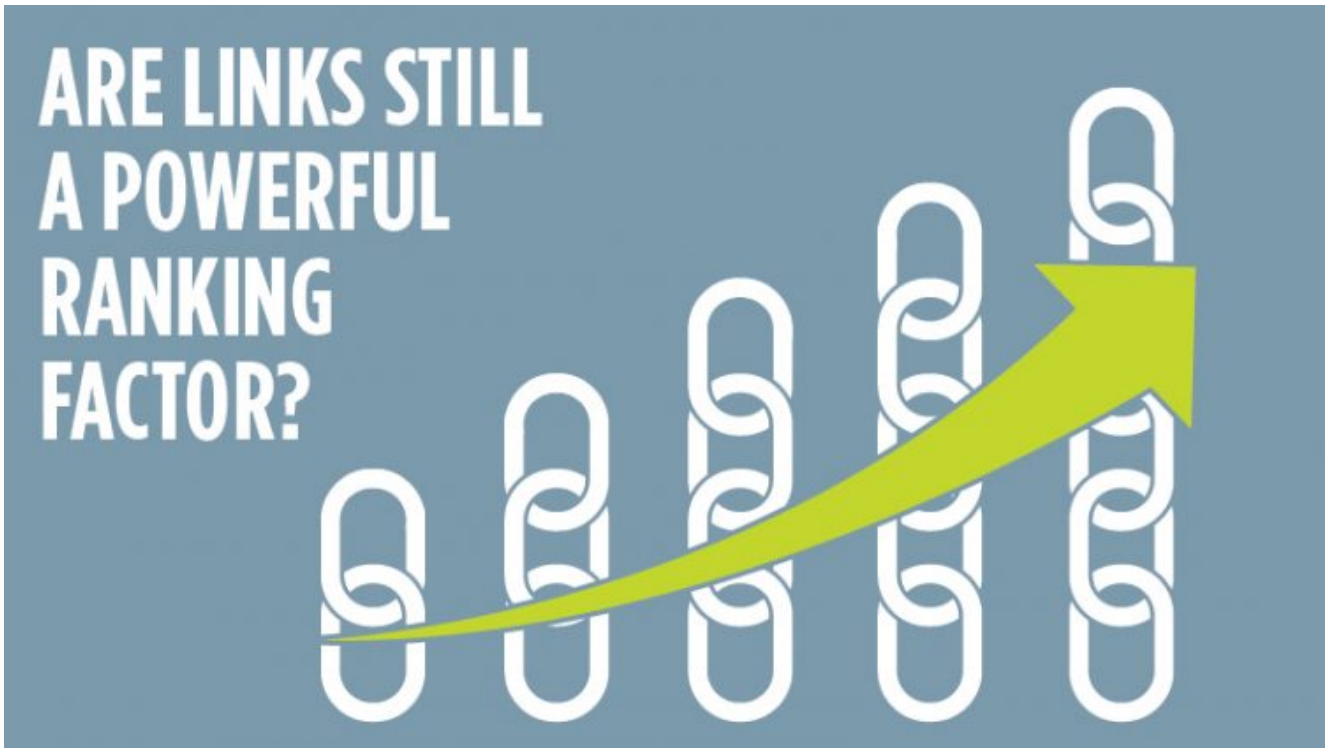
2016 SEO Strategy

Link Building
and
High Quality Content

Study Shows Link Building Still Highly Effective When Used With Great Content



You've seen the claims:

1. "So Many New Ranking Factors!"
2. "Content Quality is Everything!"
3. "User Engagement is the New King!!"
4. "Social Media Signals Rock!"
5. "RankBrain has Taken Over All Ranking!"

Puhleeze, spare us the nonsense! We want to share new data on links as a ranking factor. This report will demonstrate the continuing importance that backlinks play in rankings. We plan to share data that shows that all that talk of a decline in the importance of links as a ranking factor is grossly exaggerated.

