

Analysis for the Nature of our Literature

Aaron R. Caldwell

2021-09-18

Summary

Within this document I have included a summary of all the analyses included within our manuscript.

```
library(readr)
library(tidyverse)
library(tidyselect)
library(brms)
library(ggdist)
library(distributional)
library(broom)
library(ggpubr)
library(gtsummary)
library(labelled)
library(afex)
library(emmeans)
# Testing out data analysis
# import data -----
df_all = read_csv("df_all.csv") %>%
  mutate(support = factor(support,
                          levels = c("Unclear or not stated",
                                     "Not supported",
                                     "Partial support",
                                     "Full support")) %>%

# Coding error in 20 cases
mutate(sig_test = ifelse(is.na(sig_test), "No", sig_test)) %>%
mutate(sig_test = structure(factor(sig_test, levels = c("Yes", "No")),
                          label = "Significance Testing"),
      hypo_tested = structure(factor(hypo_tested, levels = c("Yes", "No")),
                          label = "Hypothesis Tested"),
      effect_size = structure(factor(effect_size, levels = c("Yes", "No")),
                          label = "Effect Size"))

# Get hypothesis tested set -----
hyp_tested = df_all %>%
  filter(hypo_tested == "Yes")
```

Introduction

In this study we collected data on 300 sport and exercise science research articles (100 from 3 journals). Based on the work of Büttner et al. (2020), we anticipated at least 150 (50%) of the articles would include a hypothesis that was tested. Based on the work of Fanelli (2010), Scheel, Schijen, and Lakens (2021), and Büttner et al. (2020) we hypothesized that the percentage of articles that find support for their hypothesis was greater than 80%.

Hypothesis

For this study, we hypothesized that the rate of positive results (i.e., studies that find at least partial support for their hypothesis) was greater than 80%. Therefore, the null hypothesis (H_0) was that the proportion of positive results was less than .8 and our alternative was greater than .8. There was no other effect being estimated in this study therefore the intercept of the model is what will be tested.

$$H_0 : \textit{Intercept} \leq 0.8$$

$$H_1 : \textit{Intercept} > 0.8$$

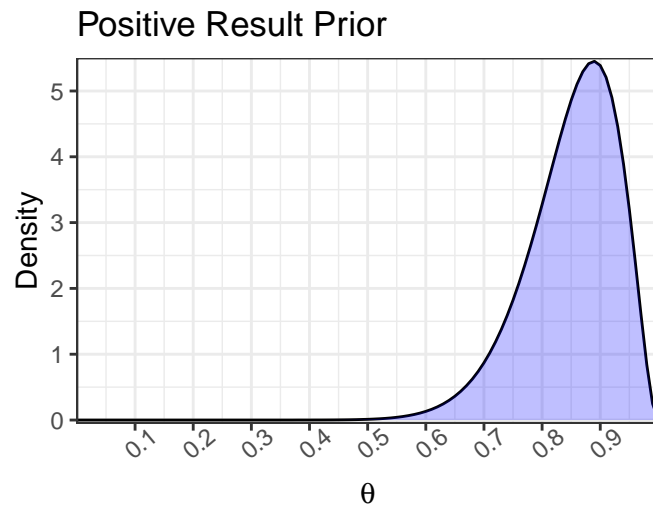
We also hypothesized that more than 60% of studies would test a hypothesis.

$$H_0 : \textit{Intercept} \leq 0.6$$

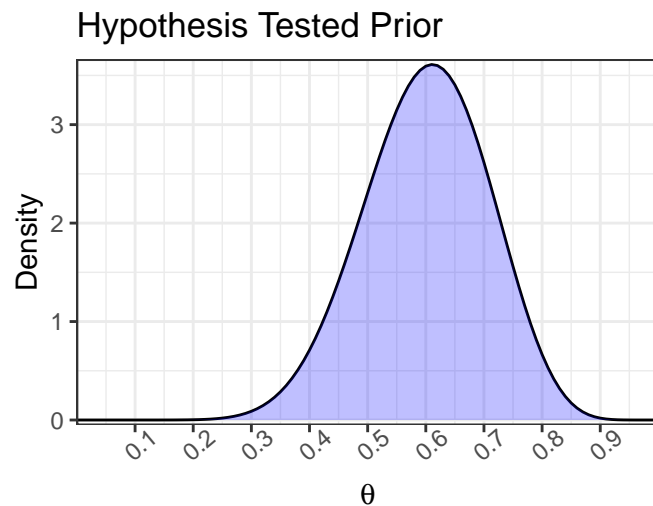
$$H_1 : \textit{Intercept} > 0.6$$

Prior Choice

The prior we selected for this analysis was informed by the previous studies assuming the true positive rate is approximately 85% (Fanelli 2010). However, we would like to avoid “spiking” the prior in favor of our hypothesis and therefore want a skeptical prior. Based on the work of Scheel, Schijen, and Lakens (2021) and Büttner et al. (2020) the estimated positive rates in original research investigations ranged from 82%-92%, and even some fields included in the survey by Fanelli (2010) observed rates as low as ~70%. Therefore, we selected a prior of $\beta(17, 3)$, and is visualized it below. This prior is centered around .85, but includes the possibility of higher (.9) and much lower (.7) proportions as compatible parameter estimates.



Similarly, we used prior centered at 60% for the secondary hypothesis test.



Bayesian Models

We constructed our Bayesian models with the following code.

```
# ANALYSIS 1: Positive Result Rate -----
#Set prior
prior_1 = set_prior("beta(17, 3)", class = "b", lb = 0, ub = 1)

#transform data
hyp_pos = hyp_tested %>%
  summarize(
    pos = sum(di_sup == "Y"),
    N = length(di_sup),
    rate = pos/N
  )

#Build model
m_final <- brm(
  pos | trials(N) ~ 0 + Intercept,
  family = binomial(link = "identity"),
  prior = prior_1,
  data = hyp_pos, sample_prior = "yes", refresh = 0
)
write_rds(m_final, "m_final.rds")

# ANALYSIS 2: Hypothesis Test Rate -----
#Set prior
prior_2 = set_prior("beta(12, 8)", class = "b", lb = 0, ub = 1)

#Generate test data
hyp_test = df_all %>%
  summarize(
    pos = sum(hypo_tested == "Yes"),
    N = length(hypo_tested),
    rate = pos/N
  )

#Build model
m_final2 <- brm(
  pos | trials(N) ~ 0 + Intercept,
  family = binomial(link = "identity"),
  prior = prior_2,
  data = hyp_test, sample_prior = "yes", refresh = 0
)
```

Now, we can import the Bayesian models to get the output for the manuscript.

```
# Import brms analysis 1 (positive) -----
m_final = read_rds("m_final.rds")
# Import brms analysis 2 (prop. of hyp. tested) -----
m_final2 = read_rds("m_final2.rds")

h_test <- hypothesis(m_final, "Intercept > 0.8")
knitr::kable(h_test$hypothesis, caption = "Hypothesis Test #1")
```

Table 1: Hypothesis Test #1

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
(Intercept)-(0.8) > 0	0.0142668	0.0269525	-0.0319772	0.0568162	2.427592	0.70825	

```
h_ci = fixef(m_final)
test_pos = posterior_interval(m_final,
                             prob = .95)
knitr::kable(test_pos, caption = "Hyp Test #1: 95% Posterior C.I.")
```

Table 2: Hyp Test #1: 95% Posterior C.I.

