University of BRISTOL

## University of Bristol – Bristol Research Portal

### General rights

We consider a hypothetical rational researcher aiming to maximise their chances of career progression. We assume that the researcher optimises the payoff that results from dividing a fixed research budget (e.g. time or money) into (i) exploratory studies that may identify novel phenomena, and (ii) confirmatory studies that attempt to replicate previous findings, and deciding how much to invest in each study. Specifically, the researcher makes three decisions: 1) the proportion of total sampling to spend on exploratory studies (denoted by $\theta$); 2) the sample size of each exploratory study ($S_E$); and 3) the sample size of each confirmatory study ($S_C$).

For simplicity we assume that there is a typical effect size for the phenomena investigated within a particular field, but we allow this average to differ between exploratory and confirmatory studies. The sample size gives an estimate of the statistical power of each study, which in turn controls the probabilities of Type I and Type II errors and so the probability of observing an effect (either a true positive or a false positive), and of acceptance by a journal. The high statistical power of confirmatory studies means that the results are informative regardless of statistical significance so we assume that all are published with fixed probability $\psi$, but that exploratory studies are only published if they get a statistically significant result.

Statistical power is a function of the probability density function of the Gaussian distribution, and so is difficult to work with, but an acceptable approximation of the dependence of the statistical power $W_i$ of study type $i$ ($i \in \{E, C\}$) with sample size $S_i$ is given by:

$$W_i = \frac{1}{2} - \frac{1}{2} \frac{H_i - r_i^2}{\sqrt{(H_i - r_i^2)^2 + \delta}} \sqrt{1 - e^{-\frac{2S_i}{\pi\sigma^2}\left(H_i - r_i^2\right)}} \tag{S1}$$

where

$$H_i = \frac{-\pi\sigma^2}{2S_i} \log\left(4\alpha\{1-\alpha\}\right).$$

In equation (A1) $\sigma^2$ is the population variance, $r_i$ is the effect size for exploratory or confirmatory studies, $\alpha$ is the Type I error rate (i.e., the critical $P$ value within a null hypothesis significance testing framework), and $\delta$ is a small number ($10^{-12}$) used in the approximation of sign($H - r^2$).

We assume that the likelihood of acceptance by a journal editor is influenced by sample size *per se*, before consideration of the effects themselves. A reasonable function that captures the idea that the probability of acceptance on the basis of sample size alone is small when $S_i$ is very small and tends to unity otherwise as $S_i$ becomes large is:

$$A = 1 - \frac{m}{S_i} \tag{S2}$$

where $m$ is a positive constant. As $m$ increases, journal editors demand larger sample sizes before accepting studies for publication (S1 Fig). This equation tends to minus infinity at small $S_i$, but this region is never optimal so this does not affect predictions.

The number of publications from exploratory studies ($N_E$) depends on two probabilities: 1) the probability of correctly identifying a novel phenomenon (i.e., a true positive), which is the product of the statistical power and the probability that an exploratory study is indeed investigating an effect that is real ($f_E$); and 2) the probability of finding a statistically significant result when the phenomenon does not exist (i.e., a false positive), which is the product of the Type I error rate ($\alpha$) and the probability that an exploratory study is in fact investigating an effect that is *not* real ($1-f_E$). We assume that all non-significant results from exploratory studies are not published (i.e., the file drawer phenomenon). $N_E$ is the product of the number of exploratory studies conducted, the probability of acceptance given the sample size, and the probability of getting a statistically significant result:

$$N_E = \frac{\theta T}{k + 2S_E} A\big[W_E f_E + \alpha(1 - f_E)\big],$$ (S3)

where $T$ is the total number of samples that can be collected (i.e., total resources) and $k$ is the set-up cost for any study. The first term in the squared brackets is the probability of a true positive result, while the second is the probability of a false positive result.

The proportion of published exploratory studies that are false is:

$$P_{F,E} = \frac{\alpha(1 - f_E)}{W_E f_E + \alpha(1 - f_E)}.$$ (S4)

Since confirmatory studies will build on the findings of exploratory studies, the probability that a confirmatory study is looking at a real effect ($f_C$) is equal to the probability that a published exploratory study is correct ($f_C = 1 - P_{F,E}$). The number of confirmatory studies that are published ($N_C$) is the sampling effort put into all confirmatory studies divided by the sampling effort of each study multiplied by the probability they are accepted:

$$N_C = \frac{(1-\theta)T}{k + 2S_C}\big[\big(1 - P_{F,E}\big)W_C + P_{F,E}\alpha + \psi P_{F,E}(1 - \alpha) + \psi\big(1 - P_{F,E}\big)\big(1 - W_C\big)\big],$$ (S5)

The terms in the squared brackets are respectively, the probability of true positive, false positive, true negative, false negative.

The probability of false negatives will therefore be $(1 - W_E)f_E$ for exploratory studies and $(1 - W_C)f_C$ for confirmatory studies, so the total probability any study commits a false negative is:

$$P_{FN} = \frac{(1-\theta)T}{k + 2S_C}(1 - W_C)f_C + \frac{\theta T}{k + 2S_E}(1 - W_E)f_E.$$ (S6)

The proportion of published studies that are false (false positive or false negative) is:

2

$$P_F = \frac{\dfrac{\theta T}{k+2S_E}A\left[\alpha(1-f_E)\right] + \dfrac{(1-\theta)T}{k+2S_C}\left[P_{F,E}\alpha + \psi\left(1-P_{F,E}\right)\left(1-W_C\right)\right]}{N_E + N_C} \ . \quad \text{(S7)}$$

The total fitness of the researcher is assumed to depend on total number of publications (with diminishing returns) with a weighting for exploratory studies. We assume that the number of confirmatory studies per exploratory study is limited to ρ (ρ=10), by calculating the number of *valuable* confirmatory studies as:

$$\hat{N}_C = \frac{N_C}{1 + e^{\frac{N_C}{N_E+\varepsilon}-\rho}} \quad \text{(S8)}$$

where $\varepsilon$ is a small number ($10^{-6}$) that controls the steepness of the slope. Equation (A8) implies that the value of confirmatory studies is around unity when $N_C < N_E$ and around zero when $N_C > N_E$.

The value of an individual researcher is given by:

$$V_R = \gamma N_E + 1 - e^{-\phi\left(N_E + \hat{N}_C\right)}, \quad \text{(S9)}$$

In equation A9 $\phi$ controls how quickly the value of the total number of publications diminishes, and $\gamma$ controls the extra weighting given to published exploratory studies. The dependence of equation (A9) on the number of published exploratory and confirmatory studies is shown in S2 Fig for representative values of $\gamma$ and $\phi$.

We simplify our analysis by assuming a fixed value for $S_C$ and finding a strategy in only two-dimensional $(\theta, S_E)$ space. We assume $S_C = 120$ so that the statistical power of confirmatory studies is $W_C \approx 0.80$ (and where $A$ is almost unity) at the values of other parameters.

We are especially interested in how the optimal strategy changes with the weight given to novel findings $\gamma$ and how quickly the value of the total number of publications diminishes $\phi$. We can find the optimal strategy $(\theta^*, S_E^*)$: the combination of proportion of effort into exploratory studies and the sample size of exploratory studies that maximises $V_T$. Matlab code is provided in S2 Text. Baseline parameter values are given in Table S1.