In fact, links remain amazingly powerful. Let's dig in!

Study Shows Links Still Rule

On Mar 23, 2016, I participated in a Hangout with Google Dublin's Andrey Lippatsev (along with Rand Fishkin and Ammon Johns). Here is an important snippet of the conversation:

> Ammon Johns: We heard that RankBrain is the third-most-important signal contributing to results now. Would it be beneficial to us to know what the first two are?
>
> Andrey Lippatsev: Yes. Absolutely. I can tell you what they are. It's **content and links going into your site**.

Do tell now?

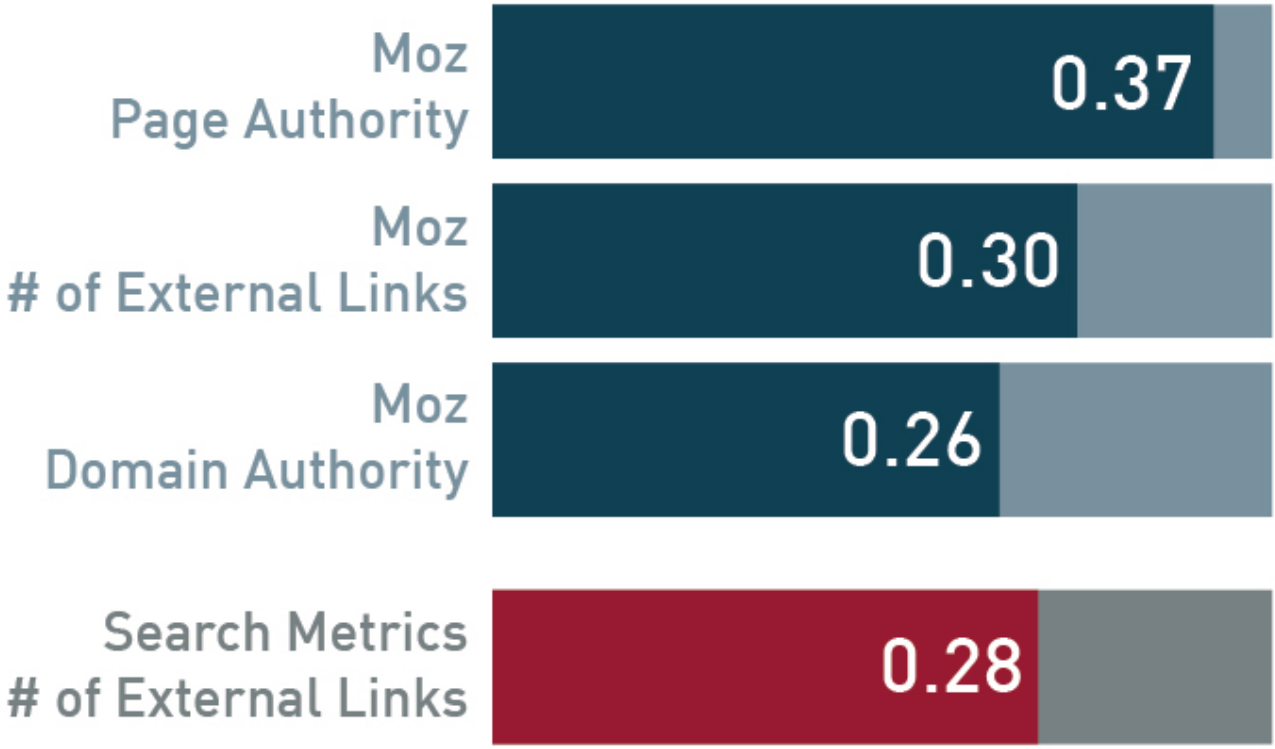
What the Moz and Searchmetrics Studies Showed

Both Moz and Searchmetrics have run groundbreaking studies on ranking factors, and each includes a look at links. You can see them here:

1. Moz Correlation Factors
2. Searchmetrics Ranking Factors

You can see some of the basic data points in the following chart:

Study Shows Links Still Rule



Each of the bars above shows the level of correlation between that factor and higher rankings. In both the Moz and Searchmetrics studies, correlation to the number of links was relatively high, but not significantly higher than other factors they examined. If links are so important that a Googler would call them one of the two most important ranking factors, why aren't these correlations showing higher numbers?