	2.5%	97.5%
b_Intercept	0.7577946	0.8629953
prior_b	0.6781015	0.9662099
lp____	-5.5422325	-3.1629022

```
h_test2 <- hypothesis(m_final2, "Intercept > 0.6")
knitr::kable(h_test2$hypothesis, caption = "Hypothesis Test #2")
```

Table 3: Hypothesis Test #2

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
(Intercept)-(0.8) > 0	0.0142668	0.0269525	-0.0319772	0.0568162	2.427592	0.70825	

```
h_ci2 = fixef(m_final2)
test_pos2 = posterior_interval(m_final2,
                              prob = .95)
knitr::kable(test_pos2, caption = "Hyp Test #2: 95% Posterior C.I.")
```

Table 4: Hyp Test #2: 95% Posterior C.I.

	2.5%	97.5%
b_Intercept	0.5811895	0.6897217
prior_b	0.3766667	0.8015812
lp____	-5.9064835	-3.2418816

```

# Main Figures

dat_mfinal = posterior_samples(m_final, "b") %>%
  mutate(Test = "Positive Result Rate")
dat_mfinal2 = posterior_samples(m_final2, "b") %>%
  mutate(Test = "Rate of Hypothesis Tests")

df_mfinal = rbind(dat_mfinal, dat_mfinal2)

# figure 1
p_fla = df_mfinal %>%
  ggplot(aes(x=b_Intercept,
             fill = Test)) +
  stat_halfeye(alpha = .75) +
  labs(fill = "Interval",
       x = "Probability",
       y = "") +
  theme_bw() +
  facet_wrap(~Test) +
  scale_fill_manual(values = c("lightgreen", "skyblue2")) +
  theme(legend.position = "none",
       axis.text.y=element_blank(),
       axis.ticks.y=element_blank(),
       text = element_text(size = 14,
                           face = "bold"))

fig_1b = df_all %>%
  select(support) %>%
  drop_na() %>%
  ggplot(aes(support,
             fill = support)) +
  geom_bar(aes(y = (..count..) / sum(..count..)),
          color = "black") +
  scale_y_continuous(labels = scales::percent,
                    limits = c(0,.5),
                    breaks = c(0,.1,.2,.3,.4,.5),
                    expand = c(0,0)) +
  labs(x = "Level of Hypothesis Support",
       y = "Relative Frequency",
       fill = "Hypothesis Support") +
  theme_classic() +
  scale_fill_viridis_d(option = "E") +
  theme(legend.position = "none")

fig_1c = df_all %>%
  select(hypo_tested) %>%
  drop_na() %>%
  ggplot(aes(hypo_tested,
             fill = hypo_tested)) +
  geom_bar(aes(y = (..count..) / sum(..count..)),
          color = "black") +
  scale_y_continuous(labels = scales::percent,
                    limits = c(0,.75),

```

```

                breaks = c(0,.25,.5,.75),
                expand = c(0,0)) +
labs(x = "Hypothesis Tested",
     y = "Relative Frequency",
     fill = "") +
theme_classic() +
scale_fill_viridis_d(option = "E") +
theme(legend.position = "none")
fig1 = ggarrange(p_f1a,
                fig_1b,
                hjust = -0.2,
                ncol = 1,
                labels = "AUTO")

```

Levels of Support

```
tab_sup = df_all %>%  
  select(hypo_tested, support) %>%  
  drop_na() %>%  
  group_by(support) %>%  
  summarize(n = n())  
tots_sup = sum(tab_sup$n)  
tab_sup$percent = tab_sup$n/tots_sup*100  
  
knitr::kable(tab_sup)
```

support	n	percent
Unclear or not stated	13	6.806283
Not supported	34	17.801047
Partial support	55	28.795811
Full support	89	46.596859

Exploratory Analyses

Statistics by Hypothesis Tested

```
ctab1 = df_all %>%
  select(effect_size, sig_test,
         hypo_tested, pval_type,
         n_just) %>%
  mutate(pval_type = structure(pval_type, label = "P-value reported"),
         n_just = structure(n_just,
                           label = "Sample Size Justification")) %>%
  tbl_summary(by = hypo_tested,
             type = list(sig_test ~ "categorical",
                        effect_size ~ "categorical",
                        n_just ~ "categorical")) %>%
  # add_p(pvalue_fun = ~style_pvalue(.x, digits = 2)) %>%
  add_overall() %>% add_n() %>% modify_header(label ~ "**Variable**") %>%
  modify_spanning_header(c("stat_1", "stat_2") ~ "**Hypothesis Tested**") %>%
  bold_labels()

ctab1
```

Variable	N	Overall, N = 300	Yes, N = 191	No, N = 109
Effect Size	300			
Yes		238 (79%)	159 (83%)	79 (72%)
No		62 (21%)	32 (17%)	30 (28%)
Significance Testing	300			
Yes		270 (90%)	181 (95%)	89 (82%)
No		30 (10%)	10 (5.2%)	20 (18%)
P-value reported	270			
A mix: some exact and relative p-value(s)		151 (56%)	113 (62%)	38 (43%)
Exact (e.g., p = .049)		91 (34%)	52 (29%)	39 (44%)
Relative (e.g., p < .05)		28 (10%)	16 (8.8%)	12 (13%)
Unknown		30	10	20
Sample Size Justification	300			
No		232 (77%)	140 (73%)	92 (84%)
Yes		68 (23%)	51 (27%)	17 (16%)

Effect Size

```
ct_effsize = table(df_all$effect_size)
ct_effsize

##
## Yes  No
## 238  62

binom_eff = binom.test(ct_effsize[1], sum(ct_effsize))
binom_eff

##
## Exact binomial test
##
## data:  ct_effsize[1] and sum(ct_effsize)
## number of successes = 238, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.7430468 0.8377215
## sample estimates:
## probability of success
##           0.7933333

eff_pr = paste0(round(binom_eff$estimate*100,2),"\% [" ,
                  round(binom_eff$conf.int[1]*100,2),", " ,
                  round(binom_eff$conf.int[2]*100,2)
                  ,"]")

eff_pr

## [1] "79.33%\% [74.3, 83.77]"
```

Significance Testing

```
ct_sig = table(df_all$sig_test)
ct_sig
```

```
##
## Yes  No
## 270  30
```

```
binom_sig = binom.test(ct_sig[1], sum(ct_sig))
binom_sig
```

```
##
## Exact binomial test
##
## data:  ct_sig[1] and sum(ct_sig)
## number of successes = 270, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.8603267 0.9315083
## sample estimates:
## probability of success
##                0.9
```

```
sig_pr = paste0(round(binom_sig$estimate*100,2),"\% [",
                 round(binom_sig$conf.int[1]*100,2),", ",
                 round(binom_sig$conf.int[2]*100,2)
                 ,"]")
```

```
sig_pr
```

```
## [1] "90\% [86.03, 93.15]"
```

