

Response To Post, “Thirty-somethings are Shrinking and Other Challenges for U-Shaped Inferences”

Uri Simonsohn & Leif Nelson (S&N) point out that the most commonly used test for testing inverted U-shapes may be inappropriate and propose an alternative. In this response, we 1) report the results of these commonly used tests, 2) apply the test proposed by S&N, and 3) discuss the implications for the Too-Much-Talent Effect.

1) Commonly Used Tests of Inverted U-Shapes

In a recently published paper, we hypothesized that the relationship between talent and performance would vary by task interdependence. We also proposed that in team sports that require high task interdependence (e.g., Basketball and Soccer) “top talent can produce diminishing marginal returns and even decrease performance by hindering intrateam coordination.” (p. 1590). This hypothesis was grounded in research showing that groups with high concentrations of dominant, high-status individuals perform worse than those with lower concentrations because a high concentration of stars increases the likelihood of status conflicts that direct behavior away from team coordination. In contrast, we proposed that in team sports that require relatively low interdependence (e.g., Baseball), talent should have a simple positive effect, with more talent producing better performance.

We tested our hypotheses by utilizing the most commonly used statistical methods. Like practically all other researchers in economics, management, marketing, psychology, and sociology¹, we examined whether the linear and quadratic effects of talent were opposite in sign and significant. For Basketball and Soccer, we found that this was true. For Baseball, as expected, only the linear effect was significant. Second, we plotted the fitted regression lines within the data range. For both Basketball and Soccer, the fitted curve suggested a downward slope at the high level of talent concentration. Thus, the results showed an inverted U-shape within the range of observed data (or really an inverted J-shape). The results reported in the paper demonstrate that these conditions for establishing an inverted U-shape were met.

We also tested the regions of significance for the slope in models with quadratic terms (Miller et al., 2013). The slopes for talent in Basketball (Figure 1a) and Soccer (Figure 1b) were positive and significant before they turned negative and significant.

% top talent	Simple Slope	Standard Error	t-value	Lower 95% CI	Upper 95% CI
0%	1.609	0.418	3.846	0.786	2.431
5%	1.426	0.364	3.923	0.711	2.141
10%	1.244	0.309	4.020	0.635	1.852
15%	1.061	0.256	4.144	0.558	1.564
20%	0.878	0.204	4.300	0.477	1.280
25%	0.696	0.156	4.469	0.390	1.002
30%	0.513	0.114	4.487	0.288	0.738
35%	0.331	0.091	3.642	0.152	0.509
40%	0.148	0.099	1.503	-0.046	0.342
45%	-0.034	0.132	-0.258	-0.295	0.226
50%	-0.217	0.178	-1.218	-0.567	0.133
55%	-0.399	0.228	-1.748	-0.848	0.050
60%	-0.582	0.281	-2.071	-1.134	-0.029
65%	-0.764	0.335	-2.283	-1.423	-0.106

% top talent	Simple Slope	Standard Error	t-value	Lower 95% CI	Upper 95% CI
0%	5.005	0.372	13.458	4.274	5.736
5%	4.579	0.326	14.052	3.938	5.220
10%	4.153	0.281	14.786	3.601	4.705
15%	3.727	0.238	15.688	3.260	4.194
20%	3.301	0.197	16.753	2.913	3.688
25%	2.875	0.161	17.814	2.557	3.192
30%	2.448	0.135	18.200	2.184	2.713
35%	2.022	0.122	16.514	1.782	2.263
40%	1.596	0.129	12.340	1.342	1.851
45%	1.170	0.153	7.664	0.870	1.470
50%	0.744	0.186	3.993	0.378	1.110
55%	0.318	0.226	1.408	-0.126	0.762
60%	-0.108	0.268	-0.403	-0.636	0.420
65%	-0.534	0.313	-1.707	-1.150	0.081
70%	-0.960	0.359	-2.677	-1.666	-0.255
75%	-1.387	0.405	-3.420	-2.184	-0.589
80%	-1.813	0.453	-4.005	-2.702	-0.923
85%	-2.239	0.500	-4.477	-3.222	-1.256
90%	-2.665	0.548	-4.864	-3.742	-1.588
95%	-3.091	0.596	-5.187	-4.263	-1.919
100%	-3.517	0.644	-5.460	-4.783	-2.251

¹ In the past five years alone, a quick search resulted in more than thirty papers published in top psychology, marketing, and management journals using this exact same approach to test u-curves.