The key to the Moz and Searchmetrics studies is understanding how the evaluations were done. They did the evaluation of each SERP on an individual basis, and then took the mean of all the results (I'll call this the "Mean of the Individual Correlations" approach). Both these studies also focused solely on commercial search terms.

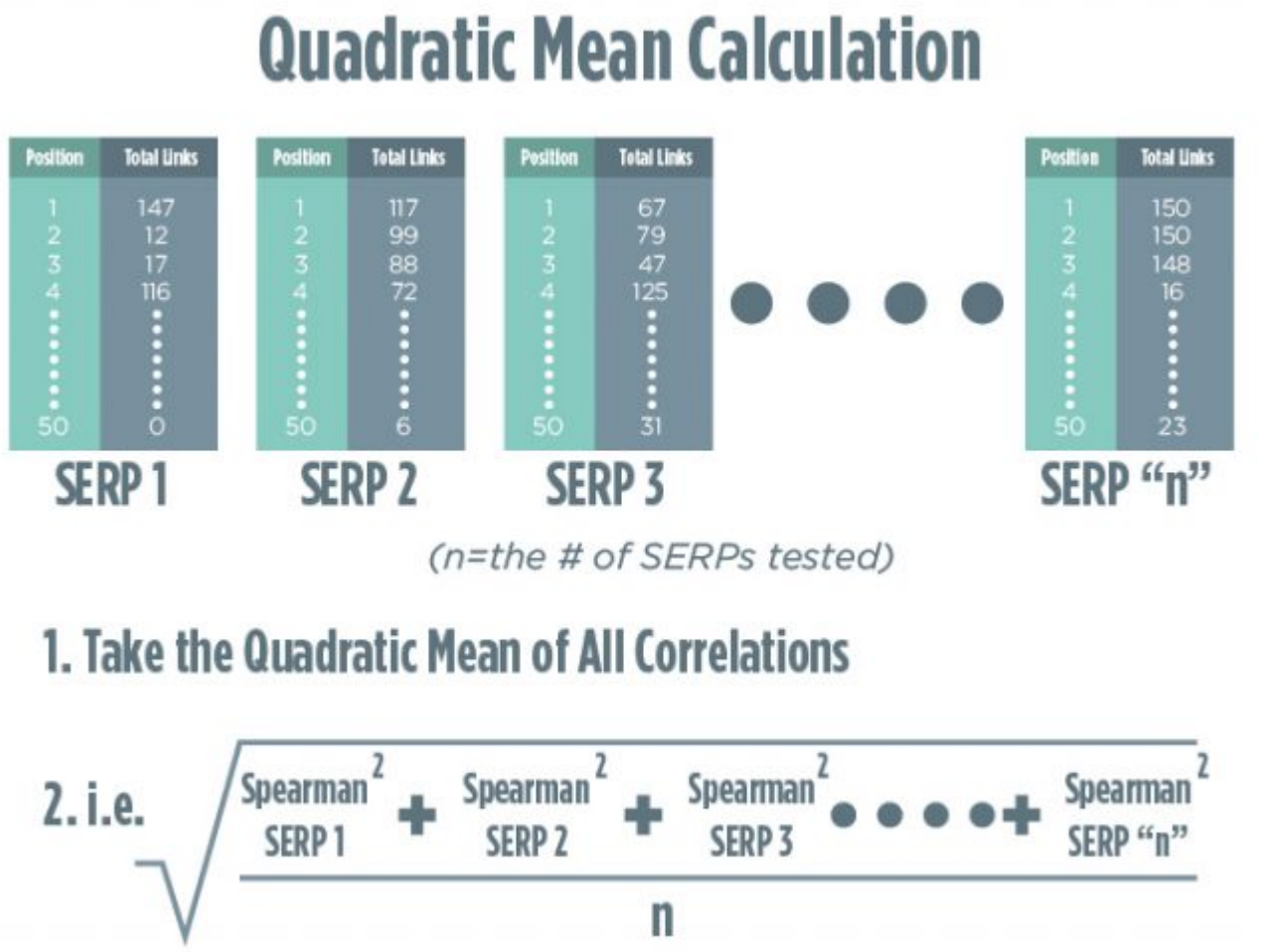
So while their approaches have validity, because of our experience with the power of links in ranking, we decided to take a deeper look into the power of links. We took some new approaches to expose more about their impact. Bottom line? I believe that the Mean of Individual Correlations approach is not providing a complete picture.