Significance Testing with NO Hypothesis Tested

```
ct_sig2 = table(subset(df_all, hypo_tested == "No")$sig_test)
ct_sig2
```

```
##
## Yes  No
##  89  20
```

```
binom_sig2 = binom.test(ct_sig2[1], sum(ct_sig2))
binom_sig2
```

```
##
## Exact binomial test
##
## data:  ct_sig2[1] and sum(ct_sig2)
## number of successes = 89, number of trials = 109, p-value = 1.418e-11
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.7309446 0.8841692
## sample estimates:
## probability of success
##              0.8165138
```

```
sig_pr2 = paste0(round(binom_sig2$estimate*100,2),"\% [",
                  round(binom_sig2$conf.int[1]*100,2),", ",
                  round(binom_sig2$conf.int[2]*100,2)
                  ,"]")
```

```
sig_pr2
```

```
## [1] "81.65\\% [73.09, 88.42]"
```

Significance Testing with Hypothesis Tested

```
ct_sig3 = table(subset(df_all, hypo_tested == "Yes")$sig_test)
ct_sig3
```

```
##
## Yes  No
## 181  10
```

```
binom_sig3 = binom.test(ct_sig3[2], sum(ct_sig3))
binom_sig3
```

```
##
## Exact binomial test
##
## data:  ct_sig3[2] and sum(ct_sig3)
## number of successes = 10, number of trials = 191, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.02538906 0.09417495
## sample estimates:
## probability of success
##           0.05235602
```

```
sig_pr3 = paste0(round(binom_sig3$estimate*100,2), "\\% [",
                  round(binom_sig3$conf.int[1]*100,2), ", ",
                  round(binom_sig3$conf.int[2]*100,2)
                  , "]" )
```

```
sig_pr3
```

```
## [1] "5.24\\% [2.54, 9.42]"
```

Manuscripts Reporting “Exact” p-values

```
ct_ptype = table(df_all$pval_type)
ct_ptype

##
## A mix: some exact and relative p-value(s)
##              151
##           Exact (e.g., p = .049)
##              91
##           Relative (e.g., p < .05)
##              28

binom_ptype = binom.test(ct_ptype[[2]],sum(ct_ptype))
binom_ptype

##
## Exact binomial test
##
## data:  ct_ptype[[2]] and sum(ct_ptype)
## number of successes = 91, number of trials = 270, p-value = 9.265e-08
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.2808696 0.3968281
## sample estimates:
## probability of success
##           0.337037

ptype_pr = paste0(round(binom_ptype$estimate*100,2),"\% [",
                    round(binom_ptype$conf.int[1]*100,2),", ",
                    round(binom_ptype$conf.int[2]*100,2)
                    ,"] of manuscripts reported exact p-values for all results (p = .045) versus only rel

ptype_pr

## [1] "33.7\\% [28.09, 39.68] of manuscripts reported exact p-values for all results (p = .045) versus
```

Manuscripts Reporting “Exact” or “Mixed” p-value types

```

ct_ptype2 = table(df_all$pval_type)
ct_ptype2

##
## A mix: some exact and relative p-value(s)
##              151
##          Exact (e.g., p = .049)
##              91
##          Relative (e.g., p < .05)
##              28

binom_ptype2 = binom.test(ct_ptype2[[2]]+ct_ptype2[[1]],sum(ct_ptype2))
binom_ptype2

##
## Exact binomial test
##
## data:  ct_ptype2[[2]] + ct_ptype2[[1]] and sum(ct_ptype2)
## number of successes = 242, number of trials = 270, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.8536141 0.9299761
## sample estimates:
## probability of success
##          0.8962963

ptype_pr2 = paste0(round(binom_ptype2$estimate*100,2),"\% [",
                    round(binom_ptype2$conf.int[1]*100,2),", ",
                    round(binom_ptype2$conf.int[2]*100,2)
                    ,"] of manuscripts reported at least *some* exact p-values (e.g., p = .045) versus r

ptype_pr2

## [1] "89.63\\% [85.36, 93] of manuscripts reported at least *some* exact p-values (e.g., p = .045) versus r

```

Preregistration by Study Type

```
ctab2 = df_all %>%
  select(clin_trial, rct, animal, prereg) %>%
  mutate(clin_trial = structure(clin_trial,
                                label = "Clinical Trial"),
         rct = structure(rct,
                         label = "RCT"),
         animal = structure(animal,
                             label = "Animal Study")) %>%
  tbl_summary(by = prereg,
              type = list(clin_trial ~ "categorical",
                          rct ~ "categorical",
                          animal ~ "categorical")) %>%
  # add_p(pvalue_fun = ~style_pvalue(.x, digits = 2)) %>%
  add_overall() %>% add_n() %>% modify_header(label ~ "**Variable**") %>%
  modify_spanning_header(c("stat_1", "stat_2") ~ "***Preregistration***") %>%
  bold_labels()

ctab2
```

Variable	N	Overall, N = 300	No, N = 273	Yes, N = 27
Clinical Trial	300			
No		260 (87%)	256 (94%)	4 (15%)
Yes		40 (13%)	17 (6.2%)	23 (85%)
RCT	300			
No		236 (79%)	227 (83%)	9 (33%)
Yes		64 (21%)	46 (17%)	18 (67%)
Animal Study	300			
No		291 (97%)	264 (97%)	27 (100%)
Yes		9 (3.0%)	9 (3.3%)	0 (0%)

Preregistration Rate

```
ct_prereg = table(df_all$prereg)
ct_prereg
```

```
##
## No Yes
## 273 27
```

```
binom_prereg = binom.test(ct_prereg[[2]],sum(ct_prereg))
binom_prereg
```

```
##
## Exact binomial test
##
## data: ct_prereg[[2]] and sum(ct_prereg)
## number of successes = 27, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.06014953 0.12824251
## sample estimates:
## probability of success
## 0.09
```

```
prereg_pr = paste0(round(binom_prereg$estimate*100,2),"\\% [",
                    round(binom_prereg$conf.int[1]*100,2)," ",
                    round(binom_prereg$conf.int[2]*100,2)
                    ,"] of manuscripts reporting preregistration or clinical trial registration informati
prereg_pr
```

```
## [1] "9\\% [6.01, 12.82] of manuscripts reporting preregistration or clinical trial registration informati
```

Sample Size by Journal

```
ctab3 = df_all %>%
  select(journal, n_just, sample_info) %>%
  mutate(n_just = structure(n_just,
    label = "Sample Size Justification"),
    sample_info = structure(sample_info,
    label = "Sample Size Information")) %>%
  tbl_summary(by = journal,
    type = list(n_just ~ "categorical",
    sample_info ~ "categorical")) %>%
  # add_p(pvalue_fun = ~style_pvalue(.x, digits = 2)) %>%
  add_overall() %>% add_n() %>% modify_header(label ~ "**Variable**") %>%
  modify_spanning_header(c("stat_1", "stat_2") ~ "**Journal**") %>%
  bold_labels()

ctab3
```

Variable	N	Overall, N = 300	EJSS, N = 100	JSAMS, N = 100	MSSE, N = 100
Sample Size Justification	300				
No		232 (77%)	81 (81%)	86 (86%)	65 (65%)
Yes		68 (23%)	19 (19%)	14 (14%)	35 (35%)
Sample Size Information	300				
No (sample size(s) unclear/unreported)		5 (1.7%)	0 (0%)	4 (4.0%)	1 (1.0%)
Partial (some information missing)		2 (0.7%)	0 (0%)	1 (1.0%)	1 (1.0%)
Yes (e.g., total and group sample sizes provided)		293 (98%)	100 (100%)	95 (95%)	98 (98%)

