YELLOW DASH), AND LOWESS CURVE (IN LIGHT BLUE) FOR EACH PANEL ARE SHOWN.



EFFECTS OF PRESTIGE, AS APPROXIMATED BY BOTH TALENT AND ADVICE NOMINATIONS, ON CHANGE IN TESTOSTERONE (INDEXED USING RESIDUAL CHANGE SCORES)

The strong association (r = .51, p < .0001) between talent and advice nominations suggest that they both tap strongly into perceived prestige within the community. To capitalize on the availability of these two indices, we summed together the number of nominations an individual received across both domains to create a single prestige distribution index that captures an individual's overall level of respect in the community. Using this prestige composite, we again reran our baseline model (Model 1, from Tables S9 and S10 above) using residual change in T as the dependent variable. An interaction between gender and prestige again emerges [b = -2.14, t(159) = -2.87, p = .005, .95 CI(-3.62, -.67)]. The simple effects (shown below in Table S14) remain virtually identical to the coefficients in our original models based on either nomination variable. Prestige significantly predicts relative T change in men but not in women.

TABLE S14. SIMPLE EFFECTS OF AGGREGATED PRESTIGE-RELATED NOMINATIONS ON TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) IN MEN AND WOMEN

	b	***SE***	***t***	***p-value***	**.95 CI**	
Males	2.0733	0.5498	3.77	<.001	0.9874	3.1591
Females	-0.0714	0.5073	-0.14	0.888	-1.0734	0.9305

NO DETECTABLE EFFECT OF DOMINANCE ON CHANGE IN TESTOSTERONE (INDEXED USING RESIDUAL CHANGE SCORES)

Our results above suggest that dominance (indexed using coercion nominations) does not predict T changes, independent of prestige (see Tables S9 and S10 above, Models 2-9). While we think this question of whether acquiring dominance by force and force-threat triggers greater T production is important and interesting, we emphasize that

independent of the results obtained, there is reason to suspect that our ability to detect any influence of dominance on endocrine responses is impacted by the small sample of dominant individuals identified using our methodology.

Of the 83 male band members studied, 51 men (61.45%) did not receive a single nomination from the community as having a coercive and forceful disposition. 32 men (38.55%) were nominated by at least one peer, and only a very small group of 17 men (20.48% of all men) were nominated by at least two peers. By direct comparison, the talent and advice nomination data were much more evenly distributed. Only 21 men (25.30%) had zero advice nomination, and a more sizable group of 62 men (74.7 %) and 44 men (53.01%) received at least 1 and at least 2 advice nominations, respectively. Thus, our coercion nominations yielded a heavily skewed distribution at zero, in conjunction with a very small sample of individuals identified as possessing any degree of dominance (i.e., who were considered forceful by at least one other person in the group).

This pattern undermines the effectiveness of our dominance variable, by adding to the challenge of capturing the distribution of dominance ranks (i.e., variation in perceived formidability) within the group. Although we cannot definitive determine whether this restricted pool of dominant men in the group reflects an anomaly of our marching band community, or our peer nomination methodology deployed, or some combination of these factors, it is clear that this pattern can skew our results and work against the detection of any measurable effects of dominance on T. Thus, it is unsurprising we obtained a null effect for dominance. In sum, these issues suggest caution in interpreting the current results regarding the predictive effects of dominance on T as providing conclusive insights into the nature of endocrine responses.

Notwithstanding these issues, we sought to further examine the data for any evidence of dominance in influencing T concentrations, by testing the predictive power of coercion nominations on relative degree of T change, without controlling for prestige. Paralleling our efforts above for our primary analyses testing the effects of prestige, we ran a series of regression models using only dominance-based rank and the same set of demographic controls as above to predict residual change in T, without including talent, advice, or other nominations as simultaneous predictors. These results, shown in Table S15 below, confirm those obtained above when the effects of dominance was unconfounded from prestige. In the baseline model (with no controls; parallels Model 1 in Tables S9 and S10), the coefficient on the gender × coercion nomination interaction [b = -.63, t(159) = -.93, p = .356, .95 CI(-1.9598, .7092)] is non-significant. Simple slopes revealed no significant effect of coerciveness on relative T change in men [b = .69, t(159) = 1.40, p = .165, .95 CI(-.2868, 1.6664)] or women [b = .06, t(159) = .14, p = .889, .95 CI[-.8449, .9740]}. The other specifications yield the same qualitative conclusions.

These results suggest that dominance has no detectable effects on T production over time, at least in this social context. However, given the restricted range of measurable dominance in this social group (at least using the methods we deployed), we caution that this pattern may be in part merely an artifact of our dataset and study context. Thus, we cannot definitively conclude whether dominance has any general effects on T release in most other social groups in which greater dominance asymmetries exist. In the future, we plan to further examine the interplay between dominance and T by exploring these links more directly in other large-scale field social groups in which extensive within-group distributions of dominance have been empirically documented and shown to affect

297 resource and influence allocation. This includes, for example, small-scale forager societies
298 (von Rueden, Gurven, & Kaplan, 2008, 2011), athletic and sports teams (Cheng, Tracy, &
299 Henrich, 2010), children and adolescent social communities (Hawley, 1999; Redhead,
300 2016), and networks of competitive Master of Business Administration students
301 (McClanahan, Maner, & Cheng, 2017).

302 TABLE S15. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING RESIDUAL
303 CHANGE SCORES) ON COERCION NOMINATIONS.

	Model 1	**Model 2**	**Model 4**	**Model 5**	**Model 5**
Coercion Nomination	0.6898	0.6817	0.7526	0.8476	0.7519
	(0.1650)	(0.1793)	(0.1442)	(0.1060)	(0.1644)
Gender (1 = Female)	-14.9080*	-14.0283*	-14.1374*	-13.3353*	-13.4944*
	(0.0149)	(0.0235)	(0.0239)	(0.0361)	(0.0343)
Gender × Coercion Nomination	-0.6253	-0.5762	-0.5727	-0.3902	-0.3795
	(0.3562)	(0.3976)	(0.4043)	(0.5785)	(0.5897)
Age		-1.0690	-1.0420	0.5436	0.5299
		(0.4249)	(0.4499)	(0.7396)	(0.7463)
Ethnicity: Other			-5.1541	0.1921	1.0179
			(0.7844)	(0.9920)	(0.9577)
Ethnicity: Native American			15.4429	17.1801	17.6819
			(0.1698)	(0.1297)	(0.1202)
Ethnicity: Asian American			2.4872	2.2660	2.2096
			(0.7757)	(0.7963)	(0.8016)
Ethnicity: African American			-8.4202	-4.1884	-2.9384
			(0.4153)	(0.6932)	(0.7849)
Ethnicity: Hispanic American			3.3591	3.4630	3.8536
			(0.5238)	(0.5150)	(0.4718)
Marching Band Experience				-4.3540+	-4.7205*
				(0.0626)	(0.0489)
Section Leader (1 = Leader)					5.6207
					(0.4639)
R^2	0.053	0.052	0.072	0.094	0.097
adj. R^2	0.035	0.028	0.017	0.032	0.028
AIC	1529.1686	1510.1884	1516.7405	1481.1302	1482.5472
BIC	1541.5436	1525.5954	1547.5545	1514.7489	1519.2222
N	163	161	161	157	157

304 *p*-values in parentheses.
305 The dominance nomination variable used in these models has been transformed using the natural logarithm.
306 The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives
307 effects relative to Caucasian.
308 $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

SOCIAL POPULARITY AND FRIENDSHIPS DO NOT PREDICT CHANGE IN TESTOSTERONE

Our central hypothesis concerns how the experience of prestige calibrates the up-regulation of testosterone. Theoretically, the skill, expertise, and success of prestigious individuals (which are precisely the same attributes that generate their prestige in the first place) attract a coterie of loyal followers who willingly pay deference to their prestigious model and desire to hang around them, in exchange for access and increased opportunity to acquire the valuable skills, know-how, and information that these models possess (Henrich, Chudek, & Boyd, 2015; Henrich & Gil-White, 2001). Indeed, these patterns are well supported by ethnographic observations (e.g., Radcliffe-Brown, 1964; Sahlins, 1963) and existing empirical work by our team and others, which converge to indicate that highly prestigious individuals receive increased social popularity and support from other members of the community (e.g., Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013; Cheng et al., 2010; von Rueden et al., 2011). Might this greater social popularity or support enjoyed by prestigious individuals be responsible for our key prestige finding? To address this possibility, we conducted additional analyses to examine our effects after controlling for who in the community is socially popular and has large networks of friends.

We first began by examining the association between social popularity and T change, we ran a regression model similar to our primary analysis, in which we regressed residual change in T on the main and interactive effects of the natural logarithm of popularity nominations and gender. Results (shown in Table S16) indicate that, in contrast to talent, popularity was not associated with relative T change. There was no significant gender × popularity nomination interaction [b = -.30, t(159) = -.47, p = .639, .95 CI(-1.58, .97)]. As expected, popularity had a non-significant effect on relative T change in both men

[b = .28, t(159) = .59, p = .555, .95 CI(-.64, 1.20)] and women [b = -.03, t(159) = -.06, p = .952, .95 CI(-.91, .85)].

TABLE S16. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) ON POPULARITY NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4	Model 5
Popularity Nomination	0.2757	0.1639	0.1470	0.2670	0.1458
	(0.5547)	(0.7306)	(0.7601)	(0.5902)	(0.7746)
Gender (1 = Female)	-12.4807*	-10.9635*	-10.5393+	-11.1870*	-11.3735*
	(0.0189)	(0.0405)	(0.0517)	(0.0407)	(0.0376)
Gender × Popularity Nomination	-0.3028	-0.1338	-0.0048	0.0108	-0.0026
	(0.6391)	(0.8368)	(0.9942)	(0.9873)	(0.9969)
Age		-0.8666	-0.8213	0.5964	0.5970
		(0.5248)	(0.5574)	(0.7180)	(0.7176)
Ethnicity: Other			-7.8138	-3.9805	-2.0976
			(0.6799)	(0.8357)	(0.9133)
Ethnicity: Native American			13.6243	14.5119	15.5576
			(0.2241)	(0.1990)	(0.1703)
Ethnicity: Asian American			1.8855	1.2097	1.4944
			(0.8305)	(0.8917)	(0.8664)
Ethnicity: African American			-8.9268	-5.6891	-3.4801
			(0.3930)	(0.5946)	(0.7494)
Ethnicity: Hispanic American			2.9580	3.0148	3.6459
			(0.5768)	(0.5736)	(0.4989)
Marching Band Experience				-3.8660+	-4.4060+
				(0.0953)	(0.0639)
Section Leader (1 = Leader)					8.0806
					(0.2993)
R^2	0.044	0.041	0.059	0.077	0.084
adj. R^2	0.026	0.017	0.003	0.014	0.014
AIC	1530.8088	1511.9733	1518.9412	1483.9994	1484.8288
BIC	1543.1838	1527.3803	1549.7553	1517.6181	1521.5037
N	163	161	161	157	157

p-values in parentheses.
The popularity nomination variable used in these models has been transformed using the natural logarithm.
The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

Moreover, results using both in-coming (Table S17) and out-going friendship nominations (Table S18) as predictors confirm the null results obtained above for social popularity nominations. Neither form of friendship nominations were associated with relative T change in men. We also found no evidence of significant gender × in-coming

348 friendship nomination interaction or gender × out-going friendship nomination interaction.
349 Combined, results across all three of these indices of social popularity—nominations of
350 popularity and friendship (both in-coming and out-going)—indicate that, in contrast to
351 prestige, which predicts individual differences in T change profiles in men, social
352 popularity is unrelated to systematic changes in T.

353 TABLE S17. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING RESIDUAL
354 CHANGE SCORES) ON IN-COMING FRIENDSHIP NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4	Model 5
Friendship Nomination (In-Coming)	1.5689	1.5164	1.6378	1.8090	1.7079
	(0.1623)	(0.1789)	(0.1548)	(0.1204)	(0.1430)
Gender (1 = Female)	-7.5947	-6.7574	-6.7233	-8.3241	-8.2220
	(0.1419)	(0.1951)	(0.2064)	(0.1288)	(0.1331)
Gender × Friendship Nomination (In-Coming)	-2.4280	-2.5348	-2.6175	-2.0196	-2.2544
	(0.2484)	(0.2322)	(0.2236)	(0.3576)	(0.3061)
Age		-1.0731	-0.9521	0.5546	0.4472
		(0.4213)	(0.4863)	(0.7422)	(0.7909)
Ethnicity: Other			-7.4887	-4.2936	-1.9353
			(0.6894)	(0.8209)	(0.9191)
Ethnicity: Native American			13.4897	13.7515	15.3803
			(0.2222)	(0.2165)	(0.1700)
Ethnicity: Asian American			4.2910	3.9218	3.8765
			(0.6272)	(0.6598)	(0.6631)
Ethnicity: African American			-8.3979	-5.2777	-3.0496
			(0.4167)	(0.6185)	(0.7768)
Ethnicity: Hispanic American			3.3923	3.4605	4.0838
			(0.5197)	(0.5163)	(0.4455)
Marching Band Experience				-3.7198	-4.3817+
				(0.1090)	(0.0668)
Section Leader (1 = Leader)					8.6144
					(0.2447)
R^2	0.055	0.054	0.072	0.088	0.097
adj. R^2	0.037	0.029	0.017	0.026	0.028
AIC	1528.9257	1509.9130	1516.6801	1482.0226	1482.5519
BIC	1541.3007	1525.3201	1547.4941	1515.6413	1519.2269

N	163	161	161	157	157

p-values in parentheses.
The in-coming friendship nomination variable used in these models has been transformed using the natural logarithm.
The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

TABLE S18. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING RESIDUAL CHANGE SCORES) ON OUT-GOING FRIENDSHIP NOMINATIONS.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Friendship Nomination (Out-Going)	1.0519	0.9857	0.9951	1.0024	0.8643
	(0.3518)	(0.3811)	(0.3832)	(0.3835)	(0.4539)
Gender (1 = Female)	-8.5203^{+}	-7.6438	-7.8126	-9.3381^{+}	-9.2841^{+}
	(0.0917)	(0.1340)	(0.1309)	(0.0815)	(0.0828)
Gender × Friendship Nomination (Out-Going)	-1.7216	-1.8756	-1.8978	-1.2241	-1.4479
	(0.4052)	(0.3683)	(0.3679)	(0.5743)	(0.5075)
Age		-0.9120	-0.8129	0.6170	0.4736
		(0.4979)	(0.5576)	(0.7210)	(0.7842)
Ethnicity: Other			-8.1476	-5.2796	-2.7101
			(0.6653)	(0.7821)	(0.8876)
Ethnicity: Native American			13.5018	13.5361	15.3517
			(0.2249)	(0.2275)	(0.1748)
Ethnicity: Asian American			1.7148	1.3028	1.3727
			(0.8450)	(0.8831)	(0.8767)
Ethnicity: African American			-8.7606	-6.0327	-3.5762
			(0.4007)	(0.5725)	(0.7421)
Ethnicity: Hispanic American			2.6474	2.6051	3.3606
			(0.6178)	(0.6280)	(0.5342)
Marching Band Experience				-3.3809	-4.0838^{+}
				(0.1444)	(0.0876)
Section Leader (1 = Leader)					8.8977
					(0.2338)
R^2	0.048	0.047	0.065	0.078	0.087
adj. R^2	0.030	0.023	0.009	0.015	0.018
AIC	1530.1266	1511.0392	1518.0413	1483.8063	1484.2659
BIC	1542.5016	1526.4463	1548.8554	1517.4250	1520.9408
N	163	161	161	157	157

p-values in parentheses.
The out-going friendship nomination variable used in these models has been transformed using the natural logarithm.
The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

COMPARISONS AROSS DISCRETE LEVELS OF PRESTIGE IN MEN

To complement the analyses above and in the main text on residual change, in which levels of prestige (as indexed by talent and advice nominations) were treated as a continuous variable (after applying a natural logarithmic transformation), we ran additional analyses by treating prestige as discrete categories of raw nomination counts. We created three ordinal bins of subjects by grouping together the talent and advice nomination count data as follows: 0 nomination, 1-4 nominations, and 5 or more nominations. These groups were defined on the basis of our reasoning that 0 nomination corresponds to the absence of group-wide recognition, and, descriptively, receiving 5 or more prestige nominations puts an individual in the top 10 percentile of the group's prestige hierarchy. It's important to note, however, that although this "discretizing" approach may generate useful insights for understanding the link between prestige and relative T change, this procedure leads to the loss of information in our prestige measures and thus should be deemed as tentative and supplemental to our more suitable primary analyses based on continuous measures.

