

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(epitab)
library(gtsummary)
library(labelled)
# 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.

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

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

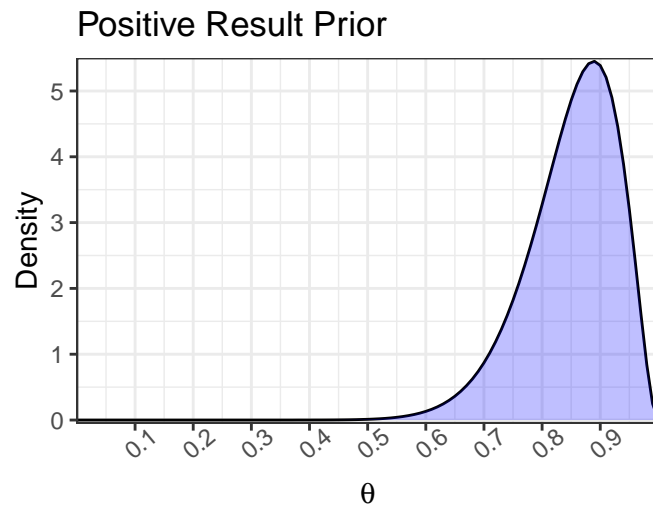
We also hypothesized that more than 60% of studies would test a hypothesis.