Sample Size Justification

```
ct_njust = table(df_all$n_just)
ct_njust

##
## No Yes
## 232 68

binom_njust = binom.test(ct_njust[[2]], sum(ct_njust))
binom_njust

##
## Exact binomial test
##
## data: ct_njust[[2]] and sum(ct_njust)
## number of successes = 68, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.1805308 0.2782908
## sample estimates:
## probability of success
## 0.2266667

njust_pr = paste0(
  round(binom_njust$estimate * 100, 2),
  "% [",
  round(binom_njust$conf.int[1] * 100, 2),
  ", ",
  round(binom_njust$conf.int[2] * 100, 2)
  ,
  "]"
)
njust_pr

## [1] "22.67% [18.05, 27.83]"
```

Sample Size Information

```
ct_samp = table(df_all$sample_info)
ct_samp

##
##           No (sample size(s) unclear/unreported)
##                                     5
##           Partial (some information missing)
##                                     2
## Yes (e.g., total and group sample sizes provided)
##                                     293

binom_samp = binom.test(ct_samp[[3]], sum(ct_samp))
binom_samp

##
## Exact binomial test
##
## data:  ct_samp[[3]] and sum(ct_samp)
## number of successes = 293, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.9525174 0.9905685
## sample estimates:
## probability of success
##           0.9766667

samp_pr = paste0(
  round(binom_samp$estimate * 100, 2),
  "% [",
  round(binom_samp$conf.int[1] * 100, 2),
  ", ",
  round(binom_samp$conf.int[2] * 100, 2)
  ,
  "]" of manuscripts reported all the required sample size information (total and group sample sizes).")
)
samp_pr

## [1] "97.67% [95.25, 99.06] of manuscripts reported all the required sample size information (total and group sample sizes)."
```

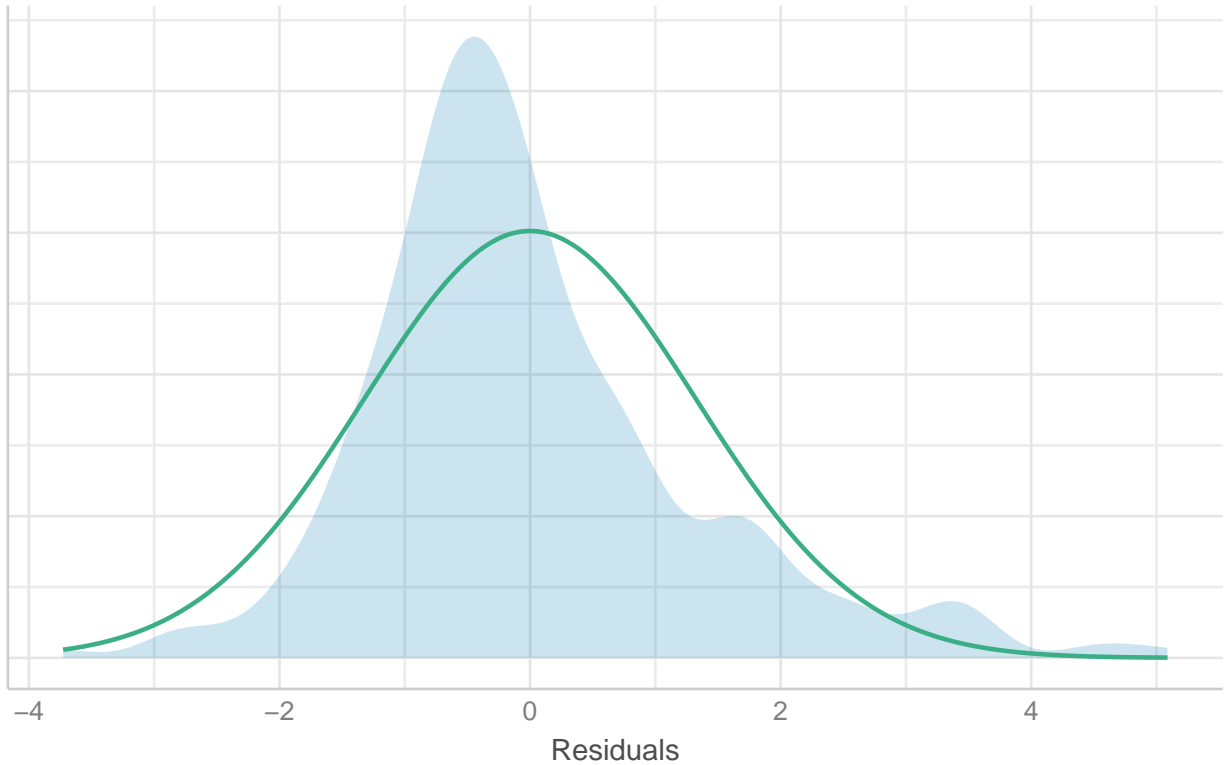
Sample Size by Discipline

```
aov_1way = df_all %>%
  select(n, sci_cat, doi) %>% mutate(n = as.numeric(n)) %>% drop_na() %>%
  afex::aov_4(formula = log(n) ~ sci_cat + (1|doi))
# Not perfect fit but close
# fit is best with log transform compared to others
# See MASS::boxcox(aov_1way); very close to best fit
plot(performance::check_normality(aov_1way))

## Warning: Non-normality of residuals detected (p < .001).
```

Normality of Residuals

Distribution should be close to the normal curve



```
knitr::kable(nice(aov_1way))
```

Effect	df	MSE	F	ges	p.value
sci_cat	9, 285	1.81	21.81 ***	.408	<.001

```
emm_samps = emmeans::emmeans(aov_1way, ~ sci_cat, type = "response")
```

```
knitr::kable(emm_samps)
```

sci_cat	response	SE	df	lower.CL	upper.CL
Applied exercise physiology (human)	30.06059	4.285230	285	22.705895	39.79756
Basic physiology (animal and cell physiology)	35.03300	15.704629	285	14.496810	84.66075
Biomechanics	25.68742	5.093465	285	17.386780	37.95089
Clinical research	41.54535	10.375167	285	25.412325	67.92044
Environmental physiology (heat, cold, & altitude)	15.86997	6.749130	285	6.871256	36.65356
Epidemiology	1161.57881	306.361126	285	691.179304	1952.12057
Motor learning/control/behavior	68.17904	21.035155	285	37.145928	125.13838
Other	165.97110	42.955841	285	99.721555	276.23320
Sport performance	50.48799	12.831597	285	30.614795	83.26161
Sport/exercise psychology	66.11571	25.667626	285	30.792118	141.96125

Other Open Science Practices

Data Availability Statements

```
ct_datstat = table(df_all$data_state)
ct_datstat

##
## No Yes
## 293 7

binom_datstat = binom.test(ct_datstat[[2]],300)
binom_datstat

##
## Exact binomial test
##
## data: ct_datstat[[2]] and 300
## number of successes = 7, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.009431502 0.047482565
## sample estimates:
## probability of success
## 0.02333333

datstat_pr = paste0(round(binom_datstat$estimate*100,2),"% [",
                    round(binom_datstat$conf.int[1]*100,2),", ",
                    round(binom_datstat$conf.int[2]*100,2)
                    ,"] of manuscripts had a data accessibility statement")
datstat_pr

## [1] "2.33% [0.94, 4.75] of manuscripts had a data accessibility statement"
```