The descriptive means for relative change in T in these three categories of nomination frequency among men are displayed below in Figures S3 (for talent nominations) and S4 (for advice nominations). Note that these two figures parallel Figure 2 shown in the main text, which instead depicts mean raw changes in T (indexed using simple difference scores) across the same nomination count categories. We next conducted one-way ANOVAs to examine how relative T change varied as a function of level of prestige. This analysis indicates that mean relative change in T differed significantly across the three groups for both talent [$F(3, 72) = 4.05$, $p = .0101$, $R^2 = .1445$] and advice [$F(3, 72) = 5.80$, $p = .0013$, $R^2 = .19$]. Follow-up comparisons revealed the same pattern of results as in the

continuous analyses. For the talent measure, compared to men who received 0 nomination and on average experienced a decline in T, a significantly higher rise in T was found in men who received 1-4 nominations [$F(1, 72) = 4.39$, $p = .0396$] or 5+ nominations [$F(1, 72) = 8.20$, $p = .0055$]. Most critically, the level of relative T increase was significantly higher across the average of the two latter groups involving 1-4 and 5+ talent nominations than the group of men who did not receive any talent nomination [$F(1, 72) = 9.68$, $p = .0027$].

A similar pattern of results were obtained for our advice measure. Men who received 0 nomination on average experienced relatively less change in T. Relative to this group, a marginally significant greater change in T was found in men who received 1-4 nominations [$F(1, 72) = 3.29$, $p = .0739$], and significantly greater T change was found in men with 5 or more nominations [$F(1, 72) = 14.70$, $p = .0003$]. Finally, the degree of T increase was significantly higher across the average of the two latter groups involving 1-4 and 5+ advice nominations than the group of un-nominated men [$F(1, 72) = 11.09$, $p = .0014$].

Importantly, these results demonstrate that our key findings do not hinge on the use of ordinary least square analytic methods, nor do they reflect trends arising from the effects of one or two highly prestigious individuals in the community. As evidenced by the stepwise pattern visible in these bar graphs, the amount of T change relative to others tracks the experience of prestige, as we go from individuals at the bottom of the prestige hierarchy to those in the middle ranks, and from there to those who earned the most prestige. In sum, these results based on discretized levels of prestige reveal the same basic findings as our analyses based on continuous data. Both sets of results indicate that men

who are recognized as prestigious in their community show a relatively greater rise in T compared to men who lack prestige.

FIGURE S3. BAR GRAPH DEPICTING CHANGE IN TESTOSTERONE (INDEXED BY RESIDUAL CHANGE SCORES) AS A FUNCTION OF RAW TALENT NOMINATION COUNT ACROSS THREE LEVELS IN MEN.

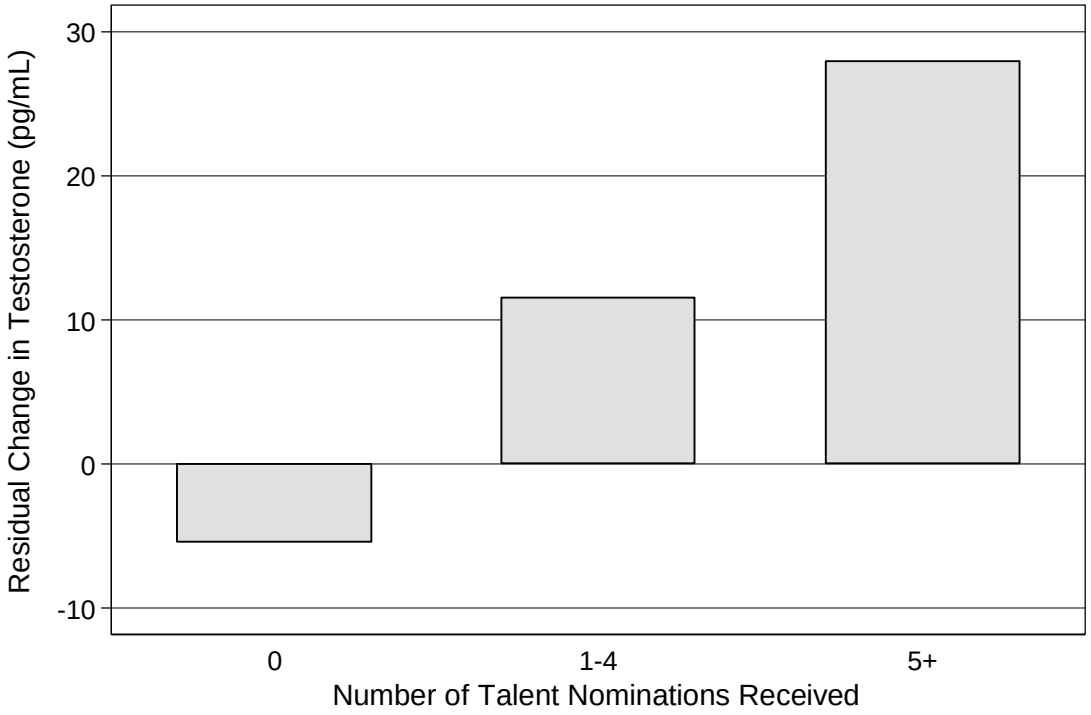
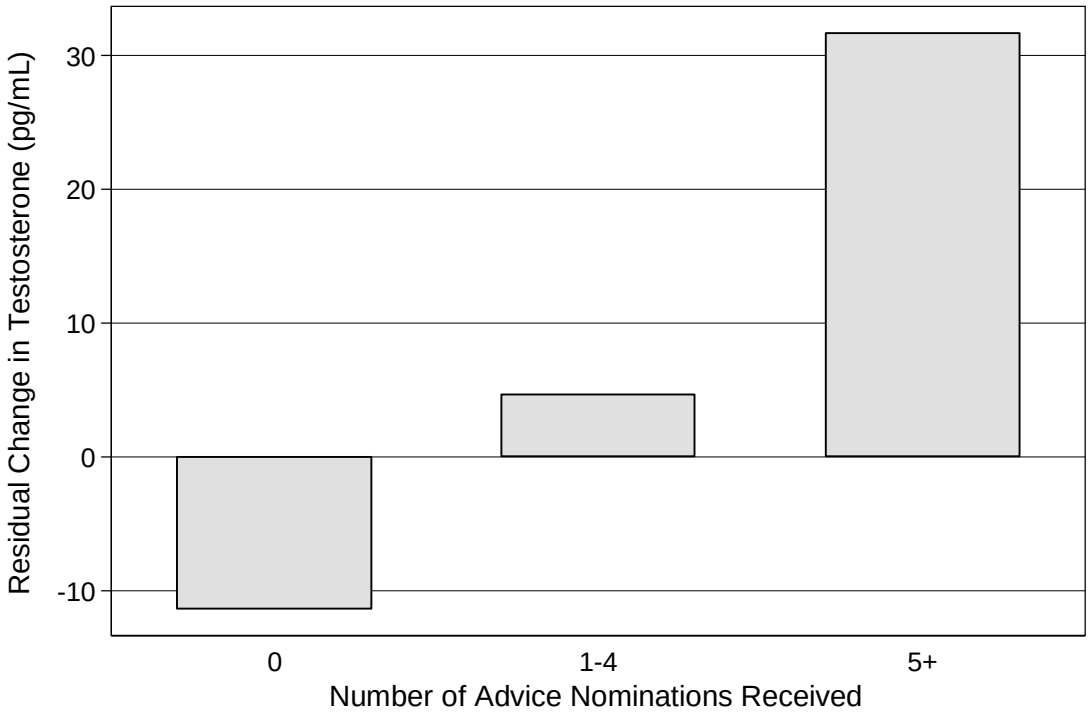


FIGURE S4. BAR GRAPH DEPICTING CHANGES IN TESTOSTERONE (INDEXED BY RESIDUAL CHANGE SCORES) AS A FUNCTION OF RAW ADVICE NOMINATION COUNT ACROSS THREE LEVELS IN MEN.



MEASURING CHANGE USING SIMPLE DIFFERENCE SCORES

The results presented above and in the main text focus primarily on the use of residual change scores to measure change, with a brief discussion of supplemental results using raw change scores (i.e., Time 2 T minus Time 1 T) as a robustness check (for a similar approach that combines both indices, see: Knight & Mehta, 2017; Mehta & Josephs, 2006). Before reporting and discussing a suite of additional statistical checks using the raw change measure below, however, it is beneficial to expand upon the presentation in the main text and remind readers how the interpretations differ for these two measures of change (also see Hand & Taylor, 1987).

As noted briefly in the main text, in the present research residual gain scores express Time 2 T as a deviation from the regression line predicting Time 2 T from Time 1 T. This means that the part of between-person variability in Time 2 T is partialled out, creating a "base-free measure of change" that is unconfounded by between-person differences in Time 1 T (Cronbach & Furby, 1970; Rogosa, Brandt, & Zimowski, 1982). Positive values on the residual change variable are thus interpreted as positive deviations from expectation (Time 2 T is higher than expected given Time 1 T), whereas negative values are negative deviations from expectation (Time 2 T is lower than expected given Time 1 T; Griffin, Murray, & Gonzalez, 1999). Because this measure captures a given individual's deviation (i.e., degree of change) from expectation *relative* to other individuals (rather than a person's *absolute* change, as indexed by raw change), it provides "a way of singling out individuals who changed more (or less) than expected" (Cronbach & Furby, 1970, p. 74). Consequently, residual scores are deemed most suitable for testing whether a trait of interest (such as status, as in our focus) is associated with change, which is precisely the kind of question in which we are interested.

Complicating interpretations, however, residual scores can be negative (reflecting a relatively weaker increase than expected) when the actual change in T from Time 1 to Time 2 is positive, and vice-versa.[1] Interpretation-wise, a positive association between residual change scores and prestige would indicate that highly prestigious individuals show a *relatively* larger increase in T, compared to their less prestigious counterparts. This association on its own, however, provides no clear direct indication of how the T of

[1] We thank the reviewer for the helpful suggestion to clarify the interpretation of the different measures of change employed.

relatively more and less prestigious individuals changed precisely. This question instead requires exploring the effect based on raw change scores.

By comparison, raw change scores (also called simple difference scores), which express change as the arithmetic difference between Time 2 T and Time 1 T, is more straightforward both in computation and interpretation. Intuitively, positive values on simple difference scores indicate an absolute rise in T from Time 1 to Time 2, whereas negative values indicate an absolute drop in T. Unlike the residual score approach, it indexes precisely how much an individual's T changed over time. However, despite its appeal in ease of interpretation, difference scores are less suitable than residual scores for exploring how change relates to other variables (i.e., what covaries with change, as in our interest here), given their lower reliability under many conditions, potential in confounding change with one or both variables that comprise the discrepancy index (i.e., Time 1 or Time 2 T), and vulnerability to ceiling effects (Cohen, Cohen, West, & Aiken, 2003; Cronbach & Furby, 1970; Griffin et al., 1999; John & Robins, 1994; Johns, 1981; Lord, 1956; McNemar, 1958; Schultheiss et al., 2005; Tucker, Damarin, & Messick, 1966). This raw change measure, on the other hand, is preferable when comparing mean change between two discrete groups (i.e., is there a difference in average change between the two populations). Given that both of these widely used approaches to capturing change have their own limitations (Burt & Obradović, 2013), although we think it is important to perform additional analyses using raw difference as robustness checks, the conclusions we draw are based on examining the convergence across both measures of change.

Our first set of analyses using raw change scores follow the same procedure deployed above for residual change scores, and is aimed at testing the predictive effect of each prestige index on raw change in T (as the dependent variable). We ran a series of 9 models (baseline and with controls). In all specifications, we used the natural logarithm of talent nomination or advice nomination, as in the main analyses. These regression results and their corresponding simple effects are summarized in Tables S19 to S22. These results suggest that overall our qualitative conclusions about the association between prestige and T change remain robust when change is assessed using difference scores rather than residual scores. For men, each logged talent nomination predicts a significant absolute increase of roughly 1.58 to 1.84 pg/mL in T, and each logged advice nomination predicts a significant increase of 1.56 to 2.00 pg/mL in T. For women, however, models predict a much weaker and non-significant change of roughly -.05 to +.49 pg/mL for talent, and -.42 to +.01 pg/mL for advice. These analyses indicate that for men, going from 0 (the 50th percentile in the community in terms of nominations received) to 5 talent nominations (the 90^{th} percentile) predicts an absolute increase in T of roughly 20.79 to 24.19 pg/mL, which is comparable to the estimated higher than expected increase of 22.81 pg/mL in the baseline model using residual change scores. Similarly, going from 0 (the 25^{th} percentile in terms of nominations received) to 10 advice nominations (the 90^{th} percentile) predicts a change in T of roughly 21.56 pg/mL to 27.65 pg/mL, also comparable to (though slightly weaker than) the estimated higher than expected increase of 25.70 pg/mL in the baseline model based on residual change scores.

