How Story Problems Strengthen Arithmetic Problem-Solving Strategies: Evidence from a Learning Trajectory Teaching Experiment in Kindergarten (L & I Revision)

Manuscript ref. JLI-23-00237

Anonymous

2023-10-24

setwd("C:/Users/pcher/Dropbox/TK_PC/AERA Soph over Time")

knitr::opts_chunk$set(echo = T, message=F, warning=F, error=F, comment=NA, cache=T, R.options=list(width=220), fig.align='center', out.width='60%', fig.asp=.75)

Data cleaning and procesing

The code below reads in and provides a preview of the raw coded data:

```
# load necessary R packages
library(dplyr)
library(ggplot2)
library(forcats)
library(hms)
library(lubridate)
library(stringr)
library(readxl)
library(kableExtra)
library(tidyr)
library(viridis)

raw <- read.csv("ArithStrat_allmergeddata_02_28_2022.csv",
                stringsAsFactors = TRUE,na.strings = "NA")
glimpse(raw)
```

```
## Rows: 12,574
## Columns: 23
## $ IRR           <fct> No, No, No, No, No, No, No, No, No, No, No, No, No, No,~
## $ PrimerCoder   <fct> VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC, VC,~
## $ CLASS         <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ CHILD_ID      <dbl> 9531101016, 9531101016, 9531101016, 9531101016, 9531101~
## $ SESSION       <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2~
## $ GRA           <fct> Jared, Jared, Jared, Jared, Jared, Jared, Jared, Jared,~
## $ WARMUP        <fct> No, No, No, No, No, No, No, No, No, No, No, No, No, No,~
## $ PROB_CAT      <fct> Start Unknown, Start Unknown, Start Unknown, Start Unkn~
## $ PROB_TYPE     <fct> Join, Join, Join, Join, Join, Join, Join, Join, Join, J~
## $ Count_LT      <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
## $ Arith_LT      <fct> Counting Strategies, Counting Strategies, Counting Stra~
```

```
## $ Equation_old   <fct> n+6=11, n+6=11, n+6=11, n+6=11, n+5=10, n+5=10, n+5=10,~
## $ Equation       <fct> N+6=11, N+6=11, N+6=11, N+6=11, N+5=10, N+5=10, N+5=10,~
## $ Event          <int> 1, 1, 1, 1, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 5, 5, 5, 5~
## $ Attempt        <int> 1, 2, 3, 4, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 1, 2, 3, 4, 5~
## $ TimeStamp.ENDS <fct> 1:28, 1:50, 2:29, 3:00, 3:19, 3:30, 3:40, 4:30, 4:53, 5~
## $ MathSupp       <fct> Manipulatives 3, Arith Scaffold 1, Arith Scaffold 1, Ar~
## $ Dynamic        <fct> You Do (Student), You Do (Student), You Do (Student), Y~
## $ CountLevel     <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
## $ ArithStrat     <fct> Reasonable Guess, Makes a set, Makes a set, Combination~
## $ Correctness    <fct> Incorrect, Correct with Support, Correct with Support, ~
## $ PrevStrat      <fct> NA, Reasonable Guess, Makes a set, Makes a set, Combina~
## $ file           <fct> 9531101016_session02_VC.xlsx, 9531101016_session02_VC.x~
```

The code below processes the raw coded data, applies exclusion criteria described in the manuscript, and collapses over levels of categorical variables with very little coded data to make the analytic data set used in the article.

```
# pre-processing
raw$GRA<-raw$GRA %>% toupper()
raw$Dynamic<-raw$Dynamic %>% toupper()
raw$Equation <- raw$Equation %>% str_trim()
# turn time stamps of each attempt into a formal R time object
raw$Time_Stamp<-raw$TimeStamp %>%
  strptime(format="%M:%S") %>%
  as_hms()

# make analytic dataset
D_arith<-raw %>%
  filter(Attempt < 11) %>%       #  filter attempts >10
  filter(!is.na(ArithStrat)) %>% #  filter NAs to INLCUDE ONLY Arithmetic events
  filter(ArithStrat != "Other") %>%    #  remove where Strategy = Other
  filter(ArithStrat != "Not Observed") %>% # remove where Strategy = Not Observed

  # remove problem categories and types not relevant to story problems
  filter(!(PROB_CAT %in%c("Counting","Composing Number","Larger Unknown",
                         "Join","Separate","Part-Part-Whole",
                         "Number Comparison","Smaller Unknown"))) %>%
  filter(!(PROB_TYPE %in%c(Prob_Type = "Composing Number", "Equalize","Compare",
                          "Counting", "Number Comparison"))) %>%

  # assign order to ArithStrats based on sophistication, define new variable Y
  mutate(Y = ordered(ArithStrat,
                     levels=c("Wild Guess", "Reasonable Guess", "Trial & Error",

  # make Time_Stamp as numeric in number of seconds ###
  mutate(time_sec=as.numeric(seconds(Time_Stamp))) %>%

  # collapse levels of Y together into new variable YC (Y Collapsed)
  mutate(YC = fct_collapse(Y,
                          "Wild Guess" = "Wild Guess",
                          "RG & TE" = c("Reasonable Guess","Trial & Error"),
                          "Makes a set" = "Makes a set",
                          "Counting All" = "Counting All",
                          "Counting On" = c("Counting On - Concrete",
                                            "Counting On - Abstract"),
```

```r
                              "Jump Strategy" = "Jump Strategy",
                              "Combination" = "Combination",
                              "Derived Combination" = "Derived Combination",
                              "Comp & Decomp" = c("Compensation",
                                                  "Decomposition"))) %>%

  # collapse problem categories & sets reference levels #
  mutate(PROB_CAT_C =fct_collapse(PROB_CAT,
                                    "Find Difference"=c("Difference Unknown",
                                                        "Find Difference"),
                                    "Result Unknown"="Result Unknown",
                                    "Start Unknown"="Start Unknown") %>%
                    relevel(ref="Result Unknown") )

# cast variables below as factors to be included as random intercepts
D_arith$CHILD_ID<-as.factor(D_arith$CHILD_ID)
D_arith$Equation<-as.factor(D_arith$Equation)
D_arith$GRA<-as.factor(D_arith$GRA)
```