Open Data Availability

```
ct_odat = table(df_all$open_data)
ct_odat

##
## No Yes
## 5 2

binom_odat = binom.test(ct_odat[[2]],300)
binom_odat

##
## Exact binomial test
##
## data: ct_odat[[2]] and 300
## number of successes = 2, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.0008083864 0.0238734976
## sample estimates:
## probability of success
## 0.006666667

odat_pr = paste0(round(binom_odat$estimate*100,2),"% [",
                  round(binom_odat$conf.int[1]*100,2),", ",
                  round(binom_odat$conf.int[2]*100,2)
                  ,"] of manuscripts reported some form of data sharing or open data")
odat_pr

## [1] "0.67% [0.08, 2.39] of manuscripts reported some form of data sharing or open data"
```

Replication Studies

```
ct_replic = table(df_all$replic)
ct_replic
```

```
##
## No
## 300
```

```
binom_replic = binom.test(0,300)
binom_replic
```

```
##
## Exact binomial test
##
## data: 0 and 300
## number of successes = 0, number of trials = 300, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.00000000 0.01222097
## sample estimates:
## probability of success
## 0
```

```
replic_pr = paste0(
  round(binom_replic$estimate * 100, 2),
  "% [",
  round(binom_replic$conf.int[1] * 100, 2),
  ", ",
  round(binom_replic$conf.int[2] * 100, 2)
  ,
  "]" of manuscripts explicitly stated they were replicating a previous study."
)
replic_pr
```

```
## [1] "0% [0, 1.22] of manuscripts explicitly stated they were replicating a previous study."
```


Analysis by Journal

Hypothesis Support

```
## Hypothesis Support -----
```

```
tab_jhyp = table(df_all$journal,df_all$support)
tab_jhyp
```

```
##
##      Unclear or not stated Not supported Partial support Full support
##  EJSS                      5          13             19          34
##  JSAMS                      2           8             17          19
##  MSSE                      6          13             19          36
```

```
chisq_support = chisq.test(tab_jhyp)
chisq_support
```

```
##
##  Pearson's Chi-squared test
##
## data:  tab_jhyp
## X-squared = 2.4004, df = 6, p-value = 0.8794
```

Hypothesis Tested

```
## Hypothesis Tested -----  
tab_jtest = table(df_all$journal,df_all$hypo_tested)  
tab_jtest
```

```
##  
##           Yes No  
##  EJSS      71 29  
##  JSAMS      46 54  
##  MSSE      74 26
```

```
chisq_jtest = chisq.test(tab_jtest)  
chisq_jtest
```

```
##  
##  Pearson's Chi-squared test  
##  
## data:  tab_jtest  
## X-squared = 20.433, df = 2, p-value = 3.656e-05
```

Significance Testing

```
## Significance Testing -----
tab_jsig = table(df_all$journal,df_all$sig_test)
tab_jsig

##
##      Yes No
##  EJSS   92  8
##  JSAMS   84 16
##  MSSE   94  6

chisq_jsig = chisq.test(tab_jsig)
chisq_jsig

##
##  Pearson's Chi-squared test
##
## data:  tab_jsig
## X-squared = 6.2222, df = 2, p-value = 0.04455
```

Effect Sizes

```
## Effect Size -----  
tab_jes = table(df_all$journal,df_all$effect_size)  
tab_jes
```

```
##  
##           Yes No  
##    EJSS    90 10  
##    JSAMS    72 28  
##    MSSE     76 24
```

```
chisq_jes = chisq.test(tab_jes)  
chisq_jes
```

```
##  
##  Pearson's Chi-squared test  
##  
## data:  tab_jes  
## X-squared = 10.897, df = 2, p-value = 0.004302
```

Sample Size Justification

```
## Sample Size Justification -----
tab_jjust = table(df_all$journal,df_all$n_just)
tab_jjust

##
##           No Yes
##  EJSS   81  19
##  JSAMS  86  14
##  MSSE   65  35

chisq_jjust = chisq.test(tab_jjust)
chisq_jjust

##
##  Pearson's Chi-squared test
##
## data:  tab_jjust
## X-squared = 13.73, df = 2, p-value = 0.001044
```

Analysis among Clinical Trials

Hypothesis Support

```
# Clinical Trial breakdown -----
df_clin = subset(df_all, clin_trial == "Yes")

## Hypothesis Support (di)
tab_clindisup = table(df_clin$di_sup)
tab_clindisup

##
##   N   Y
##  9  16

binom_clindisup = binom.test(tab_clindisup[2], sum(tab_clindisup),
                             p = .8)
binom_clindisup

##
##   Exact binomial test
##
## data:  tab_clindisup[2] and sum(tab_clindisup)
## number of successes = 16, number of trials = 25, p-value = 0.07416
## alternative hypothesis: true probability of success is not equal to 0.8
## 95 percent confidence interval:
##  0.4252063 0.8202832
## sample estimates:
## probability of success
##                0.64

tab_clinsup = table(df_clin$support)
knitr::kable(tab_clinsup)
```

Var1	Freq
Unclear or not stated	2
Not supported	9
Partial support	6
Full support	10

Hypothesis Tested

```
## Hypothesis Tested
tab_clinhypo = table(df_clin$hypo_tested)
tab_clinhypo

##
## Yes No
## 27 13

binom_clinhypo = binom.test(tab_clinhypo[2], sum(tab_clinhypo),
                             p = .6)
binom_clinhypo

##
## Exact binomial test
##
## data: tab_clinhypo[2] and sum(tab_clinhypo)
## number of successes = 13, number of trials = 40, p-value = 0.0005468
## alternative hypothesis: true probability of success is not equal to 0.6
## 95 percent confidence interval:
## 0.1857290 0.4912949
## sample estimates:
## probability of success
## 0.325
```

Sample Size Justification

```
## Sample Size Just -----
tab_clinjust = table(df_clin$n_just)
tab_clinjust

##
## No Yes
## 15 25

binom_clinjust = binom.test(tab_clinjust[2], sum(tab_clinjust))
binom_clinjust

##
## Exact binomial test
##
## data: tab_clinjust[2] and sum(tab_clinjust)
## number of successes = 25, number of trials = 40, p-value = 0.1539
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.4580148 0.7727373
## sample estimates:
## probability of success
## 0.625
```

Preregistration

```
tab_clinreg = table(df_clin$prereg)

binom_clinreg = binom.test(tab_clinreg[2], sum(tab_clinreg))

### by journal -----
tab_clinregj = table(df_clin$prereg, df_clin$journal)
```

Analysis by RCT

Hypothesis Support

```
# RCT breakdown -----

df_rct = subset(df_all, rct == "Yes")

## Hypothesis Support (di)
tab_rctdisup = table(df_rct$di_sup)
tab_rctdisup

##
## N Y
## 11 33

binom_rctdisup = binom.test(tab_rctdisup[2], sum(tab_rctdisup),
                             p = .8)
binom_rctdisup

##
## Exact binomial test
##
## data: tab_rctdisup[2] and sum(tab_rctdisup)
## number of successes = 33, number of trials = 44, p-value = 0.4493
## alternative hypothesis: true probability of success is not equal to 0.8
## 95 percent confidence interval:
## 0.5966196 0.8680726
## sample estimates:
## probability of success
## 0.75

tab_rctsup = table(df_rct$support)

knitr::kable(tab_rctsup)
```