Consistent with our prediction that the level of task interdependence would moderate these effects, the slope in baseball was always positive (Figure 1c).

Figure 1c. Baseball analyses

% top talent	Simple Slope	Standard Error	t-value	Lower 95% CI	Upper 95% CI
0%	0.698	0.155	4.502	0.393	1.002
5%	0.655	0.134	4.886	0.392	0.919
10%	0.613	0.113	5.407	0.390	0.836
15%	0.571	0.093	6.149	0.389	0.754
20%	0.529	0.073	7.280	0.386	0.672
25%	0.487	0.053	9.162	0.383	0.592
30%	0.445	0.035	12.545	0.375	0.515
35%	0.403	0.024	16.726	0.355	0.450
40%	0.361	0.028	12.816	0.305	0.416
45%	0.319	0.044	7.320	0.233	0.404
50%	0.276	0.062	4.435	0.154	0.399
55%	0.234	0.082	2.849	0.073	0.396
60%	0.192	0.103	1.873	-0.010	0.394
65%	0.150	0.123	1.218	-0.092	0.393

2) An Application of the Test Proposed by S&N

S&N point out that these widely used tests may not accurately represent whether a negative slope is evident and they propose an alternative test that they believe more accurately captures whether a negative slope exists. We agree with S&N that their test has intuitive appeal and we are happy to re-analyze our data with this new method. Additionally, S&N acknowledge at the end of their blog that there may be other methods that can capture these tests more effectively.

When we use the new test proposed by S&N on our own data with the 40% talent cutoff for Basketball and Baseball that we reported in the supplementary online materials (SOM) of the paper (i.e., an expanded and more inclusive measure of talent), we can see that for Baseball, the inflection point falls outside the data range and only reveals a significant linear relationship between talent and team performance ($p \leq .001$, Figure 2a). In contrast, we can see for Basketball that a regression with the two lines produces significant positive slopes for team performance ($p \leq .001$, Figure 2b) and intra-team coordination ($p \leq .001$, Figure 2c), as well as significant negative slopes following the inflection points ($p = .026$ and $p = .039$, respectively). In addition, intra-team coordination mediates the effect of talent on team performance. These tests clearly confirm our hypothesis that task interdependence moderates the relationship between talent and team performance. We cannot expand the Soccer data without creating inconsistencies in the talent coding because the total number of elite clubs in the Deloitte Football Money League ranking are fixed.

When we use a more narrow measure of talent (the 33.33% cutoff, reported in the main text of the paper), we can see that the inflection point in Baseball falls again outside the data range and only reveals the hypothesized linear relationship between talent and team performance ($p \leq .001$, Figure 3a). We can also see that for Basketball and Soccer a) the regression with the two lines yields significant positive slopes in Basketball performance ($p \leq .001$, Figure 3b), Basketball coordination ($p \leq .001$, Figure 3c), and Soccer ($p \leq .001$, Figure 3d), b) there is a critical inflection point that falls inside the data range, but c) there are no significant slopes following the inflection points ($p = .48$, $p = .86$, and $p = .53$, respectively). These analyses continue to confirm that concentrations of higher talent produce diminishing returns and that this effect is moderated by task interdependence.