Check analytic dataset for errors:

```r
table(D_arith$YC,useNA="always") # SHOULD BE 9 STRATEGIES WITH NO NAs
```

```
##
##          Wild Guess             RG & TE         Makes a set        Counting All
##                 169                 749                 727                2765
##         Counting On       Jump Strategy         Combination Derived Combination
##                2416                 179                1250                 306
##       Comp & Decomp                <NA>
##                 294                   0
```

```r
D_arith$YC %>% levels()          # CHECK THE ORDERING OF STRATEGIES HERE
```

```
## [1] "Wild Guess"          "RG & TE"             "Makes a set"
## [4] "Counting All"        "Counting On"         "Jump Strategy"
## [7] "Combination"         "Derived Combination" "Comp & Decomp"
```

```r
table(D_arith$PROB_TYPE,useNA="always") # SHOULD BE NO COUNTING PROB_TYPES
```

```
##
##        Cardinality            Compare   Composing Number           Counting
##                  0                  0                  0                  0
##           Equalize               Join  Number Comparison Number Composition
##                  0               5505                  0                  0
##     Part-Part-Whole           Separate         Subitizing               <NA>
##                621               2729                  0                  0
```

```r
table(D_arith$PROB_CAT_C,useNA="always") # SHOULD BE NO COUNTING PROB_CATS
```

```
##
##      Result Unknown        Cardinality   Composing Number           Counting
##                3239                  0                  0                  0
##     Find Difference               Join     Larger Unknown  Number Comparison
##                3930                  0                  0                  0
## Number Composition    Part-Part-Whole           Separate     Smaller Unknown
##                  0                  0                  0                  0
##       Start Unknown         Subitizing               <NA>
```

```
##               1686                  0                   0
```

```r
summary(D_arith$time_sec)        # SHOULD PRODUCE *NUMERIC* SUMMARY with NO NAs
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    20.0   262.0   485.0   502.8   730.0  1577.0
```

Selected tables shown in article and supplement

The code below produces several selected tables shown in the article and the supplement:

```r
# table of strategies
D_arith %>% count(YC) %>% mutate(pct=n/nrow(D_arith))
```

```
##                     YC    n        pct
## 1          Wild Guess  169 0.01908526
## 2             RG & TE  749 0.08458498
## 3          Makes a set  727 0.08210051
## 4         Counting All 2765 0.31225296
## 5          Counting On 2416 0.27284020
## 6        Jump Strategy  179 0.02021457
## 7          Combination 1250 0.14116318
## 8 Derived Combination  306 0.03455675
## 9         Comp & Decomp  294 0.03320158
```

```r
# problem categories & types
D_arith %>% group_by(PROB_CAT_C,PROB_TYPE) %>%
  summarize(n=n()) %>%
  spread(PROB_CAT_C, n)
```

```
## # A tibble: 3 x 4
##   PROB_TYPE       `Result Unknown` `Find Difference` `Start Unknown`
##   <fct>                      <int>             <int>           <int>
## 1 Join                        2102              2302            1101
## 2 Part-Part-Whole              117               420              84
## 3 Separate                    1020              1208             501
```

```r
# top equations (>30 coded attempts)
D_arith$Equation %>% table() %>%
  as.data.frame() %>%
  arrange(desc(Freq)) %>%
  slice(1:55) %>%
  kbl() %>% kable_classic(full_width=F)
```

```r
# Graduate Research Assistants (GRAs)
D_arith %>% count(GRA) %>%
  mutate(pct=100*(n/nrow(D_arith)) %>% round(3)) %>%
  kbl() %>% kable_classic(full_width=F) #<-first two letters, then numbers to id
```

```r
# total attempts & events per child
D_arith %>% group_by(CHILD_ID, SESSION) %>%
          summarize(tot_event = max(Event),tot_att=sum(Attempt)) %>%
          mutate(avg_att_per_ev = round(tot_att/tot_event,2)) %>%
          filter(SESSION < 4) # print off data for first 4 SESSIONS
```

```
## # A tibble: 113 x 5
## # Groups:   CHILD_ID [40]
```