TABLE S19. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING SIMPLE DIFFERENCE SCORES) ON TALENT NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Talent Nomination	1.6090**	1.5842**	1.7226**	1.6970**	1.7676**	1.7475**	1.6947**	1.8435**	1.8293**
	(0.0018)	(0.0040)	(0.0024)	(0.0042)	(0.0041)	(0.0050)	(0.0073)	(0.0041)	(0.0049)
Gender (1 = Female)	-2.7413	-1.8606	-1.0785	1.8715	0.7553	3.1575	4.0302	2.9270	2.9040
	(0.6247)	(0.7825)	(0.8742)	(0.8301)	(0.9352)	(0.7367)	(0.6727)	(0.7594)	(0.7621)
Gender × Talent Nomination	-1.5302*	-1.6294*	-1.6834*	-1.6073+	-1.6517+	-1.6666+	-1.5283+	-1.3499	-1.3578
	(0.0288)	(0.0332)	(0.0448)	(0.0627)	(0.0624)	(0.0618)	(0.0964)	(0.1523)	(0.1518)
Coercion Nomination		0.0763	0.4153	0.3998	0.4529	0.4653	0.5972	0.5929	0.5776
		(0.8921)	(0.5196)	(0.5423)	(0.4991)	(0.5005)	(0.3957)	(0.3990)	(0.4162)
Gender × Coercion Nomination		0.2133	-0.0993	-0.0649	-0.1279	-0.0970	-0.2112	-0.0434	-0.0383
		(0.7865)	(0.9075)	(0.9403)	(0.8843)	(0.9134)	(0.8154)	(0.9625)	(0.9671)
Popularity Nomination			-0.6643	-0.6619	-0.6860	-0.7366	-0.7886	-0.6341	-0.6465
			(0.2774)	(0.2819)	(0.2693)	(0.2440)	(0.2175)	(0.3254)	(0.3204)
Gender × Popularity Nomination			0.5027	0.5649	0.6020	0.7791	0.9151	0.6251	0.6278
			(0.5571)	(0.5158)	(0.4924)	(0.3816)	(0.3125)	(0.5029)	(0.5026)
Friendship Nomination (In-Coming)				0.2044	0.1430	0.1738	0.2540	0.3150	0.3169
				(0.8753)	(0.9135)	(0.8951)	(0.8507)	(0.8160)	(0.8155)
Gender × Friendship Nomination (In-Coming)				-1.3438	-2.6406	-2.3163	-2.4342	-2.4930	-2.4398
				(0.5849)	(0.5959)	(0.6420)	(0.6315)	(0.6244)	(0.6335)
Friendship Nomination (Out-Going)					-0.5882	-0.6363	-0.6680	-0.8397	-0.8432
					(0.6496)	(0.6232)	(0.6121)	(0.5237)	(0.5235)
Gender × Friendship Nomination (Out-Going)					1.9975	1.2419	1.0667	2.0434	1.9700
					(0.6685)	(0.7925)	(0.8242)	(0.6733)	(0.6865)
Age						-1.4978	-1.5290	0.7027	0.6857
						(0.3332)	(0.3417)	(0.7139)	(0.7218)
Ethnicity: Other							0.0492	6.1724	6.4564
							(0.9981)	(0.7685)	(0.7596)
Ethnicity: Native American							15.0843	16.6950	16.8408
							(0.2259)	(0.1797)	(0.1784)
Ethnicity: Asian American							1.9581	2.1662	2.1729
							(0.8409)	(0.8239)	(0.8240)
Ethnicity: African American							-11.4846	-6.0873	-5.7161
							(0.3156)	(0.6018)	(0.6305)
Ethnicity: Hispanic American							3.7885	4.7954	4.8808
							(0.5221)	(0.4201)	(0.4149)
Marching Band Experience								-5.4840*	-5.5530*
								(0.0388)	(0.0390)
Section Leader (1 = Leader)									1.6037
									(0.8549)
R^2	0.066	0.068	0.075	0.077	0.079	0.082	0.102	0.136	0.136
adj. R^2	0.048	0.038	0.033	0.023	0.012	0.008	-0.005	0.023	0.016
AIC	1549.4777	1553.1707	1555.8500	1559.5049	1563.1744	1544.6851	1551.1803	1512.5621	1514.5237
BIC	1561.8527	1571.7332	1580.6000	1590.4424	1600.2994	1584.7434	1606.6455	1570.6308	1575.6486
N	163	163	163	163	163	161	161	157	157

p-values in parentheses.

All nomination variables used in these models have been transformed using the natural logarithm.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

TABLE S20. SIMPLE EFFECTS OF TALENT NOMINATIONS ON TESTOSTERONE CHANGE (INDEXED USING SIMPLE DIFFERENCE SCORES) IN MEN AND WOMEN

	b	*SE*	*t*	*p-value*	.95 CI	
Model 1						
Males	1.6090	0.5071	3.17	0.002	0.6075	2.6104
Females	0.0788	0.4733	0.17	0.868	-0.8560	1.0136
Model 2						
Males	1.5842	0.5416	2.93	0.004	0.5145	2.6538
Females	-0.0453	0.5310	-0.09	0.932	-1.0940	1.0035
Model 3						
Males	1.7226	0.5575	3.09	0.002	0.6213	2.8239
Females	0.0392	0.6175	0.06	0.949	-1.1806	1.2591
Model 4						

Males	1.6970	0.5838	2.91	0.004	0.5437	2.8503
Females	0.0897	0.6277	0.14	0.887	-1.1504	1.3298
Model 5						
Males	1.7676	0.6072	2.91	0.004	0.5679	2.9674
Females	0.1160	0.6367	0.18	0.856	-1.1420	1.3739
Model 6						
Males	1.7475	0.6133	2.85	0.005	0.5355	2.9594
Females	0.0809	0.6370	0.13	0.899	-1.1779	1.3397
Model 7						
Males	1.6947	0.6223	2.72	0.007	0.4647	2.9247
Females	0.1664	0.6584	0.25	0.801	-1.1350	1.4678
Model 8						
Males	1.8435	0.6320	2.92	0.004	0.5939	3.0932
Females	0.4936	0.6815	0.72	0.470	-0.8540	1.8411
Model 9						
Males	1.8293	0.6389	2.86	0.005	0.5659	3.0928
Females	0.4715	0.6944	0.68	0.498	-0.9017	1.8447

TABLE S21. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING SIMPLE DIFFERENCE SCORES) ON ADVICE NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Advice Nomination	1.6403**	1.5608*	1.7703**	1.7363*	1.8295*	1.8667**	1.9728**	2.0014**	1.9754**
	(0.0062)	(0.0115)	(0.0061)	(0.0109)	(0.0106)	(0.0095)	(0.0078)	(0.0077)	(0.0089)
Gender (1 = Female)	0.8526	1.2871	2.3791	4.2565	3.1576	4.9823	5.0121	4.6463	4.5600
	(0.8549)	(0.8437)	(0.7219)	(0.6198)	(0.7285)	(0.5896)	(0.5910)	(0.6210)	(0.6286)
Gender × Advice Nomination	-1.9277*	-1.9829*	-2.1769*	-2.0738*	-2.1416*	-2.1764*	-2.3042*	-1.9865+	-1.9950+
	(0.0180)	(0.0196)	(0.0135)	(0.0268)	(0.0278)	(0.0255)	(0.0218)	(0.0529)	(0.0526)
Coercion Nomination		0.3089	0.6873	0.6668	0.7328	0.8409	0.9851	0.9719	0.9291
		(0.5725)	(0.2820)	(0.3084)	(0.2772)	(0.2258)	(0.1630)	(0.1706)	(0.1953)
Gender × Coercion Nomination		0.0787	-0.2784	-0.2526	-0.3249	-0.4005	-0.5098	-0.3398	-0.3284
		(0.9171)	(0.7414)	(0.7685)	(0.7107)	(0.6523)	(0.5698)	(0.7115)	(0.7216)
Popularity Nomination			-0.7229	-0.7178	-0.7498	-0.9021	-0.9990	-0.8420	-0.8729
			(0.2463)	(0.2532)	(0.2382)	(0.1677)	(0.1305)	(0.2105)	(0.1979)
Gender × Popularity Nomination			0.6596	0.6952	0.7398	0.9983	1.2710	1.1125	1.0968
			(0.4208)	(0.4053)	(0.3818)	(0.2474)	(0.1493)	(0.2205)	(0.2287)
Friendship Nomination (In-Coming)				0.2163	0.1425	0.1067	0.0092	0.1702	0.1708
				(0.8699)	(0.9152)	(0.9365)	(0.9947)	(0.9023)	(0.9022)
Gender × Friendship Nomination (In-Coming)				-0.9069	-1.6789	-1.3551	-0.9718	-1.5946	-1.4489
				(0.7247)	(0.7473)	(0.7944)	(0.8544)	(0.7647)	(0.7866)
Friendship Nomination (Out-Going)					-0.6008	-0.7014	-0.7711	-0.8318	-0.8483
					(0.6474)	(0.5933)	(0.5619)	(0.5326)	(0.5259)
Gender × Friendship Nomination (Out-Going)					1.4548	0.8209	0.3517	1.3589	1.1950
					(0.7593)	(0.8640)	(0.9423)	(0.7825)	(0.8092)
Age						-1.4326	-1.5789	0.4708	0.4376
						(0.3551)	(0.3253)	(0.8069)	(0.8209)
Ethnicity: Other							-6.6991	-1.5761	-0.7970
							(0.7467)	(0.9402)	(0.9699)
Ethnicity: Native American							18.8030	19.3165	19.7373
							(0.1369)	(0.1277)	(0.1217)
Ethnicity: Asian American							-1.7958	-2.4230	-2.2678
							(0.8565)	(0.8084)	(0.8211)
Ethnicity: African American							-12.0526	-7.7898	-6.7868
							(0.2925)	(0.5055)	(0.5697)
Ethnicity: Hispanic American							2.8080	3.3251	3.6111
							(0.6295)	(0.5715)	(0.5425)
Marching Band Experience								-5.1907+	-5.3619+
								(0.0563)	(0.0516)
Section Leader (1 = Leader)									3.9172
									(0.6525)
R^2	0.054	0.059	0.067	0.068	0.070	0.076	0.102	0.125	0.126
adj. R^2	0.036	0.029	0.025	0.013	0.002	0.001	-0.005	0.011	0.005
AIC	1551.5432	1554.6391	1557.2063	1561.0732	1564.8087	1545.7192	1551.1727	1514.5009	1516.2677
BIC	1563.9182	1573.2016	1581.9563	1592.0107	1601.9337	1585.7774	1606.6379	1572.5696	1577.3926
N	163	163	163	163	163	161	161	157	157

p-values in parentheses.

Both the advice and popularity nomination variables used in these models have been transformed using the natural logarithm.
The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

TABLE S22. SIMPLE EFFECTS OF ADVICE NOMINATIONS ON TESTOSTERONE CHANGE (INDEXED USING SIMPLE DIFFERENCE SCORES) IN MEN AND WOMEN

	b	*SE*	*t*	*p-value*	.95 CI	
Model 1						
Males	1.6403	0.5915	2.77	0.006	0.4722	2.8084
Females	-0.2874	0.5479	-0.52	0.601	-1.3695	0.7947
Model 2						
Males	1.5608	0.6100	2.56	0.011	0.3560	2.7656
Females	-0.4221	0.5789	-0.73	0.467	-1.5655	0.7213
Model 3						
Males	1.7703	0.6372	2.78	0.006	0.5116	3.0289
Females	-0.4067	0.5942	-0.68	0.495	-1.5805	0.7672
Model 4						
Males	1.7363	0.6736	2.58	0.011	0.4055	3.0671
Females	-0.3375	0.6374	-0.53	0.597	-1.5968	0.9218
Model 5						
Males	1.8295	0.7074	2.59	0.011	0.4319	3.2271
Females	-0.3121	0.6553	-0.48	0.635	-1.6068	0.9827
Model 6						
Males	1.8667	0.7100	2.63	0.009	0.4636	3.2697
Females	-0.3098	0.6534	-0.47	0.636	-1.6009	0.9814
Model 7						
Males	1.9728	0.7313	2.70	0.008	0.5272	3.4184
Females	-0.3314	0.6705	-0.49	0.622	-1.6568	0.9940
Model 8						
Males	2.0014	0.7404	2.70	0.008	0.5374	3.4654
Females	0.0149	0.6980	0.02	0.983	-1.3653	1.3951
Model 9						
Males	1.9754	0.7448	2.65	0.009	0.5026	3.4481
Females	-0.0196	0.7042	-0.03	0.978	-1.4122	1.3729

Next, we conducted further analyses to examine how T changes, as indexed by simple difference scores, varied across discretized categories of raw nomination data. These analyses, aimed at clarifying how absolute T change (increase or decrease) varies as a function of non-transformed nomination counts, mirror our analyses conducted above on

residual change scores. The means based on this measure of change are depicted in the main text in Figure 2. One-way ANOVAs reveal that mean change in T differed significantly across the three groups for both talent [$F(3, 72) = 5.48$, $p = .0019$, $R^2 = .186$; shown in Panel A] and advice [$F(3, 72) = 6.88$, $p = .0004$, $R^2 = .223$; shown in Panel B]. Men who received zero nomination experienced an absolute decline in T over time (M = -23.20 pg/mL for talent; M = -28.22 pg/mL for advice), whereas men who received 5 or more nominations experienced an increase in T (M = 7.26 pg/mL for talent; M = 11.87 pg/mL for advice). Follow-up mean comparisons confirmed that these means differ significantly from each other [$F(1, 72) = 5.67$, $p = .0199$ for talent; $F(1, 72) = 10.42$, $p = .0019$ for advice]. Taken together, our qualitative conclusions converge across the residual score and simple difference score approach to assessing change.

MEASURING PRESTIGE USING RANKINGS

In the results reported above and in the main text, position in the community's prestige-based status hierarchy as a whole was measured using the total number of nominations received. An alternative approach, which we present here, is to create a direct measure of relative rank in the community. The application of this ranking method, commonly deployed in the study of dominance relationships in non-human primate social groups living in the wild (e.g., baboons, chimpanzees; Alberts & Altmann, 1995; Archie, Altmann, & Alberts, 2012; Sapolsky, 1983), more closely maps onto the notion of a linear status hierarchy—that is, an asymmetric, connected, and transitive "pecking order". To pursue this, we constructed a rank version of our nomination data, by sorting the number of nominations received for being the most talented and then assigning ranks, placing the

552 top ranking person first. In the case of ties—that is, two or more individuals receiving the
553 same number of nominations—an average rank is assigned. For ease of interpretation, we
554 then reversed this variable by multiplying by -1 so that higher scores (i.e., negative scores
555 closer to zero) reflect highest prestige according to the community's perception. The same
556 procedure was repeated for all the other nomination domains (dominance, social
557 popularity, and in-coming and out-going friendship) to create a ranking-based measure of
558 each.

