

What is the relationship between impact and EA Forum karma?

Questions

The current karma system of EA Forum, which is described [here](#)¹, is based on the assumption that karma is positively correlated with impact (of the post/comment). I think this is pretty reasonable, but how strong is the correlation?

I am also curious about the shape of the relationship between karma and impact. For example, if post Y has twice as much karma as post X, is Y roughly 2 times as valuable as X ([linear](#) function), less than that (e.g. [logarithmic](#)), or more than that (e.g. [quadratic](#))?

Moreover, what is the impact caused by a post with a given amount of karma? Ideally, one would want an answer in terms of the effect on the expected value of the future, but I understand this is hardly feasible. So, in practice, using heuristics may be better. Examples include donations to the [Long-Term Future Fund](#) (or other), and the metric quality adjusted research papers (QARPs) used by Nuño Sempere [here](#).

Example answers

I tried to get a very preliminary sense of the answers to the questions above based on Nuño Sempere's [analysis](#) of the first 10 winners of the [EA Forum Prize](#). I estimated the mean impact in QARPs of each of the posts from the mean between the lower and upper limit of the 80 % [confidence intervals](#) provided by Nuño. The calculations are in tab "Posts" of [this](#) Sheet.

The table below contains the slope and [correlation coefficient](#) for the linear regression, with intercept fixed to 0, between the mean impact and various functions of the karma. It also has the [p-value](#) for the null hypothesis that there is no correlation. The calculations are in tab "Statistics".

Null intercept linear regression between mean impact (QARP) and...	Slope	Correlation coefficient (R)	P-value
Logarithm of the karma	$2.00 \cdot 10^{-2}$	0.468	0.243
Square root of the	$9.43 \cdot 10^{-3}$	0.466	0.245

¹ For interesting discussions of the system, see [this](#) post from Arepo, and [this](#) one from Nathan Young.

karma			
Karma	9.54×10^{-4}	0.463	0.248
Square of the karma	8.80×10^{-6}	0.453	0.249

For the logarithm, square root and square of the karma, and karma alone, the correlation coefficient is 0.5, and the p-value 0.2. The differences are quite small given the sample size of 8^2 . Interestingly, the above estimates for the correlation coefficient are similar to the 0.47 obtained by Nathan Young [here](#) for the relationship between the inflation-adjusted karma and ranking of the posts of the [Decade Review](#).

The table below shows the correspondence between the scale described [here](#) by Nuño and the karma predicted by dividing the mean impact by the slope of 0.954 $\frac{\text{mQARP}}{\text{karma}}$ obtained for the case where the mean impact is directly proportional to karma (2nd to last row in the table above). The calculations are in tab “Predictions”.

Mean impact (QARP)	Description	Example	Predicted karma
0.1 m	“A thoughtful comment”	A thoughtful comment about the details of setting up a charity	0.105
1 m	“A good blog post, a particularly good comment”	What considerations influence whether I have more influence over short or long timelines?	1.05
10 m	“An excellent blog post”	Humans Who Are Not Concentrating Are Not General Intelligences	10.5
100 m	“A fairly valuable paper”	Categorizing Variants of Goodhart's Law	105
1	“A particularly valuable paper”	The Vulnerable World Hypothesis	1.05 k
10 to 100	“A research agenda”	The Global Priorities Institute's Research Agenda	10.5 k to 105 k
100 to > 1000	“A foundational popular book on a valuable topic”	Superintelligence, Thinking Fast and Slow	105 k to > 1.05 M

² Nuño evaluated 10 posts, but only scored 8 in terms of QARPs.

> 1000	"A foundational research work"	Shannon's "A Mathematical Theory of Communication"	> 1.05 M
--------	--------------------------------	--	----------

I believe my estimate for the slope of 1 mQARP/karma overestimates the value of karma, because:

- I suspect the slope for a random sample of posts is lower than for the first 10 winners of the [EA Forum Prize](#).
- It looks like there has been karma inflation, and the 10 aforementioned posts were published 4 years ago.

I guess the value of 1 QARP is of the order of magnitude of 10 kkarma, i.e. 10 times as valuable as above, but with huge variation. In any case, one should certainly be mindful of [Goodhart's Law](#), and do not start optimising posts just for karma!