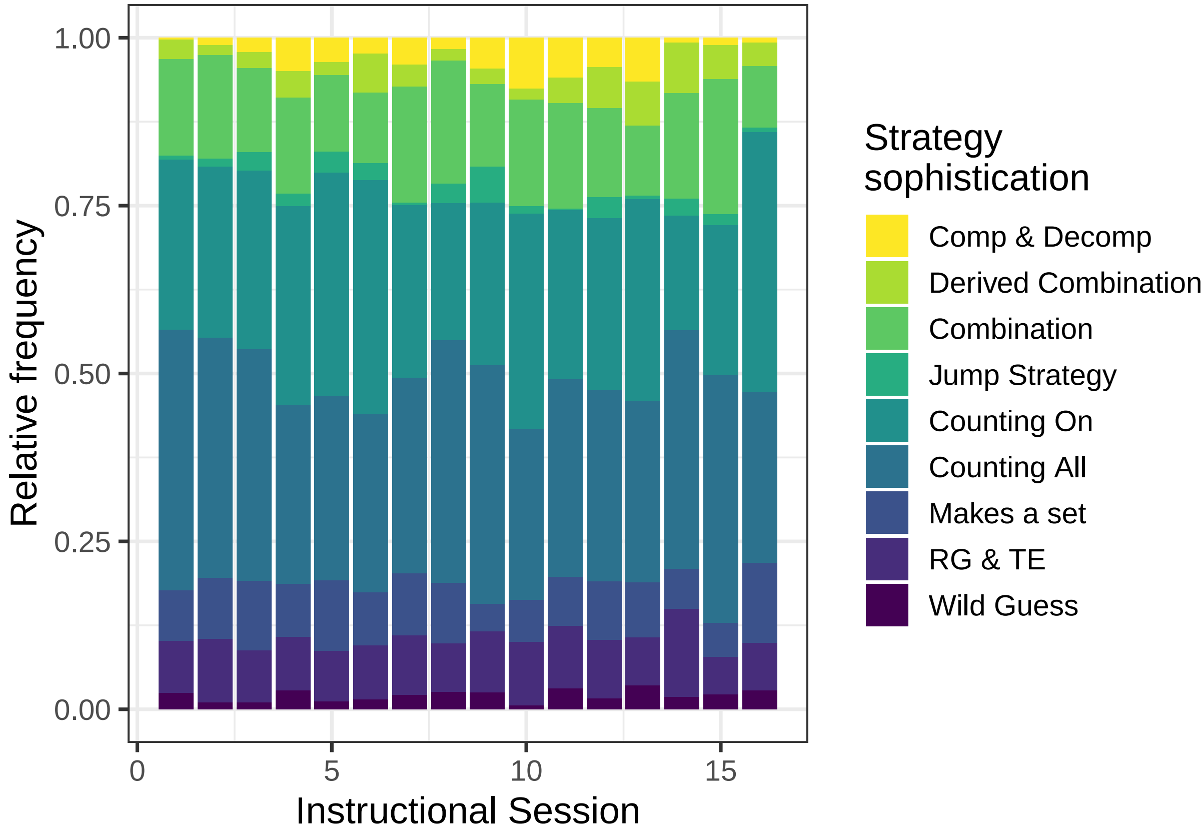
.	Freq
5+N=8	85
N+4=9	84
4+N=6	72
5+N=9	71
6+N=10	68
5+3=N	64
10-N=4	58
2+N=5	58
3+2=N	58
4+N=10	57
3+N=7	55
5-2=N	55
4+3=N	53
7+N=10	53
3+N=5	51
3+4=N	50
7+N=12	50
8+N=10	50
8+N=12	50
2+3=N	49
4+N=7	49
6+N=9	48
10-3=N	47
5+6=N	47
5+N=11	44
5+N=7	44
5+4=N	43
N+5=11	42
4+N=9	40
8+N=15	40
8+N=13	39
10-N=6	38
8+N=14	37
4+2=N	36
5+7=N	36
8+4=N	36
3+N=10	35
N+3=7	35
9+N=15	34
N+4=10	34
N+4=7	34
4+5=N	33
7-3=N	33
7+2=N	33
8+N=11	33
N+6=11	33
8-N=3	32
2+1=N	31
2+N=6	31
3+3=N	31
3+N=8	31
4+6=N	31
6+N=11	31
7+N=11	31
N+2=5	31