559 This ranking approach offers several advantages. First, using actual talent and
560 advice rankings as predictors allows a direct exploration of the quantitative effect of each
561 additional rank on testosterone changes (as opposed to the effect of each additional
562 nomination), which renders a more readily interpretable metric. Second, it also addresses
563 the prickly analytical challenge discussed above stemming from potential outliers—top-
564 ranking individuals who receive an exceptionally high number of nominations.[2] The
565 conversion of nomination count data into relative ranks provides an analytical workaround
566 by making these extreme scores less extreme. Note, however, that a (potentially less than
567 ideal) assumption introduced by rankings is that the distance between each subsequent
568 rank is presumed equal, regardless of the actual difference in nominations received. For
569 example, a difference of 1 nomination, or of 30 nominations (as in the case of extreme
570 scores), can both be assigned a 1-rank difference. By contrast, rather than treat the distance
571 between ranks as not meaningful, our nomination tally approach above emphasizes the

[2] Note, however, that far from being identified in error by a few peers, individuals who receive a large number of nominations reliably reflect the community's consensus regarding who is actively respected and admired. That is, though in a statistical sense the individual(s) with unusually high total nominations are statistical outliers, they may be considered meaningful, legitimate data points, and their inclusion is crucial for our examination of how prestige predicts endocrine responses.

distance between individuals. As such, given that all approaches to assessing status relationships among members within a social group have important—but different—limitations (Bernstein, 1981), we believe that our conclusions are best drawn in light of evaluating evidence emerging from different methods.

To examine convergence with our primary findings, we re-estimate the baseline model (i.e., Model 1 in Tables S9, S10, S19 and S21) using prestige-based rankings (instead of tallied nominations). In each case, we regress change in T on the prestige index, gender, and the interaction between prestige and gender. We present analyses for four models to explore robustness across our two ways of measuring change (residual change or raw change as the dependent variable) and two indices of prestige (talent and advice as the predictor of interest): (1) talent rank predicting residual change, (2) talent rank predicting raw change, (3) advice rank predicting residual change, and (4) advice rank predicting raw change. Results, displayed in Table S23, show that across all four specifications, the coefficient on the prestige index (which captures the simple slope in men) is positive and significant across the board, ranging from 0.14 to 0.22. Note that the coefficients denote the effect of each additional rise in rank in the community on T. As such, A 10-rank increase in prestige is associated with a higher than expected T increase of roughly 1.74 to 2.25 pg/mL (residual change), and an absolute increase of 1.39 to 2.07 pg/mL (raw change). The scatterplots for talent rank and advice rank are displayed below in Figures S5 and S6. Taken together, these analyses indicate that our primary results based on nomination count are robust to this alternate approach of mapping prestige relationships using absolute relative ranks.

TABLE S23. OLS REGRESSIONS OF TESTOSTERONE CHANGE (INDEXED USING EITHER RESIDUAL OR SIMPLE DIFFERENCE SCORES) ON PRESTIGE RANKS (INDEXED USING EITHER TALENT OR ADVICE NOMINATION RANKS).

	DV = Residual Change Prestige Index = Talent Nomination Rank (Model 1)	DV = Raw Change Prestige Index = Talent Nomination Rank (Model 2)	DV = Residual Change Prestige Index = Advice Nomination Rank (Model 3)	DV = Raw Change Prestige Index = Advice Nomination Rank (Model 4)
Prestige Index	0.2245***	0.2073**	0.1742**	0.1392*
	(0.0003)	(0.0023)	(0.0045)	(0.0385)
Gender (1 = Female)	-30.8808***	-12.8681	-26.7887**	-8.1400
	(0.0002)	(0.1584)	(0.0011)	(0.3628)
Gender × Prestige Index	-0.2336**	-0.2037*	-0.1835*	-0.1464
	(0.0054)	(0.0273)	(0.0245)	(0.1028)
R^2	0.118	0.063	0.089	0.033
adj. R^2	0.101	0.046	0.072	0.015
AIC	1517.6483	1549.9527	1522.8257	1555.0906
BIC	1530.0233	1562.3277	1535.2007	1567.4656
N	163	163	163	163

p-values in parentheses.

The prestige index is either talent nomination relative rank or advice nomination relative rank (indicated in the header). That is, in Models 1 and 2, the prestige index is talent nomination rank, and in Models 3 and 4, the prestige index is advice nomination rank. Both of these indices were computed from raw nomination counts and reversed by multiplying by -1 so that higher values indicate higher prestige.

$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < .01$, $^{***}p < .001$

FIGURE S5. SCATTERPLOT OF CHANGE IN TESTOSTERONE (INDEXED USING RESIDUAL CHANGE SCORES) AS A FUNCTION OF TALENT RANK FOR MEN AND WOMEN. THE LINE OF BEST FIT (IN GRAY), 95% CONFIDENCE INTERVAL (IN YELLOW DASH), AND LOWESS CURVE (IN LIGHT BLUE) FOR EACH PANEL ARE SHOWN.

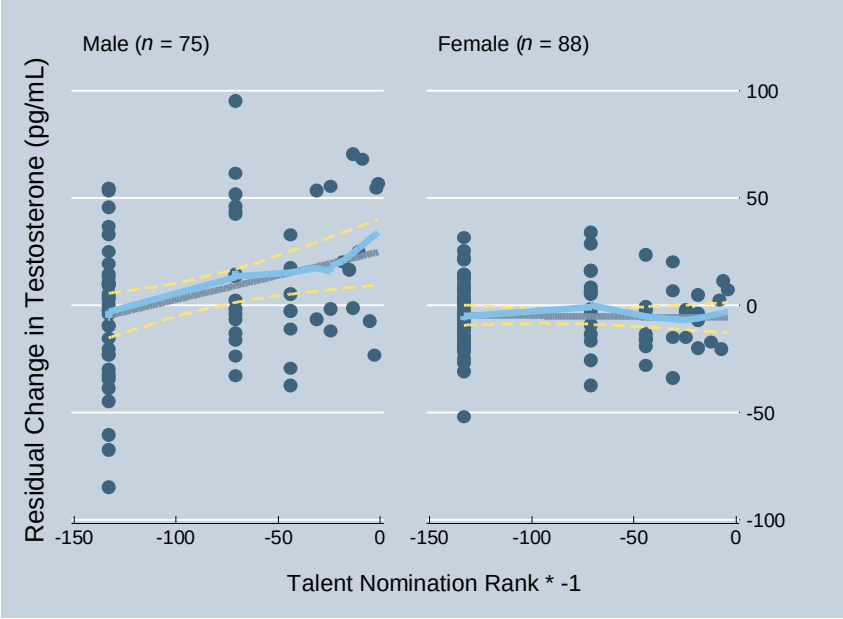
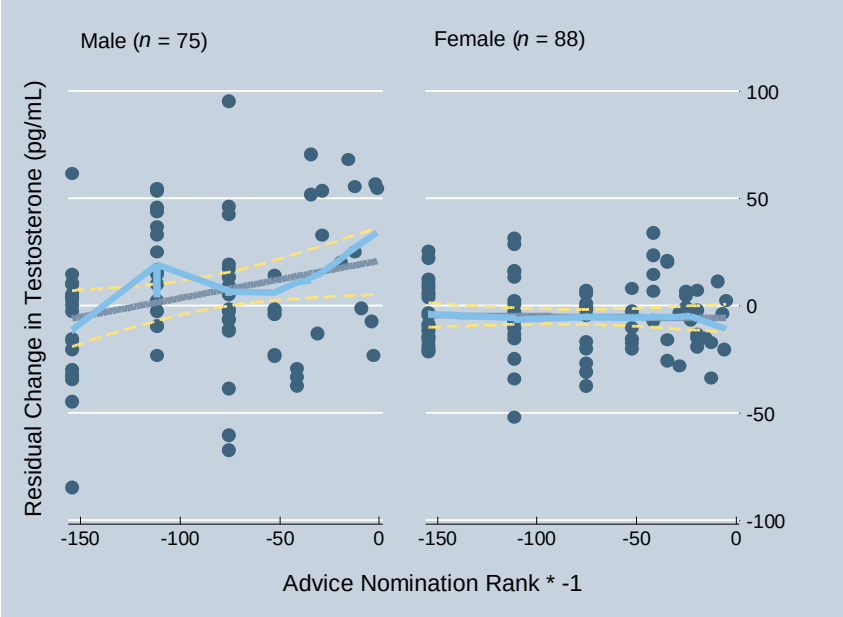


FIGURE S6. SCATTERPLOT OF CHANGE IN TESTOSTERONE (INDEXED USING RESIDUAL CHANGE SCORES) AS A FUNCTION OF ADVICE RANK FOR MEN AND WOMEN. THE LINE OF BEST FIT (IN GRAY), 95% CONFIDENCE INTERVAL (IN YELLOW DASH), AND LOWESS CURVE (IN LIGHT BLUE) FOR EACH PANEL ARE SHOWN.



DOES CHANGE IN PRESTIGE PREDICT CHANGE IN TESTOSTERONE?

A key prediction that may be derived from the influential biosocial model of status (Mazur, 1985), which proposes that primate physiology is highly reactive to and influences status allocation, is that changes in prestige may be associated with corresponding changes in testosterone. The current dataset, however, does not afford an optimal test of this question, because minimal changes in prestige standing were observed in the community. The number of nominations received at Time 1 and Time 2 were almost perfectly correlated for both talent ($r = .99$, $p < .0001$) and advice nominations ($r = .97$, $p < .0001$). This indicates almost perfect stability in prestige ranking (as assessed using the methodology employed here) across time, in that highly prestigious individuals at Time 1 retained high prestige at Time 2. This finding is consistent with existing field studies on groups and communities in humans and other primates, which also indicate that once emerged status differences generally remain highly stable overtime (Anderson, John, Keltner, & Kring, 2001; Bernstein, 1969; Bramblett, Bramblett, Bishop, & Coelho, 1982).

Nevertheless, while necessarily tentative given the limited changes in prestige ordering observed, we performed analyses to explore this prediction about the covariation between change in prestige and change in T. Paralleling the analyses reported in the main text, we regressed change in T on change in prestige ordering, also indexed using residual change scores that capture variation in Time 2 talent or advice nominations that are unexplained by Time 1 talent or advice, respectively. Either talent or advice nomination change was entered as the predictor in each of the two models, not together within the same model, given their moderate correlation (r = .27, p = .0005). In the two models, neither the coefficient on change in talent [b = .33, t(156) = .22, p = .827, .95 CI (-2.63, 3.29)] nor that on change in advice [b = -.26, t(156) = -.25, p = .800, .95 CI (-2.26, 1.74)] is significant. Subsequent models additionally using gender and its interaction with indices of prestige as predictors also yield non-significant coefficients on the change in talent × gender term [b = -2.73, t(154) = -.89, p = .372, .95 CI (-8.75, 3.30)] and the change in advice × gender term [b = -.10, t(154) = -.05, p = .960, .95 CI (-4.16, 3.96)], indicating a lack of evidence for gender difference in the (null) effect of prestige change on T change.

EXPLORING INTERACTIONS BETWEEN TESTOSTERONE AND CORTISOL

Much recent work has focused on the joint effects of T and cortisol (C) in coordinating dominant behaviors. This empirical phenomenon has received support in studies demonstrating that the effects of T in motivating and responding to social behavior vary across different levels of C (i.e., a T × C interaction), typically in acute contexts spanning minutes or hours (Mehta & Josephs, 2010; Mehta, Welker, Zilioli, & Carré, 2015; Ponzi, Zilioli, Mehta, Maslov, & Watson, 2016; Sherman, Lerner, Josephs, Renshon, & Gross, 2015; Zilioli & Watson, 2012), though not in some others (Mazur & Booth, 2014; for a review, see Mehta & Prasad, 2015). Our study, however, diverges from this work not only in its focus on the behavior-to-hormone link (rather than the hormone-to-behavior link),

654 but also by addressing the long-term, cumulative effect of prestige over months, which may
655 involve distinct processes. Moreover, theoretically, it is not clear that the effect of prestige
656 on changes in basal T is expected to be C-dependent. Thus, our analytic efforts focus on
657 isolated T effects.

658 Nevertheless, to supplement these analyses, we conducted a series of other
659 regression models to tentatively explore associations that may be anticipated by this
660 existing work on the joint effects of T and C. The first set of models explore whether T and C
661 interact to predict emergent prestige within the community. Table S24 below show 6
662 regression models using the main and interactive effects of T, C, and gender to predict our
663 prestige measures concurrently (hormones at Time 1 predicting prestige at Time 1, or
664 hormones at Time 2 predicting prestige at Time 2) or prospectively (hormones at Time 1
665 predicting prestige at Time 2). In all cases, except Model 2 (which predicts advice
666 nomination at Time 1), the coefficients on T × C and Gender × T × C terms did not reach
667 significance. However, even so, further analyses conducted separately on men and women
668 to probe the Gender × T × C interaction in Model 2 do not produce clear results. In these
669 subsequent models, no significant T × C interaction emerged in men [b = .04, t(79) = 1.38, p
670 = .171, .95 CI(-.02, .12)] or women [b = -.06, t(89) = -1.60, p = .113, .95 CI(-.13, .01)].