Var1	Freq
Unclear or not stated	3
Not supported	11
Partial support	13
Full support	20

Hypothesis Tested

```
## Hypothesis Tested
tab_rct hypo = table(df_rct$hypo_tested)
binom_rct hypo = binom.test(tab_rct hypo[2], sum(tab_rct hypo),
                             p = .6)
```

Sample Size Justification

```
## Sample Size Just -----
tab_rctjust = table(df_rct$n_just)
tab_rctjust

##
## No Yes
## 32 32

binom_rctjust = binom.test(tab_rctjust[2], sum(tab_rctjust))
binom_rctjust

##
## Exact binomial test
##
## data: tab_rctjust[2] and sum(tab_rctjust)
## number of successes = 32, number of trials = 64, p-value = 1
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.3723229 0.6276771
## sample estimates:
## probability of success
## 0.5
```

Preregistration

```
tab_rctreg = table(df_rct$prereg)

binom_rctreg = binom.test(tab_rctreg[2], sum(tab_rctreg))
```

Analysis by Discipline

```
tab_dissupp = table(df_all$sci_cat, df_all$support)
tab_dissupp
```

```
##
##                                     Unclear or not stated
## Applied exercise physiology (human)                               3
## Basic physiology (animal and cell physiology)                     3
## Biomechanics                                                       5
## Clinical research                                                  1
## Environmental physiology (heat, cold, & altitude)                 0
## Epidemiology                                                       0
## Motor learning/control/behavior                                    0
## Other                                                               0
## Sport performance                                                  0
## Sport/exercise psychology                                          1
##
##                                     Not supported
## Applied exercise physiology (human)                               18
## Basic physiology (animal and cell physiology)                     0
## Biomechanics                                                       4
## Clinical research                                                  4
## Environmental physiology (heat, cold, & altitude)                 0
## Epidemiology                                                       1
## Motor learning/control/behavior                                    0
## Other                                                               1
## Sport performance                                                  4
## Sport/exercise psychology                                          2
##
##                                     Partial support
## Applied exercise physiology (human)                               16
## Basic physiology (animal and cell physiology)                     1
## Biomechanics                                                       14
## Clinical research                                                  4
## Environmental physiology (heat, cold, & altitude)                 4
## Epidemiology                                                       1
## Motor learning/control/behavior                                    3
## Other                                                               3
## Sport performance                                                  6
## Sport/exercise psychology                                          3
##
##                                     Full support
## Applied exercise physiology (human)                               27
## Basic physiology (animal and cell physiology)                     4
## Biomechanics                                                       15
## Clinical research                                                  9
```

```
## Environmental physiology (heat, cold, & altitude) 3
## Epidemiology 5
## Motor learning/control/behavior 9
## Other 8
## Sport performance 7
## Sport/exercise psychology 2
```

```
chisq_dissupp = chisq.test(tab_dissupp)
chisq_dissupp
```

```
##
## Pearson's Chi-squared test
##
## data:  tab_dissupp
## X-squared = 40.022, df = 27, p-value = 0.051
```

Hypothesis Tested

```
tab_dishypop = table(df_all$sci_cat, df_all$hypo_tested)
tab_dishypop
```

```
##
##                                     Yes No
##   Applied exercise physiology (human)    64 25
##   Basic physiology (animal and cell physiology)    8  1
##   Biomechanics                             38 12
##   Clinical research                        18 11
##   Environmental physiology (heat, cold, & altitude)    7  3
##   Epidemiology                             7 19
##   Motor learning/control/behavior          12  7
##   Other                                    12 15
##   Sport performance                       17 12
##   Sport/exercise psychology                8  4
```

```
chisq_dishypop = chisq.test(tab_dishypop)
chisq_dishypop
```

```
##
##   Pearson's Chi-squared test
##
## data:  tab_dishypop
## X-squared = 28.438, df = 9, p-value = 0.0008056
```

Plots

```
p_dissup = df_all %>%
  group_by(support, sci_cat) %>%
  summarize(count = n(),
            .groups = 'drop') %>%
  filter(!is.na(support)) %>%
  ggplot( aes(fill=support, y=count, x=sci_cat)) +
  geom_bar(position="fill", stat="identity",
           color = "black")+
  scale_y_continuous(labels = scales::percent) +
  labs(x = "",
       y = "Relative Frequency",
       fill = "") +
  theme_classic() +
  scale_fill_viridis_d(option = "E") +
  theme(legend.position = "bottom") +
  coord_flip()+
  theme(text = element_text(face = "bold"))

p_dishypo = df_all %>%
  group_by(hypo_tested, sci_cat) %>%
  summarize(count = n(),
            .groups = 'drop') %>%
  filter(!is.na(hypo_tested)) %>%
  ggplot( aes(fill=hypo_tested, y=count, x=sci_cat)) +
  geom_bar(position="fill", stat="identity",
           color = "black")+
  scale_y_continuous(labels = scales::percent) +
  labs(x = "",
       y = "Relative Frequency",
       fill = "Hypothesis Tested") +
  theme_classic() +
  scale_fill_viridis_d(option = "E") +
  theme(legend.position = "bottom") +
  coord_flip() +
  theme(text = element_text(face = "bold"))

emm_plot = plot(emm_samps) +
  scale_x_continuous(trans = "log",
                    breaks = c(10,15,20,30,40,50,65,80,110,150,1095)) +
  labs(x = "Estimated Mean Sample Size (log scale)",
       y = "") +
  theme_bw() +
  theme(text = element_text(face = "bold"))

fig3 = ggarrange(p_dishypo,p_dissup,emm_plot,
                 ncol = 1,
                 labels = "AUTO")
```

```
# Main Figures -----
```

```
p_2a = df_all %>%
  group_by(journal, support) %>%
  summarize(count = n(),
            .groups = 'drop') %>%
  filter(!is.na(support)) %>%
  ggplot( aes(fill=support, y=count, x=journal)) +
  geom_bar(position="fill", stat="identity",
            color = "black")+
  scale_y_continuous(labels = scales::percent) +
  labs(x = "Journal",
       y = "Relative Frequency",
       fill = "Hypothesis Support") +
  theme_classic() +
  scale_fill_viridis_d(option = "E") +
  theme(legend.position = "top",
        text = element_text(face = "bold"))

p_2b = df_all %>%
  group_by(journal, hypo_tested) %>%
  summarize(count = n(),
            .groups = 'drop') %>%
  filter(!is.na(hypo_tested)) %>%
  ggplot( aes(fill=hypo_tested, y=count, x=journal)) +
  geom_bar(position="fill", stat="identity",
            color = "black")+
  scale_y_continuous(labels = scales::percent) +
  labs(x = "Journal",
       y = "Relative Frequency",
       fill = "Hypothesis Tested") +
  theme_classic() +
  scale_fill_viridis_d(option = "E") +
  theme(legend.position = "top",
        text = element_text(face = "bold"))

p_2c = df_all %>%
  group_by(journal, effect_size) %>%
  summarize(count = n(),
            .groups = 'drop') %>%
  filter(!is.na(effect_size)) %>%
  ggplot( aes(fill=effect_size, y=count, x=journal)) +
  geom_bar(position="fill", stat="identity",
            color = "black")+
  scale_y_continuous(labels = scales::percent) +
  labs(x = "Journal",
       y = "Relative Frequency",
       fill = "Effect Size Reported") +
  theme_classic() +
  scale_fill_viridis_d(option = "E") +
  theme(legend.position = "top",
        text = element_text(face = "bold"))
```

```

p_2d = df_all %>%
  group_by(journal, n_just) %>%
  summarize(count = n(),
            .groups = 'drop') %>%
  filter(!is.na(n_just)) %>%
  ggplot(aes(fill=n_just, y=count, x=journal)) +
  geom_bar(position="fill", stat="identity",
            color = "black")+
  scale_y_continuous(labels = scales::percent) +
  labs(x = "Journal",
        y = "Relative Frequency",
        fill = "Sample Size Justification") +
  theme_classic() +
  scale_fill_viridis_d(option = "E") +
  theme(legend.position = "top",
        text = element_text(face = "bold"))

fig_2 = ggarrange(p_2b, p_2a, p_2c, p_2d,
                  ncol = 1,
                  labels = "AUTO")