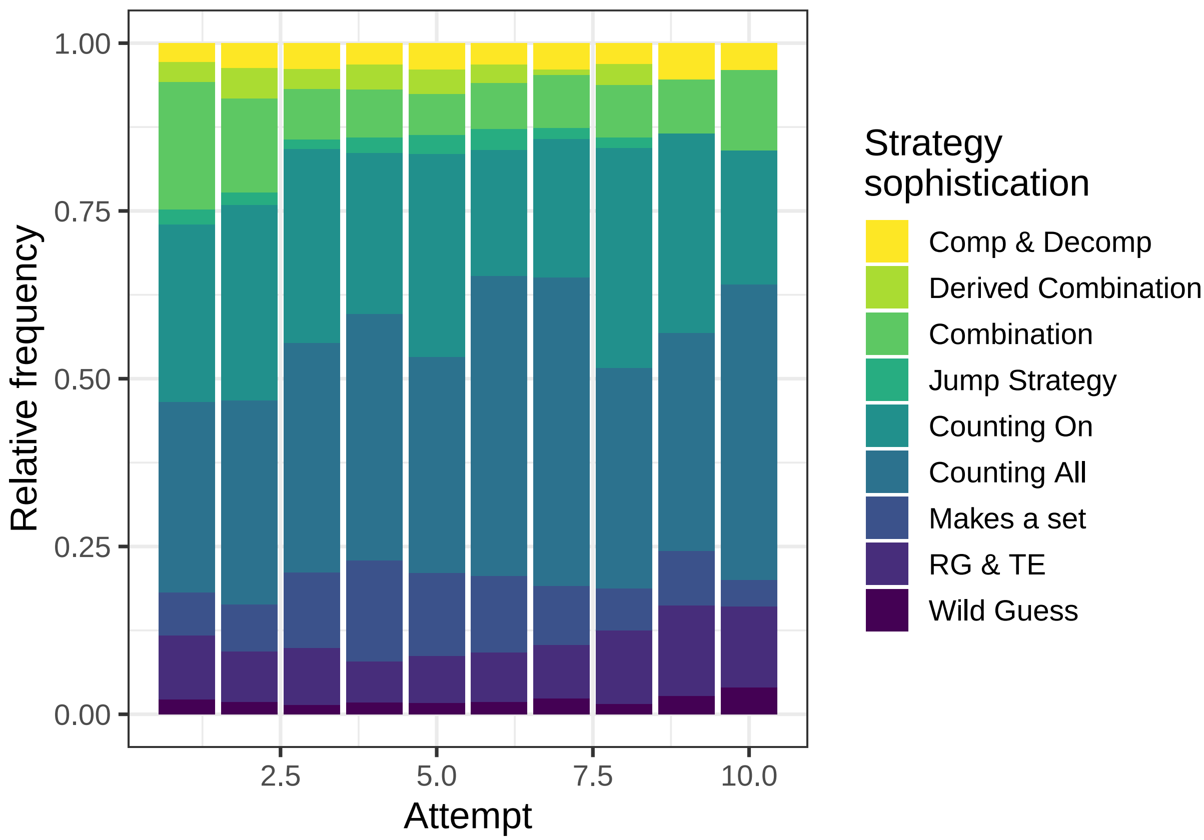
GRA	n	pct
ALI	620	7.0
ARIEL	836	9.4
CANDACE	580	6.5
COURTNEY	935	10.6
ELLEN	533	6.0
ERICA	262	3.0
HOLLAND	67	0.8
JARED	315	3.6
JASON	895	10.1
JCW	239	2.7
JOSE	687	7.8
JULIA	714	8.1
JULIE	22	0.2
LEXI	573	6.5
NEBA	510	5.8
SARA	524	5.9
TARYN	355	4.0
TONI	116	1.3
TRACI	72	0.8

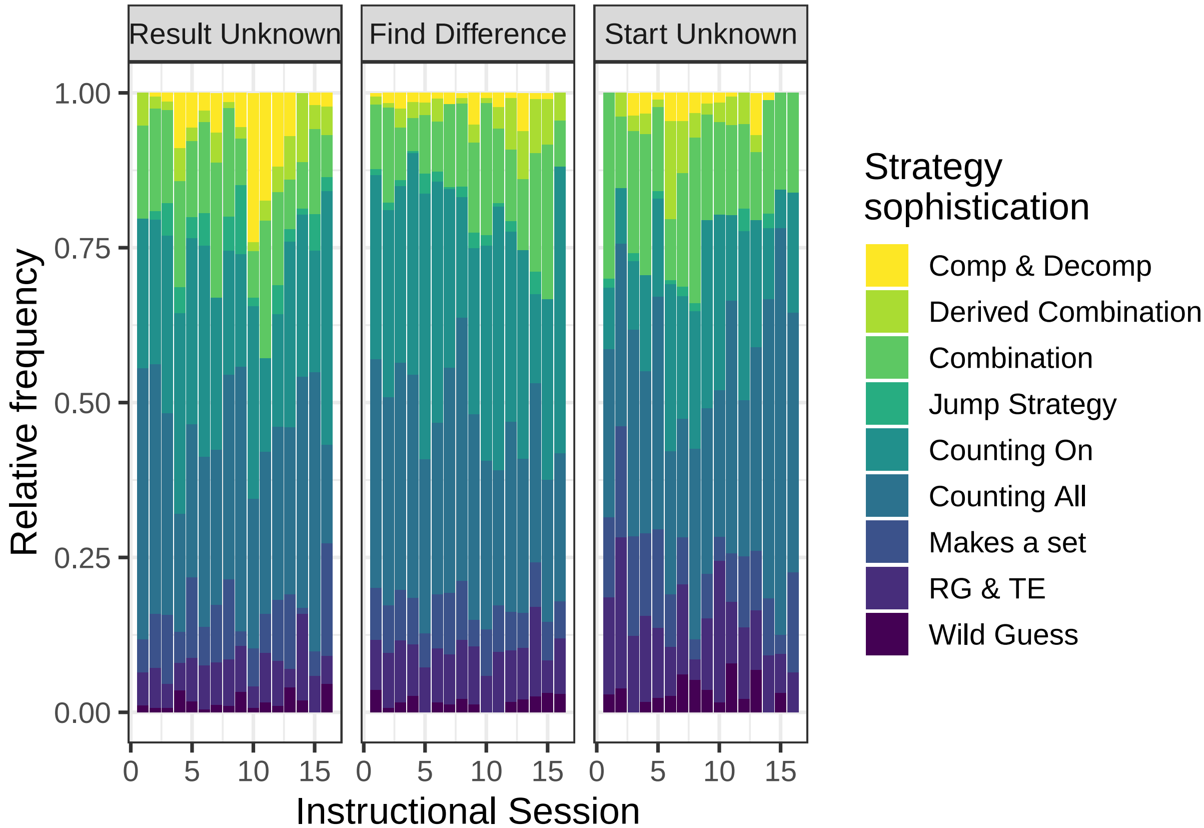
```
##    CHILD_ID   SESSION tot_event tot_att avg_att_per_ev
##    <fct>        <int>     <int>   <int>          <dbl>
##  1 9531101016       1        13      18           1.38
##  2 9531101016       2         6      73          12.2
##  3 9531101016       3         6      63          10.5
##  4 9531101019       1        13      48           3.69
##  5 9531101019       2         4      73          18.2
##  6 9531101019       3        19      66           3.47
##  7 9531101023       1         8      36           4.5
##  8 9531101023       2         4      13           3.25
##  9 9531101023       3        16      38           2.38
## 10 9531102030       1        11      24           2.18
## # i 103 more rows
```