671 TABLE S24. OLS REGRESSIONS OF PRESTIGE ON TESTOSTERONE, CORTISOL, AND
672 TESTOSTERONE × CORTISOL INTERACTION.

	Model DV					
	Talent Nomination Time 1 (Model 1)	**Advice Nomination Time 1 (Model 2)**	**Talent Nomination Time 2 (Model 3)**	**Advice Nomination Time 2 (Model 4)**	**Talent Nomination Time 2 (Model 5)**	**Advice Nomination Time 2 (Model 6)**
Testosterone at Time 1	0.0367	0.1140+	0.0467	0.0984		
	(0.6116)	(0.0767)	(0.5280)	(0.1510)		
Log(Cortisol at Time 1)	-2.6336	-6.6515	-2.7425	-5.9778		
	(0.6125)	(0.1506)	(0.6120)	(0.2317)		
Testosterone at Time 1 × Log(Cortisol at Time 1)	-0.0023	0.0475	0.0028	0.0382		

	(0.9522)	(0.1717)	(0.9435)	(0.3001)		
Gender (1 = Female)	18.6129	24.1440*	17.7271	15.5501	45.8948	56.3581
	(0.1028)	(0.0176)	(0.1414)	(0.1619)	(0.3020)	(0.1678)
Gender × Testosterone at Time 1	-0.1434	-0.2114*	-0.1001	-0.0861		
	(0.1570)	(0.0194)	(0.3665)	(0.3995)		
Gender × Log(Cortisol at Time 1)	6.4985	11.5547*	4.0459	6.3497		
	(0.2633)	(0.0259)	(0.5038)	(0.2562)		
Gender × Testosterone at Time 1 × Log(Cortisol at Time 1)	-0.0375	-0.1073*	0.0200	-0.0301		
	(0.5134)	(0.0361)	(0.7438)	(0.5934)		
Log(Testosterone at Time 2)					8.5415	11.4183
					(0.2912)	(0.1249)
Log(Cortisol at Time 2)					-3.2049	-17.7731
					(0.8635)	(0.2997)
Log(Testosterone at Time 2) × Log(Cortisol at Time 2)					0.2714	3.9366
					(0.9471)	(0.2947)
Gender × Log(Testosterone at Time 2)					-9.1999	-13.1369
					(0.3587)	(0.1540)
Gender × Log(Cortisol at Time 2)					-0.7224	19.5817
					(0.9728)	(0.3141)
Gender × Log(Testosterone at Time 2) × Log(Cortisol at Time 2)					0.7533	-4.7521
					(0.8762)	(0.2852)
R^2	0.050	0.053	0.074	0.024	0.088	0.044
adj. R^2	0.010	0.014	0.031	-0.022	0.046	-0.000
AIC	1153.2545	1111.2752	1030.0865	1004.6452	1034.4224	1007.0685
BIC	1178.6184	1136.6391	1054.5873	1029.1460	1058.9736	1031.6198
N	176	176	158	158	159	159

p-values in parentheses.
The talent and advice nomination variables in these models as outcomes have been transformed using the natural logarithm.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

Our second set of models parallel those directly above and examine whether T and C interact to predict social popularity and success in building alliances, as indexed by peer-reported popularity and friendship nominations (in-coming), both of which are conceptually and empirically overlap with prestige-based status (see Table S6). As shown in Table S25 below, across all models the coefficients on T × C and Gender × T × C terms are non-significant, except for Model 2 (which predicts friendship nomination at Time 1), in which a significant T × C interaction emerged (but no gender interaction). However, further analyses conducted separately on men and women indicate only a marginally significant T × C interaction in men [b = .03, $t(79)$ = 1.59, p = .116, .95 CI(-.01, .06)] but not in women [b = -.004, $t(89)$ = .40, p = .692, .95 CI(-.02, .02)]. Note that this T × C interaction reaches statistical significance only in the combined sample (but not in either gender separately),

on only 1 of our 2 measures of social popularity, and has a very weak estimated effect size. Given the lack of clear results here, further examination in future research is needed to draw any conclusions.

TABLE S25. OLS REGRESSIONS OF SOCIAL POPULARITY ON TESTOSTERONE, CORTISOL, AND TESTOSTERONE × CORTISOL INTERACTION.

	Model DV					
	Popularity Nomination Time 1 (Model 1)	Friendship Nomination (In-Coming) Time 1 (Model 2)	Popularity Nomination Time 2 (Model 3)	Friendship Nomination (In-Coming) Time 2 (Model 4)	Popularity Nomination Time 2 (Model 5)	Friendship Nomination (In-Coming) Time 2 (Model 6)
Testosterone at Time 1	0.0115	0.0613*	-0.0101	-0.0016		
	(0.8759)	(0.0209)	(0.8955)	(0.9170)		
Log(Cortisol at Time 1)	-2.1641	-4.6681*	1.1746	0.4586		
	(0.6818)	(0.0147)	(0.8341)	(0.6795)		
Testosterone at Time 1 × Log(Cortisol at Time 1)	-0.0076	0.0288*	-0.0286	-0.0051		
	(0.8480)	(0.0447)	(0.4894)	(0.5307)		
Gender (1 = Female)	12.8921	11.0811**	15.2553	0.4931	35.0757	12.6903
	(0.2646)	(0.0082)	(0.2220)	(0.8415)	(0.4584)	(0.1618)
Gender × Testosterone at Time 1	-0.0152	-0.0605	-0.1860	0.0045		
	(0.8823)	(0.1020)	(0.1067)	(0.8424)		
Gender × Log(Cortisol at Time 1)	3.5324	4.8543*	4.9036	-0.4719		
	(0.5487)	(0.0228)	(0.4350)	(0.7040)		
Gender × Testosterone at Time 1 × Log(Cortisol at Time 1)	0.0227	-0.0247	-0.0628	0.0118		
	(0.6970)	(0.2392)	(0.3226)	(0.3464)		
Log(Testosterone at Time 2)					0.6112	2.5058
					(0.9433)	(0.1293)
Log(Cortisol at Time 2)					4.2570	-4.7207
					(0.8302)	(0.2150)
Log(Testosterone at Time 2) × Log(Cortisol at Time 2)					-0.7114	1.0535
					(0.8702)	(0.2070)
Gender × Log(Testosterone at Time 2)					-9.1454	-2.6094
					(0.3914)	(0.2018)
Gender × Log(Cortisol at Time 2)					7.9325	3.7813
					(0.7249)	(0.3810)
Gender × Log(Testosterone at Time 2) × Log(Cortisol at Time 2)					-2.2131	-0.6820
					(0.6675)	(0.4891)
R^2	0.054	0.061	0.063	0.035	0.031	0.061
adj. R^2	0.014	0.022	0.019	-0.010	-0.014	0.018
AIC	1158.5811	798.4085	1041.7183	529.8131	1054.3901	528.4633
BIC	1183.9450	823.7723	1066.2190	554.3138	1078.9413	553.0145
N	176	176	158	158	159	159

p-values in parentheses.
The popularity and friendship nomination variables in these models as outcomes have been transformed using the natural logarithm.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

In a third set of models, we investigated the joint effects of initial T and C at Time 1 in predicting relative change in T from Time to Time 2 and whether these effects may vary

for relatively more versus less prestigious individuals. These questions are motivated by recent laboratory studies pointing towards differential patterns of acute T change among winners and losers in competitive interactions, such that individuals with a unique high T low C profile rise in T following a win (Zilioli & Watson, 2012), and those with a high T high C profile decline in T following a defeat (Mehta & Josephs, 2010). To explore this, we regressed Time 2 T on the main effects of Time 1 T, Time 1 C, prestige-based status (indexed either by talent or advice nomination), gender, and the interaction among these variables.[3] Importantly, the inclusion of Time 1 T as a covariate in these models allows us to directly examine the effect of these other predictors in explaining *relative change* in T from Time 1 to Time 2 (i.e., their effects on residual T change from Time 1 to Time 2).

Table S26 contains the regression models examined using talent nominations, and Table S27 using advice nomination index. What is of greatest relevance here is the estimated coefficient on the Talent Nomination × T × C term and the Gender × Talent Nomination × T × C interaction term, both of which are estimated in Model 4 in Tables S26 and S27. Neither of these coefficients reached significance in the model using talent nomination (*p*s = .5450 and .2177, respectively), whereas the model using advice nomination estimated a significant Advice Nomination × T × C effect [b = -.1428, $t(147)$ = -2.45, p = .015, .95 CI(-.26, -.03)]. However, subsequent follow-up analysis aimed at further investigating this interactive effect separately for individuals with relatively high and low advice nominations (grouped in relation to the median, or the top vs. bottom 50% on this

[3] Inspection reveals that raw C concentrations were skewed, and thus log-transformed. Raw T concentrations were used. To facilitate model interpretation, T and C (both at Time 1) and the prestige index (talent nomination or advice nomination) were grand-mean centered before entry into models.

variable) did not produce any clear, conclusive results.[4] We regressed Time 2 T on Time 1 T (again to capture residual change in T), Time 1 C, and the Time 1 T × Time 1 C interaction, separately for prestigious individuals in the top 50% or bottom 50% of advice nominations received. Among those who received many advice nominations, the model estimated a non-significant T × C interaction term [b = -.1537, t(120) = -1.48, p = .143, .95 CI(-.36, .05)]. Similarly, among those who received few advice nominations, the T × C interaction term [b = .0441, t(35) = .23, p = .819, .95 CI(-.34, .43)] did not reach significance. These follow-up analyses suggest that the significant Advice Nomination × T × C interaction effect lacks robusticity.

Also relevant here is whether receiving prestige deference might produce variable degrees of change in T at different levels of C. A test of this question is provided by the prestige × C interaction term in the specification in Model 3. Across both indices of prestige, we found no consistent pattern indicating that C modulates the effect of perceived talent or being sought for advice on T change. As shown in Model 3 in Tables S26 and S27 below, the coefficient on neither the Time 1 C × prestige term (ps = .277 and .091) nor the gender × Time 1 C × prestige term (p = .346 and .372) reached conventional levels of significance. This suggests that the effect of winning prestige contests on increasing testosterone in men, as reported in the main text, does not reliably vary at different levels of C.

[4] Though prior work has primarily focused on acute T changes following victory and defeat in men only (e.g., Mehta & Josephs, 2010; Zilioli & Watson, 2012), the small coefficient estimated here for the Gender × Advice Nomination × Time 1 Cortisol × Time 1 T term and its non-significance suggest that in this dataset the effect of gender is minimal. Thus in our follow-up analysis presented here, we combined men and women for greater statistical power. In analyses not shown, however, we also examined T × C effects separately for men and women, and found no consistent evidence for T and C interactions in either gender.

TABLE S26. OLS REGRESSIONS OF TESTOSTERONE AT TIME 2 ON TESTOSTERONE AT TIME 1, CORTISOL AT TIME 1, TESTOSTERONE × CORTISOL INTERACTION, AND TALENT NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4
Testosterone at Time 1	0.7509***	0.5055***	0.4541***	0.3974***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Talent Nomination	0.8507*	1.8718***	2.0701***	2.1206**
	(0.0100)	(0.0000)	(0.0000)	(0.0042)
Gender (1 = Female)		-25.8010***	-30.5854***	-31.7950***
		(0.0000)	(0.0000)	(0.0000)
Gender × Talent Nomination		-1.9205**	-2.1536***	-1.9639+
		(0.0019)	(0.0007)	(0.0827)
Log(Cortisol at Time 1)			11.6030+	16.0657
			(0.0923)	(0.1213)
Talent Nomination × Log(Cortisol at Time 1)			1.1272	0.3129
			(0.2767)	(0.8519)
Gender × Log(Cortisol at Time 1)			-9.9342	-9.4995
			(0.2059)	(0.4746)
Gender × Talent Nomination × Log(Cortisol at Time 1)			-1.1775	-1.6168
			(0.3461)	(0.4590)
Talent Nomination × Testosterone at Time 1				-0.0033
				(0.8321)
Gender × Testosterone at Time 1				0.1304
				(0.4372)
Gender × Talent Nomination × Testosterone at Time 1				0.0027
				(0.9178)
Log(Cortisol at Time 1) × Testosterone at Time 1				-0.0112
				(0.9550)
Talent Nomination × Log(Cortisol at Time 1) × Testosterone at Time 1				0.0198
				(0.5450)
Gender × Log(Cortisol at Time 1) × Testosterone at Time 1				0.1725
				(0.5286)
Gender × Talent Nomination × Log(Cortisol at Time 1) × Testosterone at Time 1				-0.0553
				(0.2177)
R^2	0.606	0.666	0.676	0.684
adj. R^2	0.601	0.658	0.659	0.652
AIC	1529.2946	1506.1513	1509.4752	1519.2418
BIC	1538.5759	1521.6201	1537.3190	1568.7418
N	163	163	163	163

p-values in parentheses.

The talent nomination and cortisol variables have been transformed using the natural logarithm and subsequently grand-mean centered. Raw testosterone (without transformation) is also grand-mean centered.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

TABLE S27. OLS REGRESSIONS OF TESTOSTERONE AT TIME 2 ON TESTOSTERONE AT TIME 1, CORTISOL AT TIME 1, TESTOSTERONE × CORTISOL INTERACTION, AND ADVICE NOMINATIONS.

	Model 1	Model 2	Model 3	Model 4
Testosterone at Time 1	0.7461***	0.4948***	0.4431***	0.4631***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Advice Nomination	0.8020*	2.0963***	2.4657***	2.8447***
	(0.0370)	(0.0001)	(0.0000)	(0.0007)
Gender (1 = Female)		-26.8580***	-30.8479***	-27.7432***
		(0.0000)	(0.0000)	(0.0002)
Gender × Advice Nomination		-2.0319**	-2.2598**	-3.5933*
		(0.0044)	(0.0033)	(0.0198)
Log(Cortisol at Time 1)			2.9760	3.9968
			(0.6704)	(0.6917)
Advice Nomination × Log(Cortisol at Time 1)			2.4503+	7.2704**
			(0.0912)	(0.0022)
Gender × Log(Cortisol at Time 1)			-0.8070	1.2324
			(0.9207)	(0.9270)
Gender × Advice Nomination × Log(Cortisol at Time 1)			-1.4648	-7.5553*
			(0.3716)	(0.0259)

Advice Nomination × Testosterone at Time 1				-0.0048 (0.8194)
Gender × Testosterone at Time 1				0.1245 (0.4748)
Gender × Advice Nomination × Testosterone at Time 1				-0.0201 (0.5989)
Log(Cortisol at Time 1) × Testosterone at Time 1				0.0680 (0.7395)
Advice Nomination × Log(Cortisol at Time 1) × Testosterone at Time 1				-0.1428^{+} (0.0154)
Gender × Log(Cortisol at Time 1) × Testosterone at Time 1				0.0408 (0.8867)
Gender × Advice Nomination × Log(Cortisol at Time 1) × Testosterone at Time 1				0.1155 (0.1289)
R^2	0.600	0.663	0.674	0.698
adj. R^2	0.595	0.655	0.657	0.667
AIC	1531.6275	1507.5018	1510.0763	1511.9488
BIC	1540.9087	1522.9705	1537.9201	1561.4488
N	163	163	163	163

p-values in parentheses.

The advice nomination and cortisol variables have been transformed using the natural logarithm and subsequently grand-mean centered. Raw testosterone (without transformation) is also grand-mean centered.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$, *** $p < .001$

Together, across these three sets of analyses, no reliable evidence of T × C interactions were obtained in predicting the effective acquisition of prestige-based status, social popularity, or inter-individual change in T from Time 1 to Time 2. However, in performing these analyses, we were concerned about statistical power and our ability to accurately estimate and detect any interactive effects in the current data, including any T × C effects. Despite our sample size being relatively sizable for field studies of T and C, these data still afford relatively limited statistical power for tests of interactions, especially in the models that contain up to 15 predictors. In this respect, the limitation of small samples, which bedevils existing work on T and status (Geniole, Bird, Ruddick, & Carré, 2017; Salvador & Costa, 2009; van Anders & Watson, 2006), also applies to this and other studies that examine T × C effects (e.g., Mehta & Josephs, 2010; Zilioli & Watson, 2012). In future work we aim to conduct more well-powered tests of T and C interactions with these considerations in mind.

EFFECTS OF DISAGGREGATED INDIVIDUAL TESTOSTERONE MEASURES.