```


Figure 1

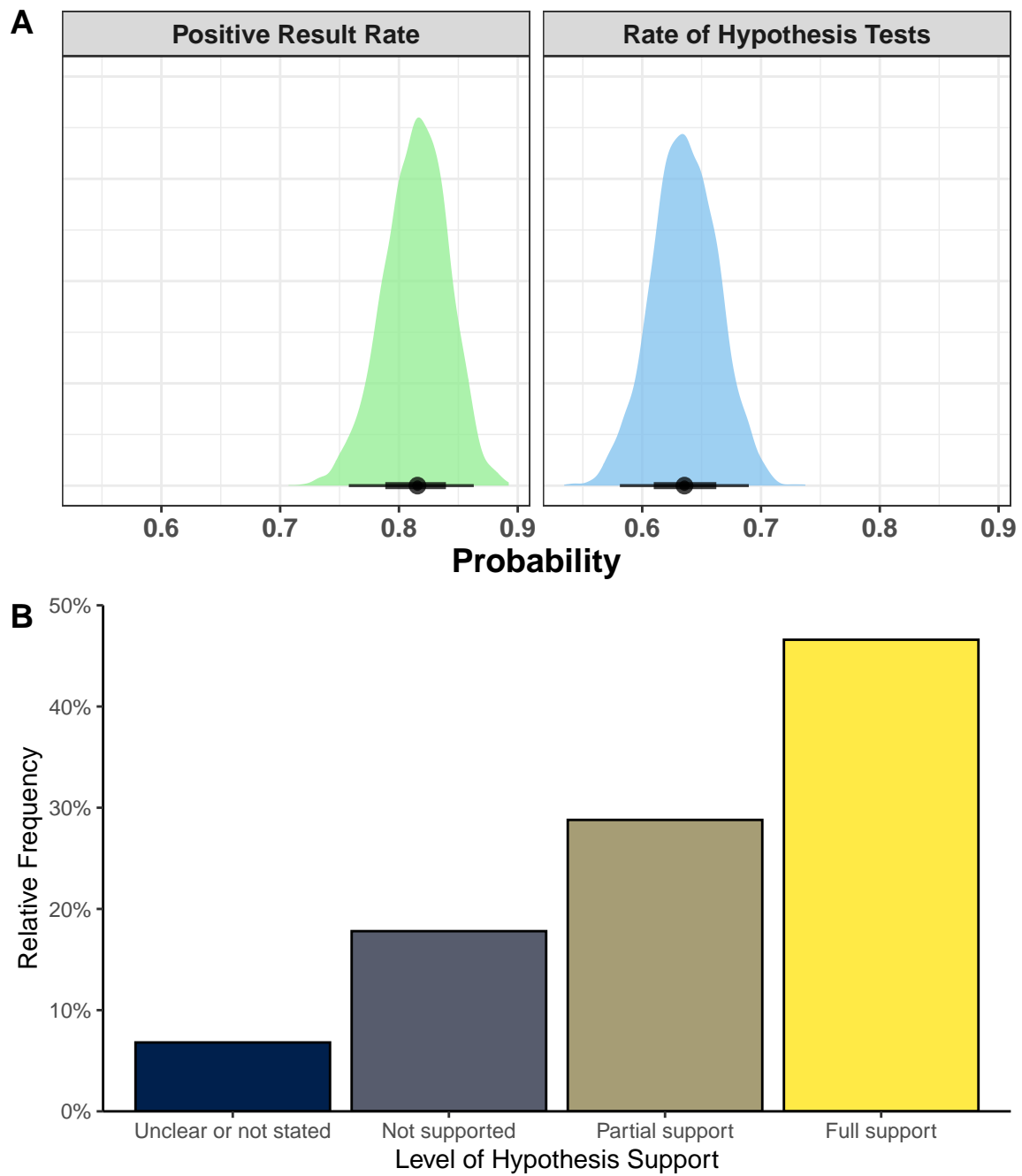


Figure 1. A) Posterior distributions from Bayesian model with the 50% and 95% percent compatibility intervals represented by the error bars at the bottom and B) Relative frequencies of the level of support reported for manuscripts with hypotheses (N = 191) with 17.8% report no support, 28.8% stating partial support, 46.6% stating full support, and 6.81% for which support was unclear or not stated.