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

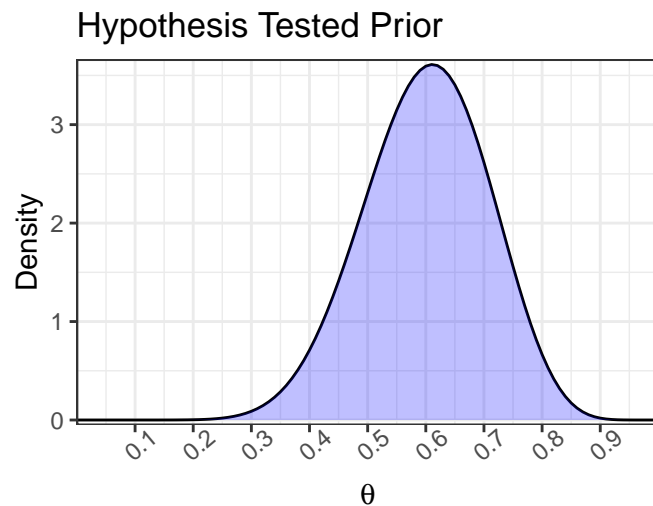
$$H_1 : \textit{Intercept} > 0.6$$

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")

```

Figure 1

```
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

fig1

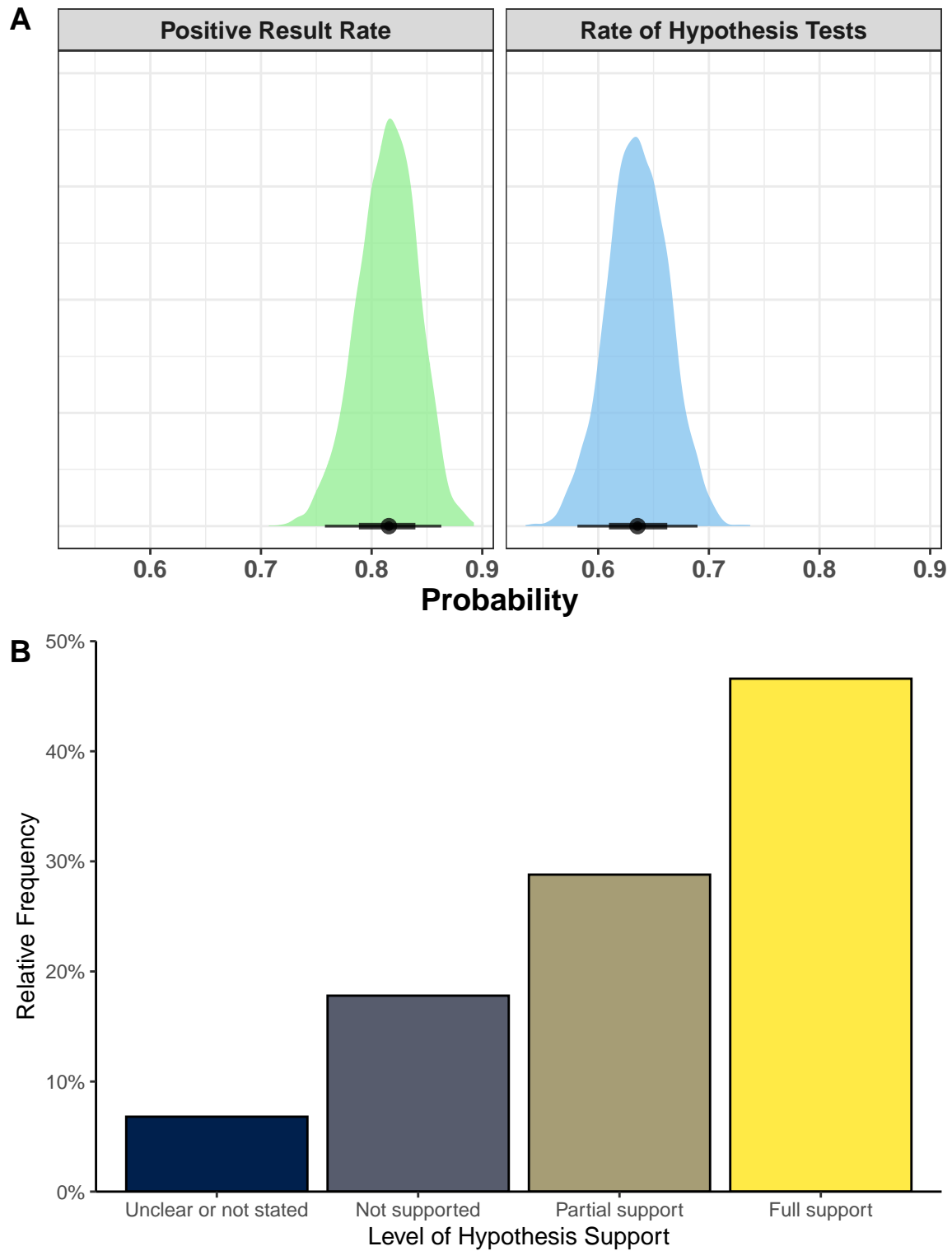


Figure 1. A) Posterior distributions from Bayesian model with the 50% and 95% percent compatibility intervals represented by the error bars at the bottom and B) Relative frequencies of the level of support reported for manuscripts with hypotheses (N = 191) with 17.8% report no support, 28.8% stating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

Exploratory Analyses

Split by Hypothesis Tested

```
ctab1 = df_all %>%
  #filter(hypo_tested == "Yes") %>%
  select(effect_size, sig_test, hypo_tested) %>%
  tbl_summary(by = hypo_tested,
              type = list(sig_test ~ "categorical",
                          effect_size ~ "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**") %>%
  #modify_caption("**Table 1. Patient Characteristics**") %>%
  bold_labels()

ctab1
```

Variable	N	Overall, N = 300	Yes, N = 191	No, N = 109	p-value
Effect Size	300				0.027
Yes		238 (79%)	159 (83%)	79 (72%)	
No		62 (21%)	32 (17%)	30 (28%)	
Significance Testing	300				<0.001
Yes		270 (90%)	181 (95%)	89 (82%)	
No		30 (10%)	10 (5.2%)	20 (18%)	

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]"

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

binom_sig2 = binom.test(ct_sig2[2], sum(ct_sig2))
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)
                  ,"]")