40% TALENT MEASURE
(reported in the SOM of the paper)

33.3% TALENT MEASURE
(reported in the main text of the paper)

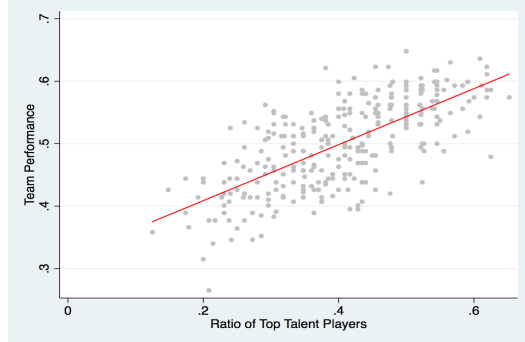


Figure 2a. MLB performance – top talent (40%). S&N test only reveals a linear relationship ($p \leq .001$)

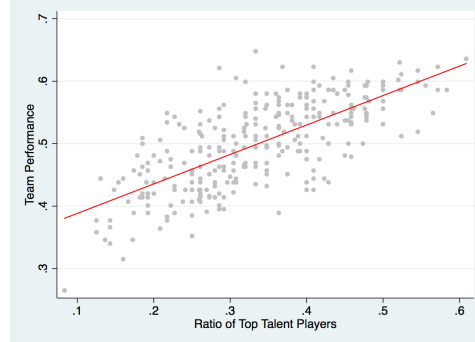


Figure 3a. MLB team performance – top talent (33%). S&N test only reveals a linear relationship ($p \leq .001$)

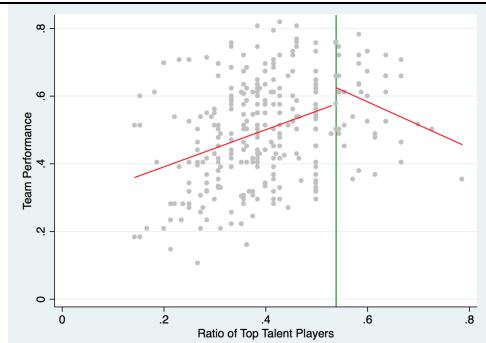


Figure 2b. NBA performance – top talent (40%). S&N test reveals that the first slope is significant and positive ($p \leq .001$) and that the second slope is significant and negative ($p = .026$).

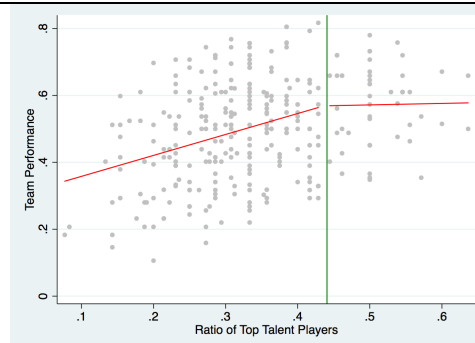


Figure 3b. NBA performance – top talent (33%). S&N test reveals that the first slope is significant and positive ($p \leq .001$) and that the second slope is not significant ($p = .48$).

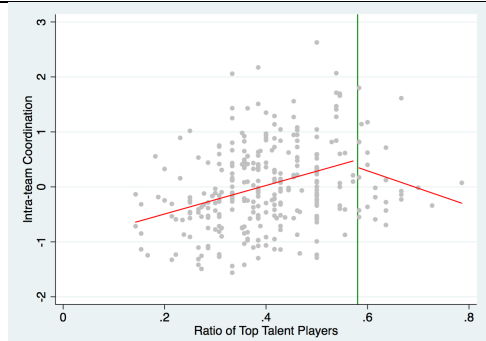


Figure 2c. NBA coordination – top talent (40%). S&N test reveals that the first slope is significant and positive ($p \leq .001$) and that the second slope is significant and negative ($p = .039$).

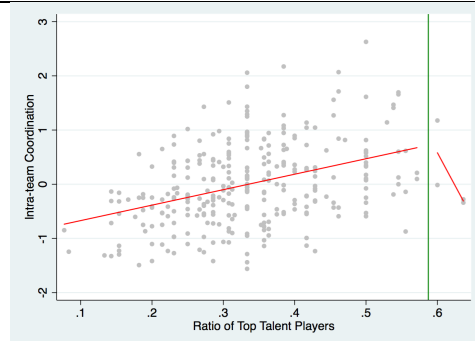


Figure 3c. NBA coordination – top talent (33%). S&N test reveals that the first slope is significant and positive ($p \leq .001$) and that the second slope is not significant ($p = .86$).