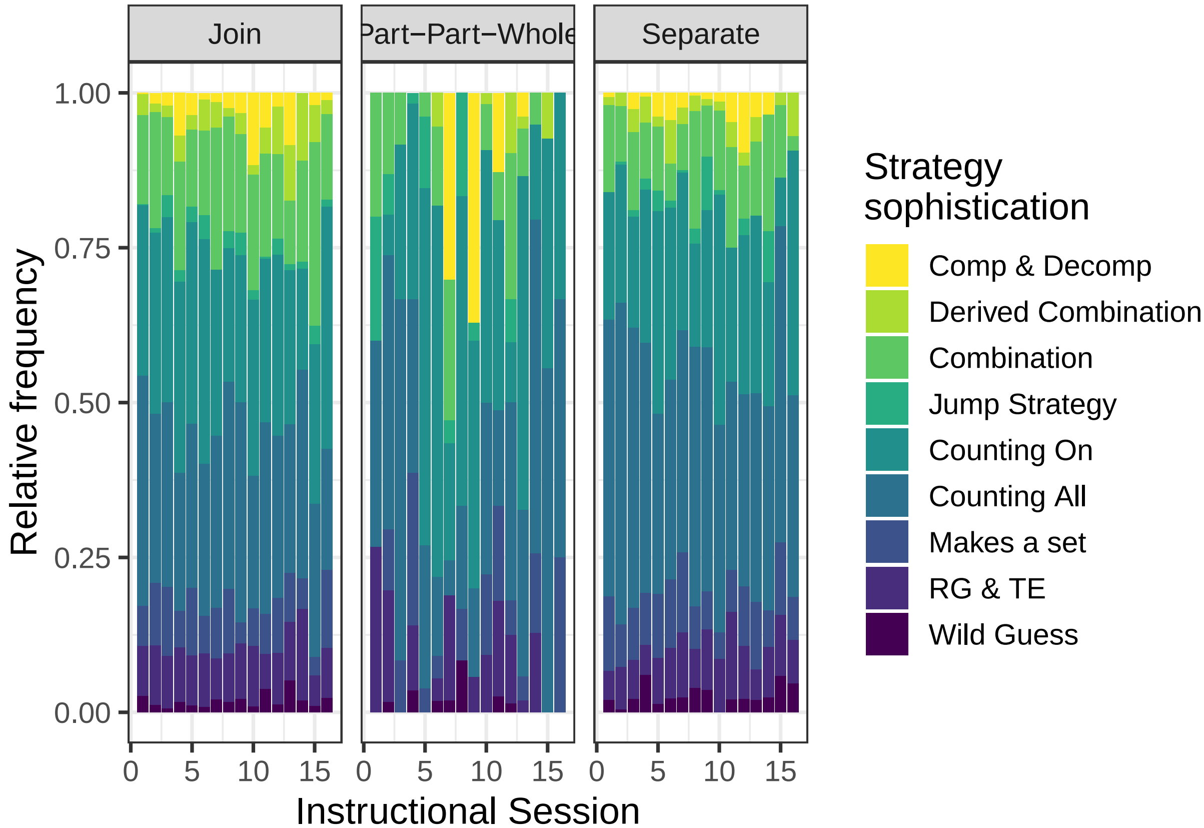
Exploratory plots

Below, we reproduce all figures from the main body of the manuscript and selected figures from the supplement:

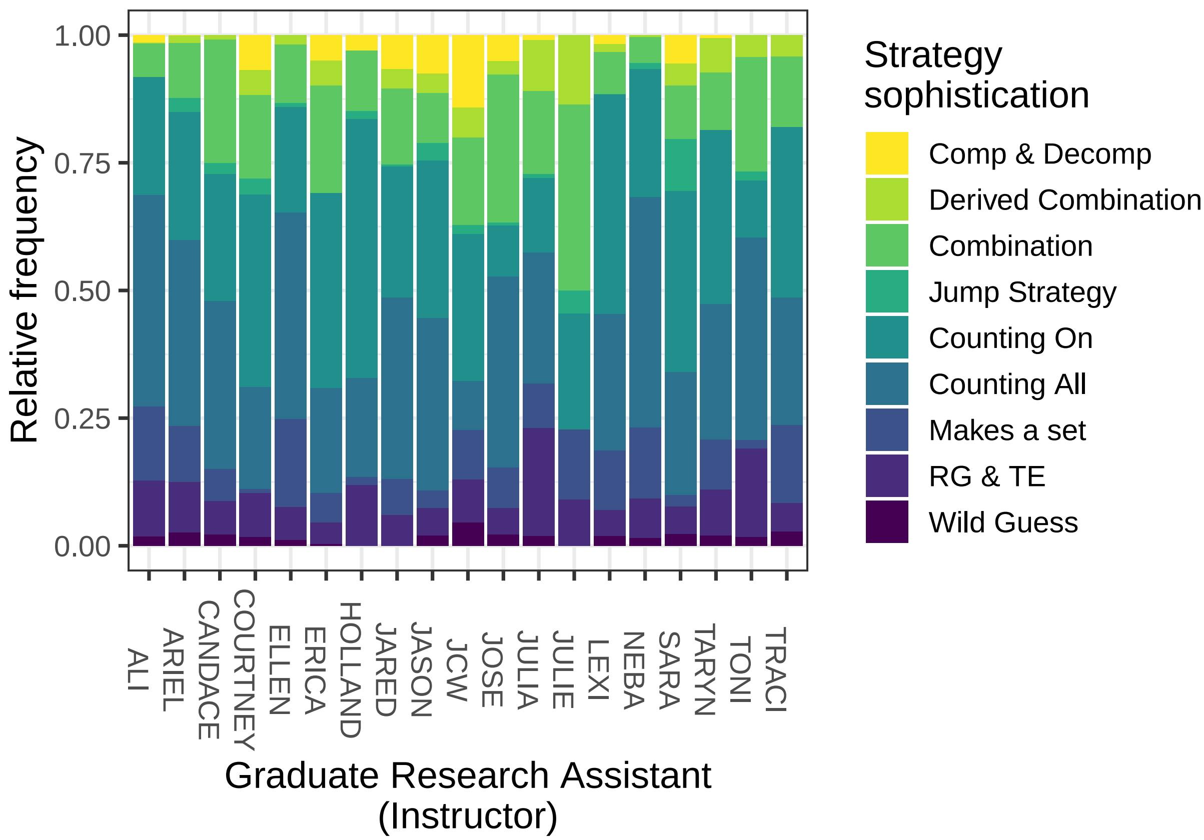



Relative frequency
1.00
0.75
0.50
0.25
0.00
2.5
5.0
7.5
10.0
Attempt
Strategy sophistication
Comp & Decomp
Derived Combination
Combination
Jump Strategy
Counting On
Counting All
Makes a set
RG & TE
Wild Guess


Result Unknown
Find Difference
Start Unknown
Relative frequency
1.00
0.75
0.50
0.25
0.00
0
5
10
15
Instructional Session
Strategy sophistication
Comp & Decomp
Derived Combination
Combination
Jump Strategy
Counting On
Counting All
Makes a set
RG & TE
Wild Guess



Join
Part−Part−Whole
Separate
Relative frequency
1.00
0.75
0.50
0.25
0.00
0
5
10
15
Instructional Session
Strategy sophistication
Comp & Decomp
Derived Combination
Combination
Jump Strategy
Counting On
Counting All
Makes a set
RG & TE
Wild Guess



Bayesian models

Here we specify and estimate longitudinal ordinal sequential logit models. In a previously-published article, Authors et al. (2023) show that category-specific effects are necessary to model arithmetic strategy sophistication in a kindergarten sample.

The model uses random intercepts for children, equations, and GRAs (Instructors), random slopes across session for children, and a Gaussian process for time within session. The Gaussian process is specified with an exponential correlation function exp(-phi/d), where d is the time elapsed in seconds between any two attempts. A model where the range parameter (phi) is estimated from the data failed to converge; thus, we profiled over a range of fixed values for phi to arrive at an optimal value, now held fixed in the model. The results of profiling are shown in the Appendix in Supplementary Table 6.

The Gaussian process for time within session is implemented via a Gaussian process basis. The knots (i.e., times) at which the basis is evaluated are defined on an equally-spaced grid between the 5th and the 95th percentile of time within session (t_knots). No-U-Turn Hamiltonian Monte Carlo (NUTS HMC), implemented in the brms package, is used to sample from the full posterior.