As described in our main text, we obtained two saliva samples at Time 1, one directly before the band rehearsal at approximately 3pm, and one immediately after the rehearsal at approximately 6pm. This procedure was repeated roughly 2 months later, at Time 2. This design yields 4 salivary T measures: Time 1 pre-rehearsal, Time 1 post-rehearsal, Time 2 pre-rehearsal, and Time 2 post-rehearsal. For the main theoretical findings in the main text and above, we examined T change using the within-day aggregate T measure (i.e., the mean across pre- and post-rehearsal T) from Time 1 to Time 2.

Use of this daily average measure rather than individual (pre- and post-rehearsal) assays is justified on the grounds of measurement accuracy and reliability. In terms of accuracy, our composite measure, which aggregates T release over two sampling times that span several hours, averages out noise introduced by T's diurnal rhythm. Although T follows a general rhythm that peaks in the morning and declines over the course of the day (dramatically before noon and more slowly in the afternoon and evening), individuals differ, however, in their specific rate of decline. This means that T concentrations at 3pm imperfectly predicts levels at 6pm, and each measurement provides some amount of unique, non-overlapping information. Thus, the average of T release at 3pm and 6pm ought to provide a more accurate picture of people's T release than each individual value. In addition, by aggregating pre- and post-rehearsal, we also average out noise resulting from the experience of events that occur. For example, we suspect that a variety of factors and events that occur during our band rehearsal context—be it the mere act of being a part of a social group, interactions with peers, engagement with the marching and musical activities, competitive mindset whilst rehearsing for an upcoming marching competition or

performance, or the mere anticipation of these events—may modulate T release. Past work has, in fact, shown that simply attending a sports event as a spectator and vicariously observing competition can lead to changes in T (Bernhardt, Dabbs Jr, Fielden, & Lutter, 1998). This suggests that relying on a single sample of either pre or post-rehearsal T levels might introduce noise and skew results.

In terms of reliability, existing research in neuroendocrinology indicates that, compared to single sample disaggregated analytic methods, the aggregation approach should be used whenever possible to derive more reliable assessments of inter-individual variation in endocrine activity (and by implication, hormone change; Dariotis, Chen, & Granger, 2016; Gunnar, 2001; Hellhammer et al., 2007; Pruessner et al., 1997). Given that it is these measures of T at each time point from which our measure of change in T subsequently derive, we expect reliability to be an especially crucial consideration here given that, complicating matters even further, change scores are critiqued for being generally less reliable than the component variables (i.e., Time 1 and Time 2 T; Allison, 1990; Kessler, 1977; Rogosa & Willett, 1983). This means that using reliable Time 1 and Time 2 T measures is crucial for obtaining the most reliable T change measure possible. Thus, as in a number of studies of T (e.g., Cadoux-Hudson, Few, & Imms, 1985; Granger et al., 2003; Granger, Shirtcliff, Booth, Kivlighan, & Schwartz, 2004; Johnsen & Zuk, 1995; Welling et al., 2008), we used T aggregates by averaging pre- and post-rehearsal concentrations to maximize the reliability of our Time 1 and Time 2 measures of T, with the aim of increasing the reliability of the Time 1 to Time 2 change in T measure that we are ultimately seeking to explain.

Nevertheless, with these limitations of single time-point assessments in mind, we explored patterns of prestige-dependent changes in T based on individual T assays by performing regression analyses parallel to those above using 5 additional ways of examining T change (using disaggregated T measures): (1) pre- to post-rehearsal at Time 1 (acute changes in T during band practice at Time 1; no Time 2 data); (2) pre- to post-rehearsal at Time 2 (acute changes in T during band practice at Time 2; no Time 1 data); (3) pre-rehearsal at Time 1 to pre-rehearsal at Time 2 (longitudinal change; no post-rehearsal data); (4) post-rehearsal at Time 1 to post-rehearsal at Time 2 (longitudinal change; no pre-rehearsal data); and relatedly (5) post-rehearsal at Time 1 (controlling for pre-rehearsal) to post-rehearsal at Time 2 (controlling for pre-rehearsal). In all models, the specified form of T change is regressed onto our prestige index, gender, and the prestige × gender interaction, and the suite of control variables deployed above. For brevity, we present and discuss the specification with the full set of controls (the specification in Model 9 from Tables S9, S10, S19, and S21).

Tables S28-S32 show, for each of the 5 change time points modeled (1-5 described above, in order), a series of regression models using the prestige index (either talent or advice, and either nomination count or rank) to predict residual or raw change scores. In the models addressing same-day, relatively acute T changes from pre- to post-rehearsal (outcomes 1 and 2), no prestige effects were significant at conventional levels for any specification using outcomes (1) and (2), which address prestige-dependent acute changes in T occurring over the course of band practice, with two exceptions. That is, in Table S29 Models 3 and 4, advice nomination counts significantly predicts both residual and raw change from pre- to post-rehearsal at Time 2 (outcome 2; *p*s = .010 and .031), but this

effect is far from reaching significance in the other specifications using the talent nomination count index or any of the prestige nomination count indices (talent or advice). Together, these results indicate that acute T responses occurring over the course of the band rehearsal is unlikely to be responsible for our major findings.

In models addressing longitudinal T changes from Time 1 to Time 2, a clear and consistent pattern of results emerged in the models predicting post-rehearsal T (outcome 4; Table S31). In all specifications, the coefficient on prestige was significant and significant across the board (the only exception was the marginal effect in Model 8, $p = .121$). These results are consistent with our primary findings reported in the main text and above based on aggregate measures of T. Outcome 5 (in Table S32) is similar to outcome 4 but differs in that here pre-rehearsal levels were partialled from post-rehearsal levels at both Time 1 and Time 2. This measure of longitudinal change allows us to conduct a preliminary exploration of whether our major finding can be alternatively explained by an acute within-day T increase from pre- to post-rehearsal that might be stronger among more prestigious individuals at Time 2 compared to Time 1—a question not addressed by outcome 4 because the post-rehearsal T longitudinal change measure examined there partly reflects any ongoing T activity pre-rehearsal. Results across models using outcome 5 show that the prestige effect largely evaporates when pre-rehearsal inter-individual variability is removed. The coefficient on prestige (indicating the simple slope in men) reaches significance in only 1 of the 8 specifications examined. Model 3 gives the prestige coefficient as 1.62 ($p = .032$), but in the remaining specifications the coefficient ranges from .01 ($p = .891$) to 1.33 ($p = .186$). The absence of a consistent significant effect of prestige in these models predicting outcome 5, combined with the null effects in models predicting

858 outcomes 1 and 2 above, suggests that within-day acute changes in T and how they change
859 longitudinally over time are unlikely to be responsible for our pattern of results.

860 Finally, in the models predicting pre-rehearsal T change across time (outcome 3;
861 Table S30), however, the coefficient on prestige, despite being positive, did not reach
862 significance. The models estimate the coefficient on our prestige index to be 0.89 and 0.68
863 (*p*s = .137 and .237) for the talent and advice nomination count measures, respectively, in
864 predicting residual change.

865 Overall, these analyses indicate that, using single time-point T assessments, the
866 pattern of results was strongest for longitudinal (i.e., Time 1 to Time 2) post-rehearsal
867 change and consistent with our primary results based on aggregated T measures. Though
868 the results for pre-rehearsal change were in the expected direction, they are weaker and
869 did not reach conventional levels of significance.

870 Note that, in our view, although these associations and comparisons of change
871 across different time points are interesting, we emphasize that, independent of these
872 results, as mentioned above there are reasons to expect the lowered accuracy and
873 reliability of these single time-point T assessments (and the resultant T change measure) to
874 work against the detection of trends. It is interesting to note, however, that the absence of a
875 robust association between prestige and pre- to post-rehearsal T change at either Time 1
876 or Time 2 provides suggestive evidence that acute, same-day changes are unlikely to
877 explain our general findings. Nevertheless, our view is that these results should be treated
878 as tentative. Future work should examine these effects (i.e., same-day acute T changes)
879 using multiple hormone assessments to better capture endocrine activity.

TABLE S28. OLS REGRESSIONS OF TESTOSTERONE CHANGE FROM PRE- TO POST-REHEARSAL AT TIME 1-ONLY ON PRESTIGE. THESE MODELS INCLUDE ADDITIONAL CONTROL VARIABLES (OTHER NOMINATION VARIABLES—COERCION INDEX, POPULARITY INDEX, FRIENDSHIP IN-COMING INDEX, FRIENDSHIP OUT-GOING INDEX—AND EACH OF THEIR INTERACTION WITH GENDER, AND AGE, ETHNICITY, MARCHING BAND EXPERIENCE, AND SECTION LEADERSHIP STATUS).

	DV = Residual Change Prestige Index = Talent Nominations Received (Model 1)	DV = Raw Change Prestige Index = Talent Nominations Received (Model 2)	DV = Residual Change Prestige Index = Advice Nominations Received (Model 3)	DV = Raw Change Prestige Index = Advice Nominations Received (Model 4)	DV = Residual Change Prestige Index = Talent Nomination Rank (Model 5)	DV = Raw Change Prestige Index = Talent Nomination Rank (Model 6)	DV = Residual Change Prestige Index Advice Nomination Rank (Model 7)	DV = Raw Change Prestige Index = Advice Nomination Rank (Model 8)
Prestige Index	-0.1178	-0.5010	0.5494	-0.0989	-0.0813	-0.1338	-0.0470	-0.1331
	(0.8752)	(0.6464)	(0.5384)	(0.9393)	(0.4596)	(0.4041)	(0.7213)	(0.4888)
Gender (1 = Female)	-33.2631**	2.2928	-31.5885**	4.2791	-42.4672**	4.1920	-43.2788**	2.7514
	(0.0034)	(0.8879)	(0.0041)	(0.7865)	(0.0034)	(0.8406)	(0.0026)	(0.8939)
Gender × Prestige Index	-0.1128	0.6801	0.4939	1.8048	0.0153	0.1227	-0.0464	0.1204
	(0.9179)	(0.6691)	(0.6740)	(0.2913)	(0.9245)	(0.6021)	(0.7944)	(0.6430)
N	165	165	165	165	165	165	165	165

p-values in parentheses.

Index variables (derived from nominations) are either nomination counts received or relative ranking, and in all cases are consistent with the prestige index deployed in that model. That is, in Models 1-4, all indices derived from nomination data are log-transformed nomination counts, and in Models 5-8, all indices used are assigned relative ranks (reversed) computed from raw nomination counts. All nomination count received variables were log-transformed to reduce skew, and all relative ranking variables were reversed by multiplying by -1 so that higher values indicate higher prestige.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < .01$, $^{***}\ p < .001$

TABLE S29. OLS REGRESSIONS OF TESTOSTERONE CHANGE FROM PRE- TO POST-REHEARSAL AT TIME 2-ONLY ON PRESTIGE. THESE MODELS INCLUDE ADDITIONAL CONTROL VARIABLES (OTHER NOMINATION VARIABLES—COERCION INDEX, POPULARITY INDEX, FRIENDSHIP IN-COMING INDEX, FRIENDSHIP OUT-GOING INDEX—AND EACH OF THEIR INTERACTION WITH GENDER, AND AGE, ETHNICITY, MARCHING BAND EXPERIENCE, AND SECTION LEADERSHIP STATUS).

	DV = Residual Change Prestige Index = Talent Nominations Received (Model 1)	DV = Raw Change Prestige Index = Talent Nominations Received (Model 2)	DV = Residual Change Prestige Index = Advice Nominations Received (Model 3)	DV = Raw Change Prestige Index = Advice Nominations Received (Model 4)	DV = Residual Change Prestige Index = Talent Nomination Rank (Model 5)	DV = Raw Change Prestige Index = Talent Nomination Rank (Model 6)	DV = Residual Change Prestige Index Advice Nomination Rank (Model 7)	DV = Raw Change Prestige Index = Advice Nomination Rank (Model 8)
Prestige Index	0.6455	0.4454	1.8911*	1.6478*	0.0007	-0.0240	0.1312	0.1064
	(0.3254)	(0.5110)	(0.0104)	(0.0312)	(0.9939)	(0.8048)	(0.2158)	(0.3341)
Gender (1 = Female)	-14.9945	-8.6184	-15.1956+	-9.1121	-18.7250	-10.4254	-17.9160	-9.8241
	(0.1223)	(0.3886)	(0.0937)	(0.3321)	(0.1138)	(0.3947)	(0.1222)	(0.4138)
Gender × Prestige Index	-0.3939	-0.1712	-0.4967	-0.2404	-0.0382	-0.0076	-0.0518	-0.0194
	(0.6779)	(0.8613)	(0.6230)	(0.8189)	(0.7781)	(0.9571)	(0.7270)	(0.9001)
N	148	148	148	148	148	148	148	148

p-values in parentheses.

Index variables (derived from nominations) are either log-transformed nomination counts or relative ranking, and in all cases are consistent with the prestige index deployed in that model. That is, in Models 1-4, all indices derived from nomination data are log-transformed nomination counts, and in Models 5-8, all indices used are assigned relative ranks computed from raw nomination counts.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < .01$, $^{***}\ p < .001$

TABLE S30. OLS REGRESSIONS OF TESTOSTERONE CHANGE PRE-REHEARSAL FROM TIME 1 TO TIME 2 ON PRESTIGE (NO POST-REHEARSAL DATA). THESE MODELS INCLUDE ADDITIONAL CONTROL VARIABLES (OTHER NOMINATION VARIABLES—COERCION INDEX, POPULARITY INDEX, FRIENDSHIP IN-COMING INDEX, FRIENDSHIP OUT-GOING INDEX—AND EACH OF THEIR INTERACTION WITH GENDER, AND AGE, ETHNICITY, MARCHING BAND EXPERIENCE, AND SECTION LEADERSHIP STATUS).

	DV = Residual Change Prestige Index = Talent Nominations Received (Model 1)	DV = Raw Change Prestige Index = Talent Nominations Received (Model 2)	DV = Residual Change Prestige Index = Advice Nominations Received (Model 3)	DV = Raw Change Prestige Index = Advice Nominations Received (Model 4)	DV = Residual Change Prestige Index = Talent Nomination Rank (Model 5)	DV = Raw Change Prestige Index = Talent Nomination Rank (Model 6)	DV = Residual Change Prestige Index Advice Nomination Rank (Model 7)	DV = Raw Change Prestige Index = Advice Nomination Rank (Model 8)
Prestige Index	0.8887	0.7772	0.6813	0.4212	0.1369	0.1308	-0.0092	-0.0580
	(0.1368)	(0.2367)	(0.3244)	(0.5792)	(0.1224)	(0.1770)	(0.9299)	(0.6113)
Gender (1 = Female)	-7.1822	8.1107	-6.6883	8.1796	-6.6507	14.5375	-6.6066	14.0690
	(0.4191)	(0.4075)	(0.4380)	(0.3887)	(0.5652)	(0.2517)	(0.5661)	(0.2642)
Gender × Prestige Index	-0.6184	-0.4136	-1.3316	-1.2283	-0.0755	-0.0442	-0.0622	-0.0205
	(0.4769)	(0.6656)	(0.1676)	(0.2467)	(0.5610)	(0.7555)	(0.6720)	(0.8983)
N	147	147	147	147	147	147	147	147

p-values in parentheses.

