

Evaluating Claims that Registration Restricts Exploratory Research

Moin Syed

2025-11-29

Power Analysis

We can use O'Mahony (2023) as a starting point for the power analysis. No effect size is reported, but we can calculate it using two methods. The first method is based on the reported test statistic:

```
##  
##      Chi squared power calculation  
##  
##          w = 0.1781715  
##          N = 303.5016  
##          df = 2  
##      sig.level = 0.05  
##          power = 0.8  
##  
## NOTE: N is the number of observations
```

This yields a value of $w = 0.178$, which would require a sample of 304. An alternative approach is to use the reported frequencies to reconstruct the 2x2 contingency table, and use that as the input:

```
##      No  Yes  
## SR 0.29 0.38  
## R   0.08 0.25  
  
##  
##      Chi squared power calculation  
##  
##          w = 0.1854456  
##          N = 280.159  
##          df = 2  
##      sig.level = 0.05  
##          power = 0.8  
##  
## NOTE: N is the number of observations
```

This yields a value of $w = 0.185$, which is similar but just a little bit larger. In this case, a sample size of 280 would be required.

O'Mahony's (2023) study included twice as many traditional reports as Registered Reports. Next we examine what happens if we make the two equal, but preserve the same relative frequencies:

```
##      No  Yes
## SR  0.21 0.29
## R   0.12 0.38

##
##      Chi squared power calculation
##
##          w = 0.1914027
##          N = 262.9913
##          df = 2
##      sig.level = 0.05
##      power = 0.8
##
## NOTE: N is the number of observations
```

Doing this results in an effect size of $w = 0.191$ and a necessary sample size of 263.

All of the preceding are 2x2 contingency tables, and thus would assume that we would collapse the two categories of registration (Registered Reports and preregistered studies) and the two categories of traditional articles (companion articles and 2010 articles). If, instead, we separated the four articles types for analysis, there are a variety of relevant scenarios worth probing with respect to their impact on the required sample size.

In Scenario 1, we assume equal frequencies within the two categorories (registered and not), at the same rate as found in O'Mahony (2023):

```
##      SR   SO   RR   PR
## No  0.11 0.11 0.06 0.06
## Yes 0.14 0.14 0.19 0.19

##
##      Chi squared power calculation
##
##          w = 0.2111002
##          N = 216.2024
##          df = 2
##      sig.level = 0.05
##      power = 0.8
##
## NOTE: N is the number of observations
```

This indicates an effect size of $w = 0.211$ and sample size of 216.

Scenario 2 examines what would happen in the rate of exploratory research in preregistered studies was the same as for standard reports:

```

##      SR   SO   RR   PR
## No  0.11 0.11 0.06 0.11
## Yes 0.14 0.14 0.19 0.14

##
##      Chi squared power calculation
##
##      w = 0.1775552
##      N = 305.6123
##      df = 2
##      sig.level = 0.05
##      power = 0.8
##
## NOTE: N is the number of observations

```

This indicates an effect size of $w = 0.178$ and sample size of 306.

Scenario 3 examines if the rate of exploratory research is higher in the matched sample of traditional articles compared with the 2010 non-registered articles, while keeping the rate the same for preregistered studies and Registered Reports.

```

##      SR   SO   RR   PR
## No  0.11 0.14 0.06 0.06
## Yes 0.14 0.11 0.19 0.19

##
##      Chi squared power calculation
##
##      w = 0.2832368
##      N = 120.0987
##      df = 2
##      sig.level = 0.05
##      power = 0.8
##
## NOTE: N is the number of observations

```

This indicates an effect size of $w = 0.283$ and sample size of 120.

Finally, Scenario 4 combines the two previous changes into a single model:

```

##      SR   SO   RR   PR
## No  0.11 0.14 0.06 0.11
## Yes 0.14 0.11 0.19 0.14

##
##      Chi squared power calculation
##
##      w = 0.2327814
##      N = 177.8038
##      df = 2
##      sig.level = 0.05

```

```
##           power = 0.8
##
## NOTE: N is the number of observations
```

This indicates an effect size of $w = 0.233$ and sample size of 178.

Taken together, this set of analyses suggests a range of effect sizes of, $w = 0.178\text{--}0.283$, and a range of sample sizes of, $N = 120\text{--}306$. We thus opted for a total sample size of 300 articles.

The following summary text will be inserted into the manuscript:

Full details of the power analysis are available at <https://osf.io/4vbtq/>; here, we provide a summary. The arguments made in the literature that registration will restrict exploratory research has often used extreme language such as that it will “put researchers in chains” (Scott, 2013), “stifle discovery” (Goldin-Meadow, 2016) and serves as a “stranglehold” on research (McDermott, 2022). The severity of this language clearly implies a large effect, that we would see a dramatically lower rate of exploratory work in registered compared to non-registered studies. The only available evidence, however, indicates a moderate difference in the other direction, with a greater prevalence of exploratory research in Registered Reports compared with standard reports (O’Mahony, 2023). The effect size in O’Mahony (2023) is not reported, but we can use the reported test statistic, $\chi^2(1, N = 510) = 16.19$, $p < 0.001$, to calculate it as $\phi = 0.18$. This corresponded to an approximate difference in prevalence of 20%.

It is difficult to say what effect size is the smallest that would be meaningful. Our goal in this project, however, is to address the extreme claims about the damaging impact of registration. Thus, absent any evidence, we would thus primarily be interested in large effects. O’Mahony (2023) provides evidence of a moderate effect, albeit in the other direction. That this effect was found in a registered thesis using a large sample of articles suggests that there is low risk of bias. Thus, we used $\phi = 0.18$ as our estimate for the present study. The `pwr.chisq.test` function from the `pwr` package in R (Champely et al., 2020) indicated a necessary total sample size of 304 for a 2x2 contingency table (registered vs. not registered), with $\alpha = .05$ and $\text{power} = .80$. That assumes that we are collapsing the two registered articles types together and two traditional articles types together and comparing them against each other. If, instead, we compared all four articles types, but assumed that the pattern of exploratory research would be the same across the two broad categories, then that would yield an effect size of Cramer’s $V = 0.21$ and a suggested sample size of 216. Relaxing these assumptions in different ways leads to a range of effect sizes from Cramer’s $V = 0.18\text{--}0.28$ and a range of sample sizes from $N = 120\text{--}306$. Taken together, the power analysis indicates a total sample of 300 articles will be sufficient for the aims of the study.