We present three variant of the model presented in the manuscript (m_1, m_1b, and m_1c), which are the baseline model (originally shown in the manuscript), the baseline model with more iterations, and the baseline model with more uncertain prior distributions, respectively. We demonstrate that the estimates presented with the baseline model are robust to increasing the number of MCMC samples and not sensitive to changing the uncertainty in the prior distributions.

```
# Profile over range parameter
rng_phi<-seq(20/3, 420/3,length=25)
rng_phi
```

```
##  [1]   6.666667  12.222222  17.777778  23.333333  28.888889  34.444444
##  [7]  40.000000  45.555556  51.111111  56.666667  62.222222  67.777778
## [13]  73.333333  78.888889  84.444444  90.000000  95.555556 101.111111
## [19] 106.666667 112.222222 117.777778 123.333333 128.888889 134.444444
## [25] 140.000000
```

```
# result of profiling: rng_phi[12]
rng_phi[12] #effective range = 3*rng_phi[12]= 203.33 sec. Any two attempts within 203.33 seconds of one
```

```
## [1] 67.77778
```

```
# Knots (times) to estimate a Gaussian process basis
t_knots<-quantile(D_arith$time_sec,
                  seq(0.05,0.95,by=0.05)) %>% as.numeric()
t_knots
```

```
##  [1]  96 136 178 218 262 306 350 396 440 485 535 582 631 682 730 780 832 886 944
```

```
# Specify and estimate Bayesian model
library(brms)
```

```
## Warning: package 'brms' was built under R version 4.2.3

## Loading required package: Rcpp

## Warning: package 'Rcpp' was built under R version 4.2.3

## Loading 'brms' package (version 2.20.4). Useful instructions
## can be found by typing help('brms'). A more detailed introduction
## to the package is available through vignette('brms_overview').

##
## Attaching package: 'brms'

## The following object is masked from 'package:stats':
##
##     ar
```

```
# formula specification
f_1<-bf(YC ~ cs(scale(SESSION)) + cs(scale(Attempt)) +
          cs(PROB_CAT_C) + cs(PROB_TYPE) +
          (scale(SESSION)|CHILD_ID) +
          (1|Equation) + (1|GRA) +
          s(time_sec,bs='gp',
            m=c(2, rng_phi[12], 1),# powered exp, p = 1 (3rd element) = exp corr
            k=21))                 # number of bases = number of knots + 2

# prior specification
my_priors<-c(set_prior("normal(0, 1.5)", class = "b"),
             set_prior("normal(0, 2.5)", class = "Intercept"),
             set_prior("normal(0, 1.5)", class = "sds"),
             set_prior("lkj(3)", class = "cor"),
             set_prior("normal(0, 1.5)", class = "sd"))

# estimate model (pre-computed) #
# baseline model
m_1<-brm(f_1, data=D_arith, prior=my_priors,
         family=sratio(),
         knots=list(time_sec = t_knots),
```

```
        control = list(adapt_delta=0.8), refresh=50,
        cores=3,iter=1500,warmup=500,chains=3)
```

```
## Warning: Rows containing NAs were excluded from the model.

## Compiling Stan program...

## Start sampling
```

```
m_1
```

```
##  Family: sratio
##   Links: mu = logit; disc = identity
## Formula: YC ~ cs(scale(SESSION)) + cs(scale(Attempt)) + cs(PROB_CAT_C) + cs(PROB_TYPE) + (scale(SESSI
##    Data: D_arith (Number of observations: 8843)
##   Draws: 3 chains, each with iter = 1500; warmup = 500; thin = 1;
##          total post-warmup draws = 3000
##
## Smooth Terms:
##                   Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sds(stime_sec_1)     1.21      0.90     0.05     3.37 1.00     2756     1371
##
## Group-Level Effects:
## ~CHILD_ID (Number of levels: 40)
##                              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)                    0.76      0.09     0.59     0.97 1.00      661
## sd(scaleSESSION)                 0.19      0.04     0.12     0.28 1.00      868
## cor(Intercept,scaleSESSION)      0.03      0.19    -0.35     0.38 1.00     1203
##                              Tail_ESS
## sd(Intercept)                    1228
## sd(scaleSESSION)                 1691
## cor(Intercept,scaleSESSION)      1817
##
## ~Equation (Number of levels: 1255)
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.71      0.04     0.64     0.78 1.00      646     1426
##
## ~GRA (Number of levels: 19)
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.41      0.11     0.25     0.67 1.00      680     1275
##
## Population-Level Effects:
##                              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## Intercept[1]                    -4.90      0.24    -5.38    -4.44 1.01      544
## Intercept[2]                    -3.26      0.19    -3.65    -2.88 1.01      381
## Intercept[3]                    -2.89      0.19    -3.26    -2.52 1.01      402
## Intercept[4]                    -0.85      0.18    -1.20    -0.50 1.01      358
## Intercept[5]                     0.29      0.18    -0.06     0.64 1.01      360
## Intercept[6]                    -1.52      0.21    -1.93    -1.10 1.00      434
## Intercept[7]                     1.39      0.20     0.99     1.79 1.01      408
## Intercept[8]                     0.99      0.24     0.52     1.44 1.01      478
## stime_sec_1                     -0.11      1.44    -2.94     2.66 1.00     5731
## scaleSESSION[1]                  0.05      0.09    -0.11     0.22 1.00     2594
## scaleSESSION[2]                  0.13      0.05     0.03     0.23 1.00     1354
## scaleSESSION[3]                  0.24      0.05     0.14     0.34 1.00     1376
```

```