Index variables (derived from nominations) are either log-transformed nomination counts or relative ranking, and in all cases are consistent with the prestige index deployed in that model. That is, in Models 1-4, all indices derived from nomination data are log-transformed nomination counts, and in Models 5-8, all indices used are assigned relative ranks computed from raw nomination counts.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < .01$, $^{***}\ p < .001$

TABLE S31. OLS REGRESSIONS OF TESTOSTERONE CHANGE POST-REHEARSAL FROM TIME 1 TO TIME 2 ON PRESTIGE (NO PRE-REHEARSAL SAMPLE). THESE MODELS INCLUDE ADDITIONAL CONTROL VARIABLES (OTHER NOMINATION VARIABLES—COERCION INDEX, POPULARITY INDEX, FRIENDSHIP IN-COMING INDEX, FRIENDSHIP OUT-GOING INDEX—AND EACH OF THEIR INTERACTION WITH GENDER, AND AGE, ETHNICITY, MARCHING BAND EXPERIENCE, AND SECTION LEADERSHIP STATUS).

	DV = Residual Change Prestige Index = Talent Nominations Received (Model 1)	DV = Raw Change Prestige Index = Talent Nominations Received (Model 2)	DV = Residual Change Prestige Index = Advice Nominations Received (Model 3)	DV = Raw Change Prestige Index = Advice Nominations Received (Model 4)	DV = Residual Change Prestige Index = Talent Nomination Rank (Model 5)	DV = Raw Change Prestige Index = Talent Nomination Rank (Model 6)	DV = Residual Change Prestige Index Advice Nomination Rank (Model 7)	DV = Raw Change Prestige Index = Advice Nomination Rank (Model 8)
Prestige Index	2.4936**	2.5332**	3.3448***	3.0037**	0.3148**	0.3469**	0.2776*	0.2370
	(0.0011)	(0.0044)	(0.0002)	(0.0041)	(0.0043)	(0.0074)	(0.0327)	(0.1209)
Gender (1 = Female)	-26.4960*	-1.0332	-23.7155*	1.4407	-35.4224*	-3.7535	-32.1235*	-0.1275
	(0.0167)	(0.9355)	(0.0258)	(0.9081)	(0.0109)	(0.8163)	(0.0209)	(0.9937)
Gender × Prestige Index	-1.9674^{+}	-1.7863	-2.5185*	-2.5257^{+}	-0.3062^{+}	-0.2933	-0.1538	-0.0675
	(0.0729)	(0.1615)	(0.0325)	(0.0695)	(0.0549)	(0.1170)	(0.3799)	(0.7437)
N	154	154	154	154	154	154	154	154

p-values in parentheses.

Index variables (derived from nominations) are either log-transformed nomination counts or relative ranking, and in all cases are consistent with the prestige index deployed in that model. That is, in Models 1-4, all indices derived from nomination data are log-transformed nomination counts, and in Models 5-8, all indices used are assigned relative ranks computed from raw nomination counts.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < .01$, $^{***}\ p < .001$

TABLE S32. OLS REGRESSIONS OF TESTOSTERONE CHANGE POST-REHEARSAL FROM TIME 1 TO TIME 2 ON PRESTIGE (CONTROLLING FOR PRE-REHEARSAL TESTOSTERONE). THESE MODELS INCLUDE ADDITIONAL CONTROL VARIABLES (OTHER NOMINATION VARIABLES—COERCION INDEX, POPULARITY INDEX, FRIENDSHIP IN-COMING INDEX, FRIENDSHIP OUT-GOING INDEX—AND EACH OF THEIR INTERACTION WITH GENDER, AND AGE, ETHNICITY, MARCHING BAND EXPERIENCE, AND SECTION LEADERSHIP STATUS).

	DV = Residual Change Prestige Index = Talent Nominations Received (Model 1)	DV = Raw Change Prestige Index = Talent Nominations Received (Model 2)	DV = Residual Change Prestige Index = Advice Nominations Received (Model 3)	DV = Raw Change Prestige Index = Advice Nominations Received (Model 4)	DV = Residual Change Prestige Index = Talent Nomination Rank (Model 5)	DV = Raw Change Prestige Index = Talent Nomination Rank (Model 6)	DV = Residual Change Prestige Index Advice Nomination Rank (Model 7)	DV = Raw Change Prestige Index = Advice Nomination Rank (Model 8)
Prestige Index	0.4974	1.0747	1.6217*	1.3282	0.0126	0.1547	0.0921	0.0973
	(0.4471)	(0.2083)	(0.0317)	(0.1861)	(0.8912)	(0.2149)	(0.3846)	(0.4980)
Gender (1 = Female)	-9.6437	18.2682	-9.6840	19.4429	-13.2463	19.2364	-12.2406	21.3241
	(0.3095)	(0.1404)	(0.2786)	(0.1052)	(0.2620)	(0.2278)	(0.2914)	(0.1761)
Gender × Prestige Index	-0.1032	-0.2729	-0.3606	-0.4936	-0.0217	-0.0790	0.0084	0.0997
	(0.9125)	(0.8236)	(0.7229)	(0.7170)	(0.8712)	(0.6620)	(0.9548)	(0.6187)
N	144	144	144	144	144	144	144	144

p-values in parentheses.

Index variables (derived from nominations) are either log-transformed nomination counts or relative ranking, and in all cases are consistent with the prestige index deployed in that model. That is, in Models 1-4, all indices derived from nomination data are log-transformed nomination counts, and in Models 5-8, all indices used are assigned relative ranks computed from raw nomination counts.

The ethnicity dummies use Caucasian as the reference group, so the coefficient on each dummy variable gives effects relative to Caucasian.

$^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < .01$, $^{***}\ p < .001$

EXTENDED DISCUSSION

PRIOR EVIDENCE SUPPORTS NOTION THAT A HORMONAL PROFILE OF HIGH-RANK (I.E., A STRONG TESTOSTERONE-SOCIAL RANK LINK) IS UNLIKELY

A review of the relevant empirical literature reveals that, as in the present dataset, the association between T and indices of emergent rank (whether imposed through force, merit, or a mix of the two) tends to be weak or null in most studies, casting serious doubt on any robust and straightforward link between the two. Despite some early findings that T is positively (albeit weakly) correlated with rank in many species, including non-human male primates (Rose, Holaday, & Bernstein, 1971), most subsequent studies have generated conflicting results (Eaton & Resko, 1974; Gordon, Rose, & Bernstein, 1976; for a review, see Sapolsky, 1991). Most studies of humans have similarly produced null

associations. In personality research, T is found to be uncorrelated with trait dominance, assertiveness, competitiveness, and other status-relevant traits that serve as proxies to individuals' rank experiences in their everyday social relationships (Akinola, Page-Gould, Mehta, & Lu, 2016; Dabbs Jr., Hopper, & Jurkovic, 1990; Johnson, Burk, & Kirkpatrick, 2007; Josephs, Sellers, Newman, & Mehta, 2006; Slatcher, Mehta, & Josephs, 2011; van der Meij, Buunk, van de Sande, & Salvador, 2008). Moreover, confirming these findings, a recent meta-analysis reveals the lack of any straightforward effect of T on trait dominance and rank and power in real-world situations (van der Meij, Schaveling, & van Vugt, 2016).

In perhaps the most relevant study that assessed actual rank distributions in small laboratory groups (Mazur, Welker, & Peng, 2015), groups of three men took part in a leaderless, unguided 10-minute conversation. Emergent rank and leadership—as measured by a combination of variables including speaking time and group members' nominations of who led the group—was not found to be significantly associated with basal T assessed either pre- or post-interaction, even in groups for which high rank is rewarded with monetary incentives (paralleling the evolutionary incentives to high-ranking individuals). Finally, in a field study of male executives, similarly no straightforward associations emerged between T and rank; those who manage and supervise a greater number of subordinates, which indicates achieving higher rank in the modern workplace, were not found to have a higher basal T, though they appear to possess a unique hormonal profile of high T and low cortisol (Sherman et al., 2015).

In light of this existing evidence, our finding in this dataset—that at Time 1 (that is, in the initial weeks of the group's formation) the observed associations between T and our measures of prestige, despite being in the positive direction, were modest and did not

reach conventional levels of significance (*r*s = .13 and .16 for talent and advice nominations, respectively)—is not only non-surprising, but in fact anticipated by prior work. Our results add to and complement this existing body of evidence, which together challenge the folk wisdom and notion that there exists a single, robust physiological determinant (or a set of determinants) of rank in primates (Eisenegger, Naef, Snozzi, Heinrichs, & Fehr, 2010; Knight & Mehta, 2017; Sapolsky, 1991; Whitten, 2000).

WHY HAS TESTOSTERONE AT TIME 1 NOT RISEN IN RESPONSE TO CONCURRENT OR PRE-EXISTING PRESTIGE STANDING?

Our results reveal that, consistent with our primary finding of a rising T profile among highly prestigious men, prestige standing at Time 1 prospectively predicts T concentrations at Time 2 in men, but is not concurrently associated with T at Time 1 (though it trends in the predicted positive direction) in either men or women. This pattern opens up the question of why, prestige-based rank in the initial weeks of the group's formation is capable of modulating T reactivity patterns two months later, but its effect on T is not already detectable in these initial weeks? That is, why hasn't prestige experience in the first weeks already led to spikes in T to produce a correlation between Time 1 prestige and T? One potential explanation for these observed results is the issue of (lack of) time. In this large organization of over 200 individuals, figuring out the one's position in the local prestige hierarchy requires, among many other things, the time and opportunity to repeatedly interact with other group members, observe the deference signals directed from others toward the self, and refine and update one's assessment of the broader distribution of deference in the local group and one's position vis-à-vis this hierarchy. Thus, in these initial weeks during which individuals are still likely accumulating information about the

emerging prestige hierarchy and performing cognitive assessments of their own standing within, any endocrine changes in response to experiences at this time are expected to be minimal.

A related question is whether, among the returning band members (who constitute approximately 62% of our sample; see Table S19 below for frequency breakdown of marching band experience), T at the start of the current season (i.e., Time 1) has already risen among the more prestigious men, who might have also enjoyed high prestige in their former community in the year prior (perhaps as a result of their superior musical skills, knowledge, talent, or other relatively stable, locally valued attributes). We find no evidence that this is the case; among returning male members (n = 49), Time 1 T is uncorrelated with talent (r = .10; p = .4831) or advice (r = .15; p = .2913) nominations. Note, however, that this test only makes sense if prestige at Time 1 in the current community indeed tracks prestige in the community prior, an empirical question that we lack data from prior band seasons to evaluate. Thus, any interpretation of these results must remain tentative.

TABLE S33. DESCRIPTIVE SUMMARY OF MARCHING BAND EXPERIENCE (IN YEARS) IN POOLED SAMPLE.

Marching Band Experience (years, including the current)	Frequency	Percent	Cumulative Percent
1	61	34.46	34.46
2	50	28.25	62.71
3	37	20.9	83.62
4	13	7.34	90.96
5	7	3.95	94.92
6	2	1.13	96.05
Unknown (not reported)	7	3.95	100
Total	177	100	-

We suspect that, even if continuity exists in men's prestige from year to year (and by implication, band to band), elevations in T experienced by returning members who earned substantial prestige in the previous community are expected to have largely waned in the summer months prior to our study, when the organization disbanded. That is, when experiences (and reminders) of social victory and deference ceased, the elevated T of these previously prestigious individuals is expected to return to their baseline levels, an individual difference that is underlain in part by genetic components (Crabbe et al., 2007).

This flexibility has advantages over a persistently elevated T profile, given the significant costs associated with maintaining high androgen levels, including increased energetic demands, depressed immune function, and increased risk of parasitic infestation (and associated mortality; see Wingfield, Lynn, & Soma, 2001). This means that, just as T should rise in response to perceiving a prestige-status asymmetry in one's favor, T should also wane when this asymmetry is no longer reinforced or the local environment ceases to present opportunities for status-advancement (such as when the community disbands, as in our sample). In one study that demonstrates this cost-benefit trade-off for T, male baboons showed elevated T levels during critical periods when high rank was being contested, counterbalanced by decreased T levels during non-critical periods (Beehner, Bergman, Cheney, Seyfarth, & Whitten, 2006; also see Knight & Mehta, 2017). Thus, the lack of a significant positive association between prestige and T at Time 1 in our subsample of men is not entirely surprising. The logic above suggests that, if anything, it may reflect the dampening and return of T to baseline in previously high-ranking men, consistent with the hallmark of an adaptive system.

REFERENCES

Akinola, M., Page-Gould, E., Mehta, P. H., & Lu, J. G. (2016). Collective hormonal profiles predict group performance. *Proceedings of the National Academy of Sciences*, *113*(35), 9774–9779. https://doi.org/10.1073/pnas.1603443113

Alberts, S. C., & Altmann, J. (1995). Preparation and activation: determinants of age at reproductive maturity in male baboons. *Behavioral Ecology and Sociobiology*, *36*(6), 397–406. https://doi.org/10.1007/BF00177335

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology*, *20*(1), 93–114.

Anderson, C., John, O. P., Keltner, D., & Kring, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology*, *81*(1), 116–132. https://doi.org/10.1037/0022-3514.81.1.116

Archie, E. A., Altmann, J., & Alberts, S. C. (2012). Social status predicts wound healing in wild baboons. *Proceedings of the National Academy of Sciences*, *109*(23), 9017–9022. https://doi.org/10.1073/pnas.1206391109

Barnett, V., & Lewis, T. (1994). *Outliers in statistical data* (3rd Ed.). West Sussex, UK: John Wiley & Sons.

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*(4), 469–479. https://doi.org/10.1007/s00265-005-0071-2

Bernhardt, P. C., Dabbs Jr, J. M., Fielden, J. A., & Lutter, C. D. (1998). Testosterone changes during vicarious experiences of winning and losing among fans at sporting events. *Physiology & Behavior*, *65*(1), 59–62. https://doi.org/10.1016/S0031-9384(98)00147-4

Bernstein, I. S. (1969). Stability of the status hierarchy in a pigtail monkey group (Macaca nemestrina). *Animal Behaviour*, *17*, 452–458. https://doi.org/10.1016/0003-3472(69)90146-8

Bernstein, I. S. (1981). Dominance: The baby and the bathwater. *Behavioral and Brain Sciences*, *4*(3), 419–429.