Figure 2

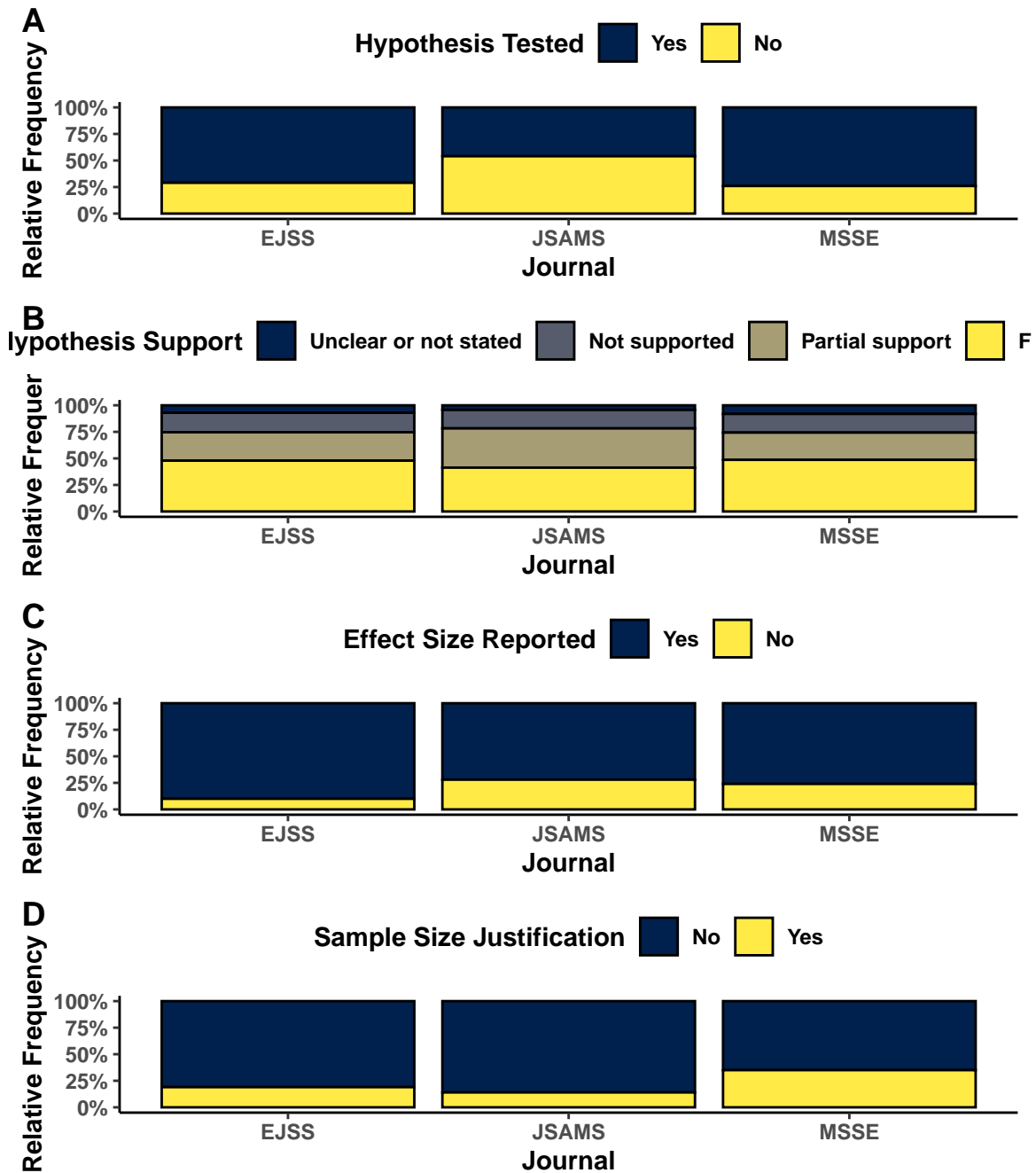


Figure 2. Relative frequencies, by journal, for A) level of reported support for hypotheses, B) indication of whether a hypothesis was tested, C) indication of whether an effect size was reported, or D) indication of if sample size was justified by the authors. Journals included the European Journal of Sport Science (EJSS), the Journal of Science and Medicine in Sport (JSAMS), and Medicine and Science in Sport and Exercise (MSSE),

Figure 3

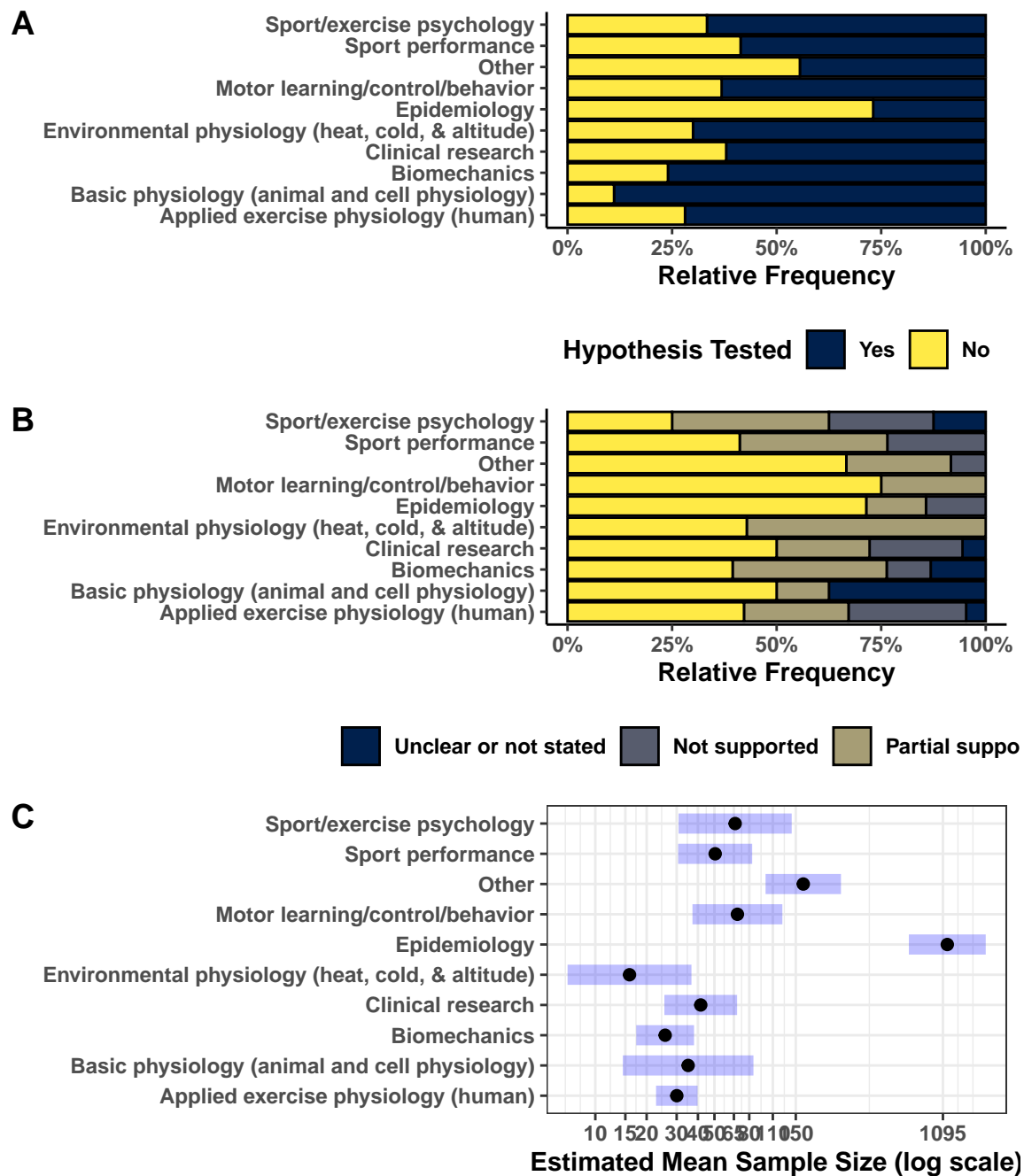


Figure 3. The breakdown, by discipline, for A) indication of whether a hypothesis was tested B) level of reported support for hypotheses, and C) the estimated total sample size (grey bands indicate 95% confidence intervals).

References

- Büttner, Fionn, Elaine Toomey, Shane McClean, Mark Roe, and Eamonn Delahunt. 2020. “Are Questionable Research Practices Facilitating New Discoveries in Sport and Exercise Medicine? The Proportion of Supported Hypotheses Is Implausibly High.” *British Journal of Sports Medicine*. <https://doi.org/10.1136/bjsports-2019-101863>.
- Fanelli, Daniele. 2010. “‘Positive’ Results Increase down the Hierarchy of the Sciences.” *PLOS ONE* 5 (4): e10068. <https://doi.org/10.1371/journal.pone.0010068>.
- Scheel, Anne M., Mitchell R. M. J. Schijen, and Daniël Lakens. 2021. “An Excess of Positive Results: Comparing the Standard Psychology Literature with Registered Reports.” *Advances in Methods and Practices in Psychological Science* 4 (2): 251524592110074. <https://doi.org/10.1177/25152459211007467>.