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

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

```

ct_ptype = table(df_all$pval_type)
binom_ptype = binom.test(ct_ptype[[2]],sum(ct_ptype))
ptype_pr = paste0(round(binom_ptype$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

ct_ptype2 = table(df_all$pval_type)

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

ctab2 = df_all %>%
  filter(sig_test == "Yes") %>%
  select(effect_size, pval_sig, pval_type) %>%
  mutate_all(as.factor) %>%
  drop_na() %>%
  contingency_table(independents=list("Effect Size Reported" = "effect_size",
                                     "p-value Type" = "pval_type"),
                   outcomes=list("Significant p-value" = "pval_sig"),
                   crosstab_funcs=list(freq())) #>%
  #neat_table(caption = "Statistics Reported")

# Prereg Descriptives -----

ctab_prereg = df_all %>%
  select(clin_trial, rct, animal, prereg) %>%
  mutate_all(as.factor) %>%
  drop_na() %>%
  contingency_table(
    independents = list(
      "Clinical Trial" = "clin_trial",
      "RCT" = "rct",
      "Animal Study" = "animal"
    ),
    outcomes = list("Preregistration" = "prereg"),
    crosstab_funcs=list(freq()))

ct_prereg = table(df_all$prereg)

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

# Sample Size Information -----

ctab_ss = df_all %>%

```

```

select(journal, n_just, sample_info) %>%
mutate_all(as.factor) %>%
drop_na() %>%
contingency_table(
  independents = list(
    "Sample Size Justification" = "n_just",
    "Sample Size Reported" = "sample_info"
  ),
  outcomes = list("Journal" = "journal"),
  crosstab_funcs = list(freq())
)

ct_njust = table(df_all$n_just)

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

ct_samp = table(df_all$sample_info)

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

# Sample Size Analysis -----

# problems with mediway not sure if result is accurate
samp_1way = df_all %>%
  select(n, sci_cat) %>%
  mutate(n = as.numeric(n)) %>%
  drop_na() %>%
  WRS2::mediway(formula = n ~ sci_cat)

# checking residuals; aov are fair
library(afex)
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))

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

# Other open Science Practices -----

ct_datstat = table(df_all$data_state)

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

ct_odat = table(df_all$open_data)

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

ct_replic = table(df_all$replic)

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

p_n = df_all %>%
  ggplot(aes(x = as.numeric(n))) +
  geom_boxplot(fill = "skyblue3",
              alpha = 0.55) +
  labs(title = "Sample Size by Discipline",
       x = "Total Sample Size (log scale)") +
  theme_bw() +
  facet_wrap(~sci_cat,
            scales = "free",
            ncol = 2) +
  scale_x_continuous(trans = "log10") +
  theme(axis.ticks.y = element_blank(),

```



```

axis.text.y = element_blank())

# By Journal ----

## Hypothesis Support -----
tab_jhyp = table(df_all$journal, df_all$support)
chisq_support = chisq.test(tab_jhyp)

## Hypothesis Tested -----
tab_jtest = table(df_all$journal, df_all$hypo_tested)
chisq_jtest = chisq.test(tab_jtest)

## Significance Testing -----
tab_jsig = table(df_all$journal, df_all$sig_test)
chisq_jsig = chisq.test(tab_jsig)

## Effect Size -----
tab_jes = table(df_all$journal, df_all$effect_size)
chisq_jes = chisq.test(tab_jes)

## Sample Size Justification -----
tab_jjust = table(df_all$journal, df_all$n_just)
chisq_jjust = chisq.test(tab_jjust)

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

## Hypothesis Support (di)
tab_clindisup = table(df_clin$di_sup)
binom_clindisup = binom.test(tab_clindisup[2], sum(tab_clindisup),
                             p = .8)
tab_clinsup = table(df_clin$support)

## Hypothesis Tested
tab_clinhypo = table(df_clin$hypo_tested)
binom_clinhypo = binom.test(tab_clinhypo[2], sum(tab_clinhypo),
                             p = .6)

## Sample Size Just -----
tab_clinjust = table(df_clin$n_just)
binom_clinjust = binom.test(tab_clinjust[2], sum(tab_clinjust))

## Pregreg -----
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)

# RCT breakdown -----