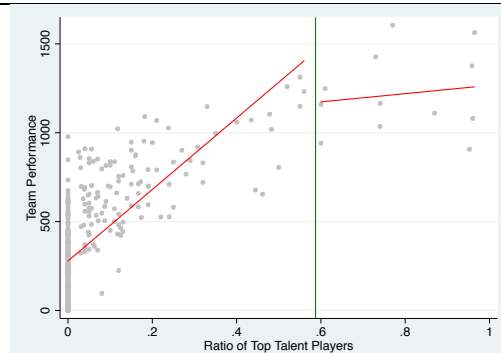


Figure 3d. Soccer performance – top talent (Top 20+ Clubs). S&N test reveals that the first slope is significant and positive ($p \leq .001$) and that the second slope is not significant ($p = .53$).

3) Implications of the S&N Test for the Too-Much-Talent Effect

So what does the new test proposed by S&N imply for our original predictions that a) task interdependence alters the relationship between talent and performance and b) when interdependence is high, top talent can produce diminishing marginal returns, and even decrease performance by hindering intra-team coordination?

First, the results of the new test proposed by S&N show that the relationship between talent and performance varies by whether the team task requires high or low levels of interdependence. The test continues to provide support for our prediction that increases in top talent no longer produce positive effects on performance at some inflection point for tasks that require high levels of interdependence (i.e., Basketball and Soccer). For tasks that require lower levels of interdependence (i.e., Baseball), the inflection point always fell outside the data range such that increases in top talent continue to produce positive effects. Regardless of the test used, this hypothesis is always supported.

The finding that the effect of top talent becomes flat (null) at some point is an important finding: Even under the assumption of diminishing marginal returns, the cost-benefit ratio of adding more talent can decline as hiring top talent is often more expensive than hiring average talent. In Hollywood, high concentrations of stars often have a negative effect on profitability; more stars may increase revenue but this benefit is completely dominated by increases in production costs (De Vany & Walls, 2004). Similarly, football clubs like Chelsea, Real Madrid, and Paris Saint-Germain continue to hire top-talented players like Fernando Torres, Kaka, and David Luiz for exorbitant amounts, while the increased costs of hiring such players more than outweigh the benefits they add during the game.

Second, the results of the new test proposed by S&N suggest that the strongest version of our arguments – that more talent can even lead to worse performance – may not be as robust as we initially thought, though further analyses show that this also depends on the cutoff point used to define ‘top talent’. Notably, our results are robust to this particular test when we use the slightly broader measure of talent discussed in the SOM of our paper. Here, the test proposed by S&N still shows that talent has a negative and significant effect on team performance and intra-team coordination beyond the inflection point in Basketball. The Basketball and Baseball analyses are particularly good tests because they cover the same years, the same country, the same talent measure, and the same performance measure.

Finally, the scatter plots suggest that in interdependent sports, some teams benefit from more talent whereas others are hurt by it. Although this could be random variance, it may also reflect critical moderators that we are currently exploring.

Regardless of the test used, we can conclude that interdependence significantly alters the relationship between top talent and team performance. When players are highly interdependent, top talent can produce diminishing marginal returns and sometimes – but perhaps not always – decrease team performance by hindering intra-team coordination.

We thank Uri Simonsohn and Leif Nelson for proposing this alternative test and for beginning a discussion about how to accurately capture whether a slope changes in sign and significance following the inflection point; this conversation will clearly benefit the entire social sciences. We also thank them for their collaborative approach and for sharing an early draft of their blog post with us.

Roderick Swaab
INSEAD

Michael Schaerer
INSEAD

Eric Anicich
Columbia
University

Richard Ronay
Free University
Amsterdam

Adam Galinsky
Columbia
University

References

De Vany, A. S., & Walls, W. D. (2004). Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar. *Journal of Economic Dynamics and Control*, 28, 1035-1057.

Miller, J.W., Stromeyer, W.R., & Schwieterman, A. (2013). Extensions of the Johnson-Neyman technique to linear models with curvilinear effects: Derivations and analytical tools. *Multivariate Behavioral Research*, 48, 267-300