Our Study Results

Based on consultations with a couple of experts, I did a different type of calculation, based on the Quadratic Mean. The reason for doing this is that it leverages the square of the correlation variables (where the correlation value is R, the quadratic mean uses R squared).

It's actually the R squared value that has some meaning in statistics. For example, if R is 0.8, then R squared is 0.64, and you can say about that 64% of the variability in Y is explained by X. As our consultants explained it to me, there is no meaningful sentence involving the correlation variable R, but R squared gives you something meaningful to say about the correlated relationship.

Here is a visual on what this calculation process looks like:



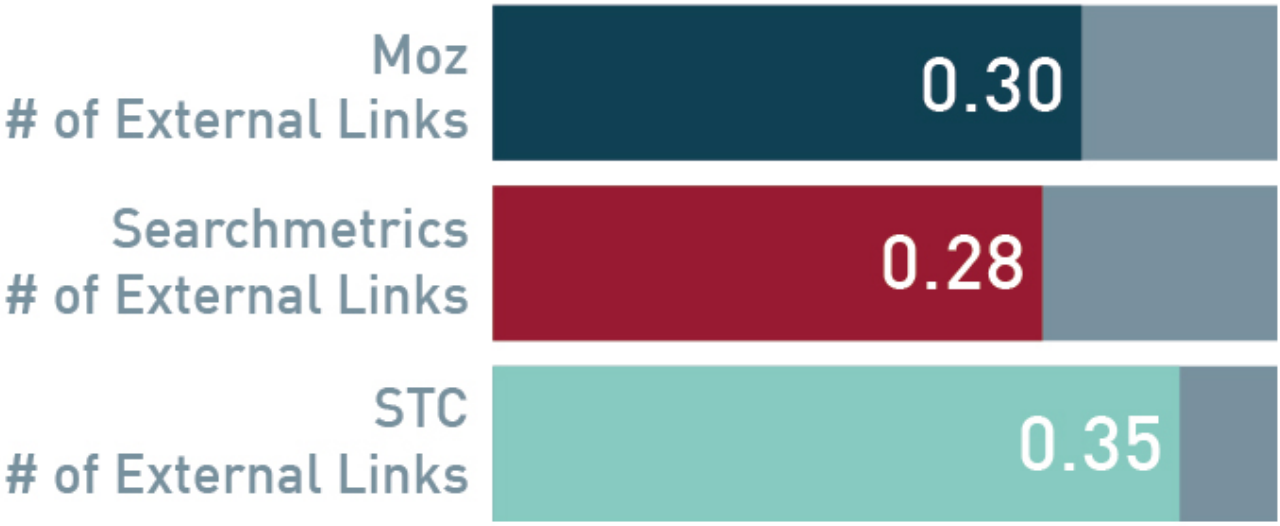
In addition to the different calculation approach, I also used a mix of different query types. We tested commercial head terms, commercial long tail terms, and also informational queries. In fact, 2/3 of our queries were informational in nature. This may be one reason why we showed somewhat different results:

Study Shows Links Still Rule

Our total links correlation result was higher than the correlation scores of DA and PA as ranking factors. We asked Rand Fishkin of Moz for a comment on why that may be, and here is what he had to say:

> We use a different, broader corpus of keywords to generate PA/DA algorithms, and thus, it makes sense that on different types of keyword queries, they'll have different levels of correlation. Very interesting to note that raw link counts tend to do better on these particular corpuses. My hope is actually that, one day in the future, Moz will be able to show how various metrics of all kinds correlate to any given set of keywords (wouldn't it be awesome if you could see the correlations for the thousand keywords you're rank tracking, for example?!).

Also of note is how high our total link score correlation was in comparison to Moz and Searchmetrics. The score I showed above used a different methodology, but even when we use the exact same methodology they did (the Mean of Individual Correlations approach), we get higher results. You can see the head to head comparison of all our link score calculations using that approach here:



Taking a Moment to Give Thanks: all of our data was pulled from the Moz API, and the good folks on the Moz team were quite supportive of us in all aspects of the entire study.

Aggregate Evaluation of Links as a Ranking Factor

I think that the Mean of the Individual Correlations and Quadratic Mean approaches are both highly valid ones, but one of the limits with them is that a small number of results with a highly negative correlation can drag down the total score in a significant way.

For that reason, I chose to take some other approaches to the analysis as well. The first of those was to measure the links in a more aggregated manner. To do this, we normalized the quantity of links. Then we took the total of all the search results by ranking position. The equations for this look more like this:

Total Links By SERP Position Calculation

1. Normalize total links for each SERP so highest "link total"=1

2. Sum Total links by position

$$\text{Pos 1 Total} = \sum_{i=1}^{n} \text{links to Position 1}$$

$$\text{Pos 2 Total} = \sum_{i=1}^{n} \text{links to Position 2}$$

$$\text{Pos 50 Total} = \sum_{i=1}^{n} \text{links to Position 50}$$

Position	Total Links	Normalized Links
1	147	147/147
2	12	12/147
3	17	17/147
4	116	116/147
⋮	⋮	⋮
50	0	0

i=SERP #

3. Calculate Spearman & Pearson Correlations on the Resulting Array

The value of this is that it smooths out the impact of the negative correlations in a different way. When you look at the correlations in this manner, here is what you get:

I also took another look at this. In this view, I continued to use the normalized totals of the links, but grouped them in ranking groups of 10. I.e., I summed the normalized link totals for the top 10, did the same for ranking positions 11 to 20, 21 to 30, and so forth. I then calculated the correlations to see how they looked in terms of what it would take to rank in each 10 position block. Those calculations looked more like this:

Total Links by Block of 10 SERPs Calculation

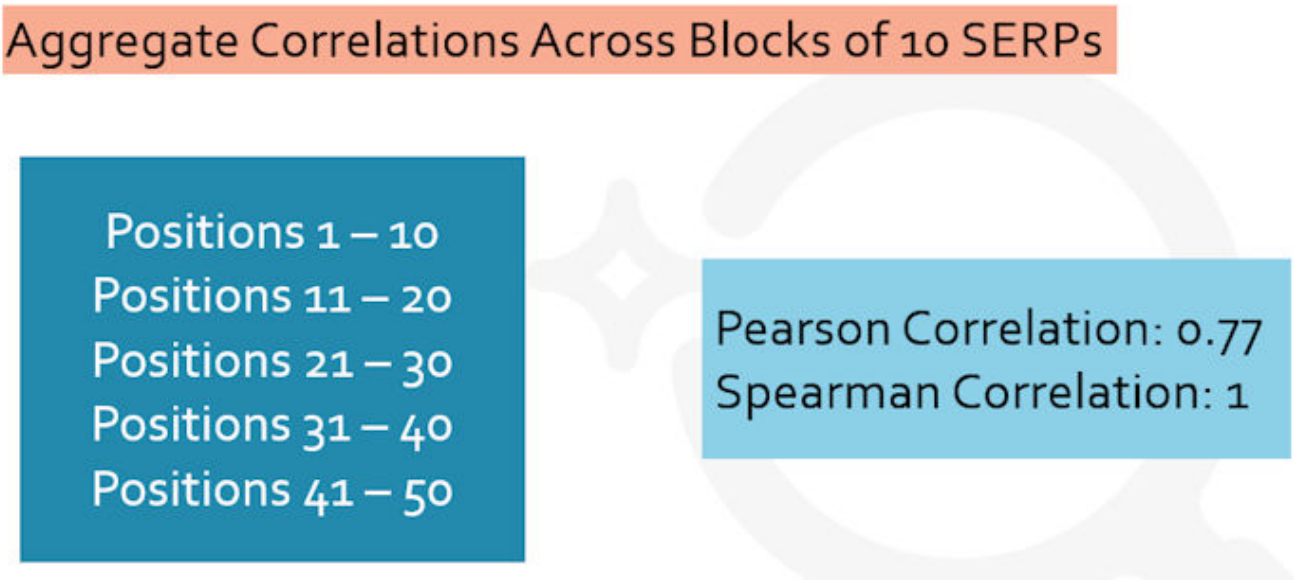
Similar to Model 2- We Still Normalize the Data

1. $\sum_{i=1}^{10}$ All the Links , $\sum_{i=11}^{20}$ All the Links , • • • • $\sum_{i=41}^{50}$ All the Links

Where i=the ranking position

2. Take the Pearson/Spearman for these 5 data points

This gives us a bit more granular approach than simply aggregating all the ranking positions into the SERP positions, but still smooths out some of the limitations of the Mean of Individual Correlations method. Here is what we saw for results here:



I've now gone from numbers in the 0.39 range to near-perfect correlations, so what do these numbers tell us? I believe that the aggregated approaches to the calculations tell us that links are far more important than the "mean" based calculations reveal.

Digging Into What's Going On Here

To help with this analysis, we did some manual analysis of a few hundred results to figure out what percentage or results were likely **not as influenced by links**. What types of results are these? Here you go:

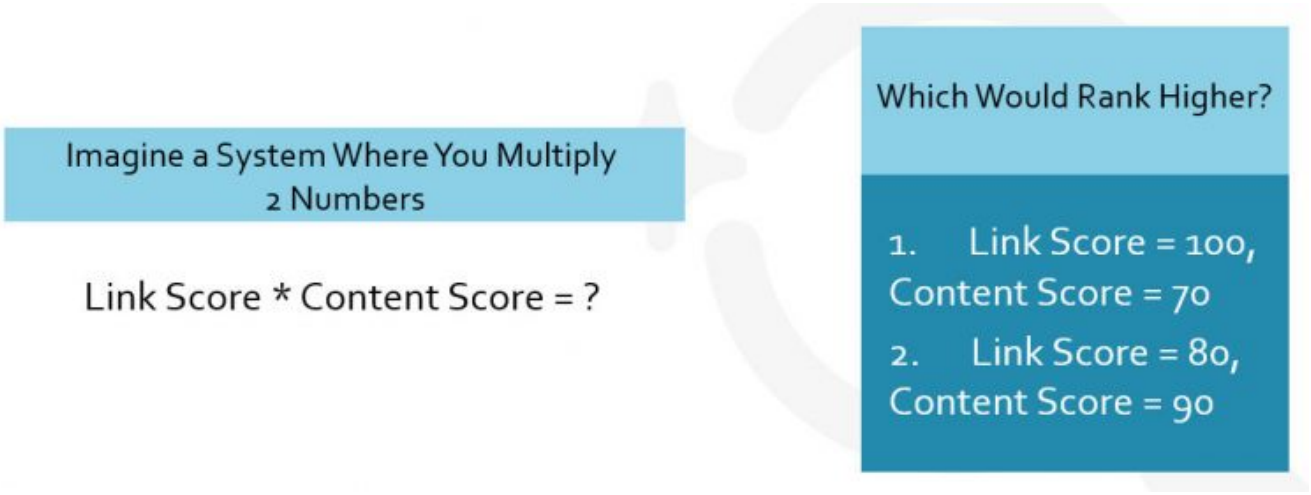
1. Local results (not maps results, but results that are locally influenced)
2. Query deserves diversity
3. In Depth Articles