```

```

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

## Hypothesis Support (di)
tab_rctdisup = table(df_rct$di_sup)
binom_rctdisup = binom.test(tab_rctdisup[2], sum(tab_rctdisup),
                             p = .8)
tab_rctsup = table(df_rct$support)

## Hypothesis Tested
tab_rcthyppo = table(df_rct$hyppo_tested)
binom_rcthyppo = binom.test(tab_rcthyppo[2], sum(tab_rcthyppo),
                             p = .6)

## Sample Size Just -----
tab_rctjust = table(df_rct$n_just)
binom_rctjust = binom.test(tab_rctjust[2], sum(tab_rctjust))

tab_rctreg = table(df_rct$prereg)

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

## Sample Size Info -----
# All studies reported sample size information
#tab_clinssj = table(df_clin$sample_info)
#binom_clinssj = binom.test(tab_clinssj)

# Breakdown by Discipline -----

## Hypothesis Tested -----

tab_dissupp = table(df_all$sci_cat, df_all$support)
chisq_dissupp = chisq.test(tab_dissupp)

tab_dishypop = table(df_all$sci_cat, df_all$hyppo_tested)
chisq_dishypop = chisq.test(tab_dishypop)

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")

```

Figure 3

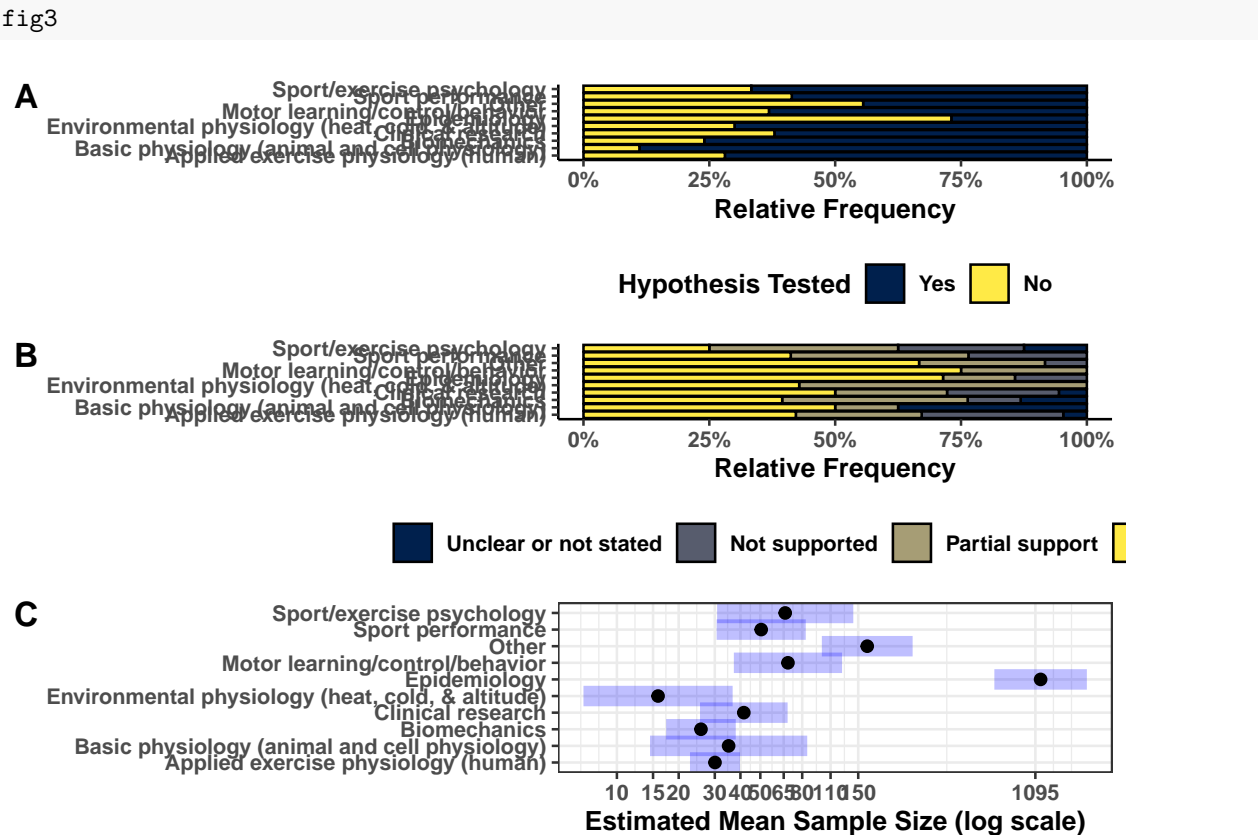


Figure 3. The breakdown, by discipline, for A) indication of whether a hypothesis was tested B) level of reported support for hypotheses, and C) the estimated total sample size (grey bands indicate 95% confidence intervals).

```
# 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 2

fig_2

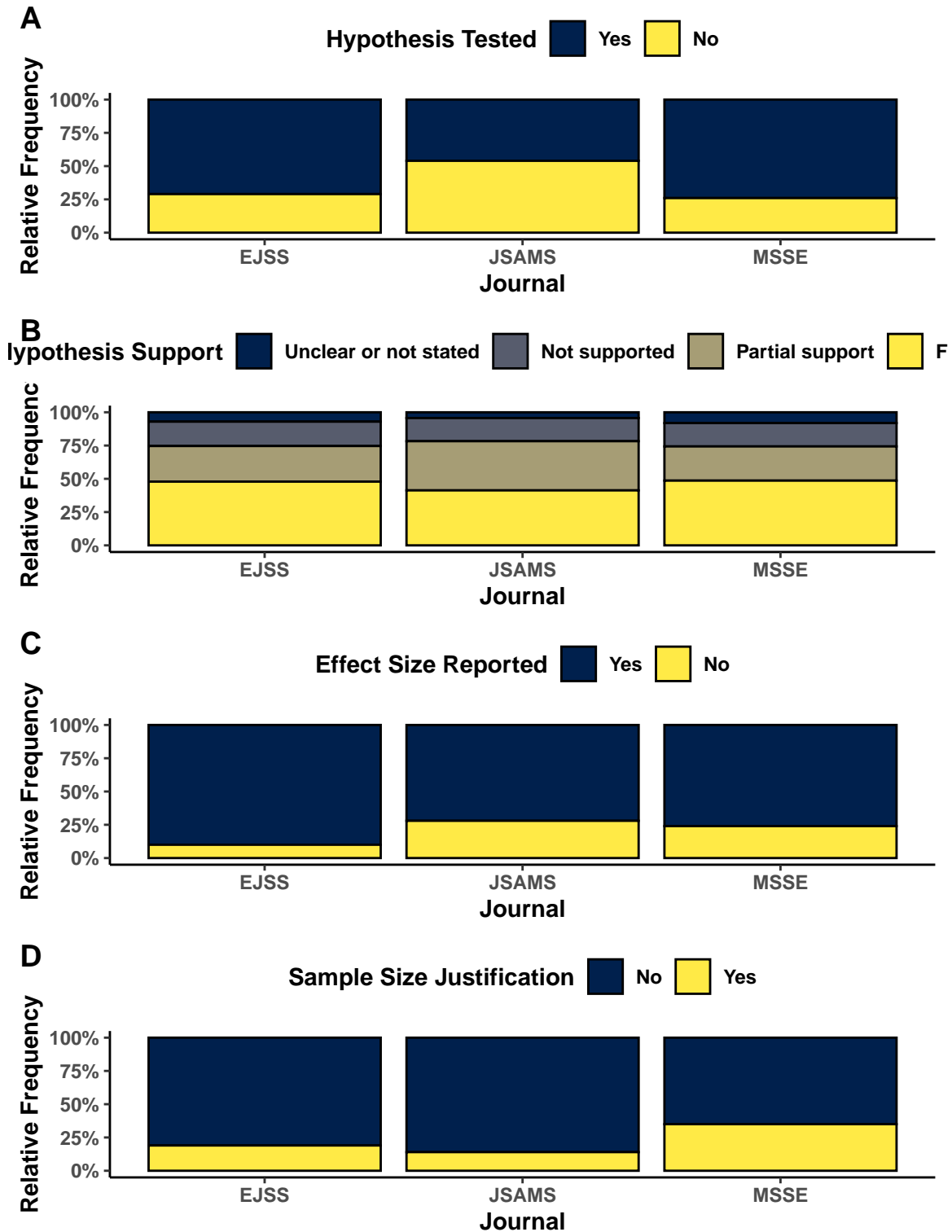


Figure 2. Relative frequencies, by journal, for A) level of reported support for hypotheses, B) indication of whether a hypothesis was tested, C) indication of whether an effect size was reported, or D) indication of if sample size was justified by the authors. Journals included the European Journal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

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>.