Bramblett, C. A., Bramblett, S. S., Bishop, D. A., & Coelho, A. M. (1982). Longitudinal stability in adult status hierarchies among vervet monkeys (Cercopithecus aethiops). *American Journal of Primatology*, *2*(1), 43–51. https://doi.org/10.1002/ajp.1350020107

Burt, K. B., & Obradović, J. (2013). The construct of psychophysiological reactivity: Statistical and psychometric issues. *Developmental Review*, *33*(1), 29–57. https://doi.org/10.1016/j.dr.2012.10.002

Cadoux-Hudson, T. A., Few, J. D., & Imms, F. J. (1985). The effect of exercise on the production and clearance of testosterone in well trained young men. *European Journal of Applied Physiology and Occupational Physiology*, *54*(3), 321–325. https://doi.org/10.1007/BF00426153

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103–125. https://doi.org/10.1037/a0030398

Cheng, J. T., Tracy, J. L., & Henrich, J. (2010). Pride, personality, and the evolutionary foundations of human social status. *Evolution and Human Behavior*, *31*(5), 334–347. https://doi.org/10.1016/j.evolhumbehav.2010.02.004

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*, 829–836.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.

Crabbe, P., Bogaert, V., De Bacquer, D., Goemaere, S., Zmierczak, H., & Kaufman, J. M. (2007). Part of the interindividual variation in serum testosterone levels in healthy men reflects differences in androgen sensitivity and feedback set point: Contribution of the androgen receptor polyglutamine tract polymorphism. *The Journal of Clinical Endocrinology and Metabolism*, *92*(9), 3604–3610. https://doi.org/10.1210/jc.2007-0117

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin*, *74*(1), 68–80. https://doi.org/10.1037/h0029382

Dabbs Jr., J. M., Hopper, C. H., & Jurkovic, G. J. (1990). Testosterone and personality among college students and military veterans. *Personality and Individual Differences*, *11*(12), 1263–1269. https://doi.org/10.1016/0191-8869(90)90153-I

Dariotis, J. K., Chen, F. R., & Granger, D. A. (2016). Latent trait testosterone among 18–24 year olds: Methodological considerations and risk associations. *Psychoneuroendocrinology*, *67*, 1–9. https://doi.org/10.1016/j.psyneuen.2016.01.019

Eaton, G. G., & Resko, J. A. (1974). Plasma testosterone and male dominance in a Japanese macaque (Macaca fuscata) troop compared with repeated measures of testosterone in laboratory males. *Hormones and Behavior*, *5*(3), 251–259. https://doi.org/10.1016/0018-506X(74)90033-6

Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., & Fehr, E. (2010). Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature*, *463*(7279), 356–359. https://doi.org/10.1038/nature08711

Geniole, S. N., Bird, B. M., Ruddick, E. L., & Carré, J. M. (2017). Effects of competition outcome on testosterone concentrations in humans: An updated meta-analysis. *Hormones and Behavior*, *92*, 37–50. https://doi.org/10.1016/j.yhbeh.2016.10.002

Gordon, T. P., Rose, R. M., & Bernstein, I. S. (1976). Seasonal rhythm in plasma testosterone levels in the rhesus monkey (Macaca mulatta): A three year study. *Hormones and Behavior*, *7*(2), 229–243. https://doi.org/10.1016/0018-506X(76)90050-7

Granger, D. A., Shirtcliff, E. A., Booth, A., Kivlighan, K. T., & Schwartz, E. B. (2004). The "trouble" with salivary testosterone. *Psychoneuroendocrinology*, *29*(10), 1229–1240. https://doi.org/10.1016/j.psyneuen.2004.02.005

Granger, D. A., Shirtcliff, E. A., Zahn–Waxler, C., Usher, B., Klimes–Dougan, B., & Hastings, P. (2003). Salivary testosterone diurnal variation and psychopathology in adolescent males and females: Individual differences and developmental effects. *Development and Psychopathology*, *15*(2), 431–449. https://doi.org/10.1017/S0954579403000233

Griffin, D., Murray, S., & Gonzalez, R. (1999). Difference score correlations in relationship research: A conceptual primer. *Personal Relationships*, *6*(4), 505–518. https://doi.org/10.1111/j.1475-6811.1999.tb00206.x

Gunnar, M. R. (2001). The role of glucocorticoids in anxiety disorders: A critical analysis. In M. W. Vasey & M. R. Dadds (Eds.), *The developmental psychopathology of anxiety* (pp. 143–159). New York, NY: Oxford University Press. https://doi.org/10.1093/med:psych/9780195123630.003.0007

Hamilton, L. C. (1992). How robust is robust regression? *Stata Technical Bulletin*, *1*(2).

Hand, D. J., & Taylor, C. C. (1987). *Multivariate analysis of variance and repeated measures: A practical approach for behavioural scientists*. London, UK: Chapman & Hall.

Hawley, P. H. (1999). The ontogenesis of social dominance: A strategy-based evolutionary perspective. *Developmental Review*, *19*(1), 97–132.

Hellhammer, J., Fries, E., Schweisthal, O. W., Schlotz, W., Stone, A. A., & Hagemann, D. (2007). Several daily measurements are necessary to reliably assess the cortisol rise after awakening: State-

and trait components. *Psychoneuroendocrinology*, *32*(1), 80–86. https://doi.org/10.1016/j.psyneuen.2006.10.005

Henrich, J., Chudek, M., & Boyd, R. (2015). The Big Man Mechanism: How prestige fosters cooperation and creates prosocial leaders. *Phil. Trans. R. Soc. B*, *370*(1683), 20150013. https://doi.org/10.1098/rstb.2015.0013

Henrich, 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*(3), 165–196. https://doi.org/10.1016/S1090-5138(00)00071-4

John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, *66*(1), 206–219. https://doi.org/10.1037/0022-3514.66.1.206

Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. *Organizational Behavior and Human Performance*, *27*(3), 443–463. https://doi.org/10.1016/0030-5073(81)90033-7

Johnsen, T. S., & Zuk, M. (1995). Testosterone and aggression in male red jungle fowl. *Hormones and Behavior*, *29*(4), 593–598. https://doi.org/10.1006/hbeh.1995.1288

Johnson, R. T., Burk, J. A., & Kirkpatrick, L. A. (2007). Dominance and prestige as differential predictors of aggression and testosterone levels in men. *Evolution and Human Behavior*, *28*(5), 345–351. https://doi.org/10.1016/j.evolhumbehav.2007.04.003

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*(6), 999–1013. https://doi.org/10.1037/0022-3514.90.6.999

Kessler, R. C. (1977). The use of change scores as criteria in longitudinal survey research. *Quality and Quantity*, *11*(1), 43–66. https://doi.org/10.1007/BF00143986

Knight, E. L., & Mehta, P. H. (2017). Hierarchy stability moderates the effect of status on stress and performance in humans. *Proceedings of the National Academy of Sciences*, *114*(1), 78–83. https://doi.org/10.1073/pnas.1609811114

Lord, F. M. (1956). The measurement of growth. *ETS Research Bulletin Series*, *1956*(1), i-22. https://doi.org/10.1002/j.2333-8504.1956.tb00058.x

Mazur, A. (1985). A biosocial model of status in face-to-face primate groups. *Social Forces*, *64*(2), 377–402.

Mazur, A., & Booth, A. (2014). Testosterone is related to deviance in male army veterans, but relationships are not moderated by cortisol. *Biological Psychology*, *96*, 72–76. https://doi.org/10.1016/j.biopsycho.2013.11.015

Mazur, A., Welker, K. M., & Peng, B. (2015). Does the biosocial model explain the emergence of status differences in conversations among unacquainted men? *PloS One*, *10*(11), e0142941. https://doi.org/10.1371/journal.pone.0142941

McClanahan, K. J., Maner, J. K., & Cheng, J. T. (2017, August). *Two ways to stay at the top? Dominance as a short-term strategy*. Presented at the Annual Meeting of the Academy of Management, Atlanta, GA.

McNemar, Q. (1958). On growth measurement. *Educational and Psychological Measurement*, *18*(1), 47–55. https://doi.org/10.1177/001316445801800104

Mehta, P. H., & Josephs, R. A. (2006). Testosterone change after losing predicts the decision to compete again. *Hormones and Behavior*, *50*(5), 684–692. https://doi.org/10.1016/j.yhbeh.2006.07.001

Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and Behavior*, *58*(5), 898–906.

Mehta, P. H., & Prasad, S. (2015). The dual-hormone hypothesis: A brief review and future research agenda. *Current Opinion in Behavioral Sciences*, *3*, 163–168. https://doi.org/10.1016/j.cobeha.2015.04.008

1199 Mehta, P. H., Welker, K. M., Zilioli, S., & Carré, J. M. (2015). Testosterone and cortisol jointly
1200 modulate risk-taking. *Psychoneuroendocrinology*, *56*, 88–99.
1201 https://doi.org/10.1016/j.psyneuen.2015.02.023
1202 Ponzi, D., Zilioli, S., Mehta, P. H., Maslov, A., & Watson, N. V. (2016). Social network centrality and
1203 hormones: The interaction of testosterone and cortisol. *Psychoneuroendocrinology*, *68*, 6–13.
1204 https://doi.org/10.1016/j.psyneuen.2016.02.014
1205 Pruessner, J. C., Gaab, J., Hellhammer, D. H., Lintz, D., Schommer, N., & Kirschbaum, C. (1997).
1206 Increasing correlations between personality traits and cortisol stress responses obtained by
1207 data aggregation. *Psychoneuroendocrinology*, *22*(8), 615–625.
1208 https://doi.org/10.1016/S0306-4530(97)00072-3
1209 Radcliffe-Brown, A. R. (1964). *The Andaman Islanders*. New York, NY: Free Press.
1210 Redhead, D. J. (2016). Rank differentiation among adolescent hierarchies in Romanian state care.
1211 *Anthropology of East Europe Review*, *33*(2), 26–45.
1212 Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of
1213 change. *Psychological Bulletin*, *92*(3), 726–748.
1214 https://doi.org/http://dx.doi.org/10.1037/0033-2909.92.3.726
1215 Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the
1216 measurement of change. *Journal of Educational Measurement*, *20*(4), 335–343.
1217 Rose, R. M., Holaday, J. W., & Bernstein, I. S. (1971). Plasma testosterone, dominance rank and
1218 aggressive behaviour in male rhesus monkeys. *Nature*, *231*(5302), 366–368.
1219 https://doi.org/10.1038/231366a0
1220 Rousseeuw, P. J., & Leroy, A. M. (2003). *Robust regression and outlier detection*. Hoboken, NJ: John
1221 Wiley & Sons.
1222 Sahlins, M. D. (1963). Poor man, rich man, big-man, chief—political types in Melanesia and
1223 Polynesia. *Comparative Studies in Society and History*, *5*(3), 285–303.
1224 https://doi.org/10.1017/S0010417500001729
1225 Salvador, A., & Costa, R. (2009). Coping with competition: Neuroendocrine responses and cognitive
1226 variables. *Neuroscience & Biobehavioral Reviews*, *33*(2), 160–170.
1227 https://doi.org/10.1016/j.neubiorev.2008.09.005
1228 Sapolsky, R. M. (1983). Endocrine aspects of social instability in the olive baboon (Papio anubis).
1229 *American Journal of Primatology*, *5*(4), 365–379. https://doi.org/10.1002/ajp.1350050406
1230 Sapolsky, R. M. (1991). Testicular function, social rank and personality among wild baboons.
1231 *Psychoneuroendocrinology*, *16*(4), 281–293. https://doi.org/10.1016/0306-
1232 4530(91)90015-L
1233 Schultheiss, O. C., Wirth, M. M., Torges, C. M., Pang, J. S., Villacorta, M. A., & Welsh, K. M. (2005).
1234 Effects of implicit power motivation on men's and women's implicit learning and
1235 testosterone changes after social victory or defeat. *Journal of Personality and Social*
1236 *Psychology*, *88*(1), 174–188. https://doi.org/10.1037/0022-3514.88.1.174
1237 Sherman, G. D., Lerner, J. S., Josephs, R. A., Renshon, J., & Gross, J. J. (2015). The interaction of
1238 testosterone and cortisol is associated with attained status in male executives. *Journal of*
1239 *Personality and Social Psychology*. https://doi.org/http://dx.doi.org/10.1037/pspp0000063
1240 Slatcher, R. B., Mehta, P. H., & Josephs, R. A. (2011). Testosterone and self-reported dominance
1241 interact to influence human mating behavior. *Social Psychological and Personality Science*,
1242 *2*(5), 531–539. https://doi.org/10.1177/1948550611400099
1243 Tucker, L. R., Damarin, F., & Messick, S. (1966). A base-free measure of change. *Psychometrika*,
1244 *31*(4), 457–473. https://doi.org/10.1007/BF02289517
1245 van Anders, S. M., & Watson, N. V. (2006). Social neuroendocrinology: Effects of social contexts and
1246 behaviors on sex steroids in humans. *Human Nature*, *17*(2), 212–237.
1247 https://doi.org/10.1007/s12110-006-1018-7

van der Meij, L., Buunk, A. P., van de Sande, J. P., & Salvador, A. (2008). The presence of a woman increases testosterone in aggressive dominant men. *Hormones and Behavior*, *54*(5), 640–644. https://doi.org/10.1016/j.yhbeh.2008.07.001

van der Meij, L., Schaveling, J., & van Vugt, M. (2016). Basal testosterone, leadership and dominance: A field study and meta-analysis. *Psychoneuroendocrinology*, *72*, 72–79. https://doi.org/10.1016/j.psyneuen.2016.06.005

von Rueden, C., Gurven, M., & Kaplan, H. (2008). The multiple dimensions of male social status in an Amazonian society. *Evolution and Human Behavior*, *29*(6), 402–415. https://doi.org/10.1016/j.evolhumbehav.2008.05.001

von Rueden, C., Gurven, M., & Kaplan, H. (2011). Why do men seek status? Fitness payoffs to dominance and prestige. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1715), 2223.

Welling, L. L. M., Jones, B. C., DeBruine, L. M., Smith, F. G., Feinberg, D. R., Little, A. C., & Al-Dujaili, E. A. S. (2008). Men report stronger attraction to femininity in women's faces when their testosterone levels are high. *Hormones and Behavior*, *54*(5), 703–708. https://doi.org/10.1016/j.yhbeh.2008.07.012

Whitten, P. L. (2000). Evolutionary endocrinology of the cercopithecoids. In P. F. Whitehead & C. J. Jolly (Eds.), *Old world monkeys* (pp. 269–297). Cambridge, UK: Cambridge University Press.

Wingfield, J. C., Lynn, S. E., & Soma, K. K. (2001). Avoiding the "costs" of testosterone: Ecological bases of hormone-behavior interactions. *Brain, Behavior and Evolution*, *57*(5), 239–251. https://doi.org/10.1159/000047243

Zilioli, S., & Watson, N. V. (2012). The hidden dimensions of the competition effect: Basal cortisol and basal testosterone jointly predict changes in salivary testosterone after social victory in men. *Psychoneuroendocrinology*, *37*(11), 1855–1865. https://doi.org/10.1016/j.psyneuen.2012.03.022