Our analysis suggested that around 6% of the results came from these types of results. So that's one factor right there, albeit not one that explains the difference between the "mean" based calculations and the aggregate calculations.

However, the more important issue is understanding the role of content quality. Since Google has called content and links the two most important ranking factors, you can imagine a greatly simplified equation based on simply multiplying a "link score" times the "content score".

Now in this completely hypothetical discussion, I would argue that the content score probably counts for more than the link score. After all, if the content isn't relevant, it shouldn't rank. That just makes sense. In addition, the level of relevance of a piece of content is highly variable.

Here is a major oversimplification of what the impact of this might be:



In the above chart, it's actually scenario 2 that would rank higher, even though its "Link Score" is significantly lower. Now imagine that the link score runs only from 1 to 100, and the content related scoring factors (relevance and quality) have an exponential decay (were minor variations in quality and relevance have a large impact on the "Content Score"), links will clearly not be able to overcome lower relevance or weaker content.

Important Note: the above discussion was meant to illustrate the basic point about the significance of content scoring and its impact on ranking algorithms. It's all my own speculation based on the analysis of the data I've seen, and the case studies that I'll share in the next section).

Cementing the Point With Case Studies

We do a lot of high end content marketing campaigns with our clients, many of which are Fortune 500 companies. Here is a sampling of the results across many of our clients:

Keyword	Search Volume	Original Rank	Current Rank	# of New Links
keyword 1	135,000	7	1	11
keyword 2	165,000	18	1	7
keyword 3	49,500	6	1	3
keyword 4	450,000	20	1	25
keyword 5	201,000	15	3	12
keyword 6	159,500	3	2	15
keyword 7	27,100	6	1	5
keyword 8	110,000	5	1	13
keyword 9	68,600	7	1	4
keyword 10	165,000	12	2	4

The sample results shown here have been repeated hundreds of times by us. However, we don't find that links can rescue poor quality content, or cause low relevance content to rank. Also, all of our efforts focus on the getting recognition from, or content published on, very high authority sites. Our data isn't based on high volume, low quality link building.

Study Shows Links Still Rule

Summary

The Google algorithm continues to evolve, and we see many things that are impacting overall organic search traffic. Some of the biggest causes are:

1. More real estate allocated to paid search
2. More content from other sources, such as image search, YouTube, and the other factors I mentioned above
3. Some pages that have less than 10 web results
4. Portions of the web results that are clearly less driven by links, such as local web, query deserves diversity, and in-depth article results

As a result, there are fewer than ten results on the first page of the SERPs which are driven by factors other than links. That doesn't mean that links aren't involved at all in rankings for those pages just that they matter less.

But, our study data strongly suggests that links continue to play a major role in rankings. In addition, our case study data makes the whole thing a slam dunk. *When you aren't facing page relevance or quality issues, links can, and do, continue to significantly impact rankings*.

Note: even though our data focused on showing the correlation based on the quantity of links, that doesn't mean that the quality of the links doesn't matter, as it truly does.

So what's the real upshot of all this? Simple:

1. DO build great content and user experiences. In today's digital marketing world, that's just table stakes.
2. DO proactively market your business and do the types of things that cause people to write about you and link to you.

You MUST do BOTH

.