## scaleSESSION[4]                0.24      0.04     0.16      0.32 1.00      880
## scaleSESSION[5]                0.32      0.05     0.23      0.42 1.00     1280
## scaleSESSION[6]                0.12      0.10    -0.07      0.31 1.00     2655
## scaleSESSION[7]                0.55      0.07     0.41      0.68 1.00     1666
## scaleSESSION[8]                0.24      0.11     0.03      0.46 1.00     2877
## scaleAttempt[1]                0.26      0.08     0.10      0.43 1.00     7525
## scaleAttempt[2]                0.28      0.04     0.19      0.37 1.00     4260
## scaleAttempt[3]               -0.02      0.04    -0.09      0.05 1.00     5498
## scaleAttempt[4]               -0.03      0.03    -0.09      0.03 1.00     5579
## scaleAttempt[5]               -0.07      0.04    -0.15      0.00 1.00     3961
## scaleAttempt[6]               -0.15      0.08    -0.30      0.02 1.00     3831
## scaleAttempt[7]                0.50      0.07     0.37      0.63 1.00     4083
## scaleAttempt[8]                0.24      0.11     0.03      0.45 1.00     4004
## PROB_CAT_CFindDifference[1]   -0.31      0.20    -0.70      0.09 1.00     2547
## PROB_CAT_CFindDifference[2]   -0.67      0.11    -0.89     -0.45 1.00     1905
## PROB_CAT_CFindDifference[3]   -0.19      0.11    -0.41      0.03 1.00     1708
## PROB_CAT_CFindDifference[4]   -0.24      0.08    -0.40     -0.08 1.00     1374
## PROB_CAT_CFindDifference[5]   -0.59      0.09    -0.78     -0.41 1.00     1557
## PROB_CAT_CFindDifference[6]    0.86      0.19     0.50      1.25 1.00     3216
## PROB_CAT_CFindDifference[7]   -0.48      0.14    -0.77     -0.20 1.00     2460
## PROB_CAT_CFindDifference[8]   -0.97      0.23    -1.44     -0.50 1.00     3484
## PROB_CAT_CStartUnknown[1]     -1.37      0.22    -1.80     -0.94 1.00     2204
## PROB_CAT_CStartUnknown[2]     -1.37      0.14    -1.63     -1.11 1.00     1501
## PROB_CAT_CStartUnknown[3]     -0.86      0.14    -1.12     -0.59 1.00     1386
## PROB_CAT_CStartUnknown[4]     -0.69      0.11    -0.91     -0.49 1.00     1390
## PROB_CAT_CStartUnknown[5]     -0.20      0.12    -0.43      0.04 1.00     1851
## PROB_CAT_CStartUnknown[6]      1.20      0.29     0.65      1.79 1.00     4133
## PROB_CAT_CStartUnknown[7]     -0.81      0.16    -1.12     -0.49 1.00     2194
## PROB_CAT_CStartUnknown[8]     -1.44      0.27    -1.97     -0.92 1.00     3364
## PROB_TYPEPartMPartMWhole[1]    0.19      0.39    -0.52      1.01 1.00     3684
## PROB_TYPEPartMPartMWhole[2]   -0.25      0.16    -0.57      0.06 1.00     3836
## PROB_TYPEPartMPartMWhole[3]   -0.29      0.17    -0.61      0.04 1.00     4608
## PROB_TYPEPartMPartMWhole[4]   -0.28      0.12    -0.52     -0.05 1.00     3818
## PROB_TYPEPartMPartMWhole[5]   -0.51      0.14    -0.79     -0.24 1.00     4580
## PROB_TYPEPartMPartMWhole[6]   -1.34      0.30    -1.88     -0.75 1.00     4767
## PROB_TYPEPartMPartMWhole[7]    0.27      0.25    -0.21      0.76 1.00     4705
## PROB_TYPEPartMPartMWhole[8]    1.03      0.37     0.29      1.75 1.00     4576
## PROB_TYPESeparate[1]          -0.31      0.17    -0.64      0.05 1.00     3269
## PROB_TYPESeparate[2]           0.09      0.10    -0.11      0.29 1.00     2608
## PROB_TYPESeparate[3]          -0.12      0.10    -0.32      0.07 1.00     2699
## PROB_TYPESeparate[4]          -0.53      0.07    -0.67     -0.38 1.00     2005
## PROB_TYPESeparate[5]          -0.21      0.09    -0.39     -0.03 1.00     2535
## PROB_TYPESeparate[6]          -0.39      0.20    -0.77      0.02 1.00     3977
## PROB_TYPESeparate[7]           0.12      0.14    -0.16      0.39 1.00     2713
## PROB_TYPESeparate[8]           0.12      0.23    -0.34      0.57 1.00     3656
##                            Tail_ESS
## Intercept[1]                   1000
## Intercept[2]                    782
## Intercept[3]                    722
## Intercept[4]                    881
## Intercept[5]                    850
## Intercept[6]                    924
## Intercept[7]                   1022
## Intercept[8]                   1144
```

```
## stime_sec_1                     2075
## scaleSESSION[1]                 2283
## scaleSESSION[2]                 1888
## scaleSESSION[3]                 2009
## scaleSESSION[4]                 1779
## scaleSESSION[5]                 1685
## scaleSESSION[6]                 2215
## scaleSESSION[7]                 2161
## scaleSESSION[8]                 2152
## scaleAttempt[1]                 2120
## scaleAttempt[2]                 2309
## scaleAttempt[3]                 2056
## scaleAttempt[4]                 2356
## scaleAttempt[5]                 2441
## scaleAttempt[6]                 2015
## scaleAttempt[7]                 2421
## scaleAttempt[8]                 2121
## PROB_CAT_CFindDifference[1]     2165
## PROB_CAT_CFindDifference[2]     2159
## PROB_CAT_CFindDifference[3]     1844
## PROB_CAT_CFindDifference[4]     1817
## PROB_CAT_CFindDifference[5]     1790
## PROB_CAT_CFindDifference[6]     2431
## PROB_CAT_CFindDifference[7]     2187
## PROB_CAT_CFindDifference[8]     2343
## PROB_CAT_CStartUnknown[1]       2262
## PROB_CAT_CStartUnknown[2]       1805
## PROB_CAT_CStartUnknown[3]       1772
## PROB_CAT_CStartUnknown[4]       1874
## PROB_CAT_CStartUnknown[5]       2168
## PROB_CAT_CStartUnknown[6]       1696
## PROB_CAT_CStartUnknown[7]       2425
## PROB_CAT_CStartUnknown[8]       2164
## PROB_TYPEPartMPartMWhole[1]     2038
## PROB_TYPEPartMPartMWhole[2]     2256
## PROB_TYPEPartMPartMWhole[3]     2400
## PROB_TYPEPartMPartMWhole[4]     2298
## PROB_TYPEPartMPartMWhole[5]     2382
## PROB_TYPEPartMPartMWhole[6]     2160
## PROB_TYPEPartMPartMWhole[7]     2451
## PROB_TYPEPartMPartMWhole[8]     2120
## PROB_TYPESeparate[1]            2230
## PROB_TYPESeparate[2]            2282
## PROB_TYPESeparate[3]            2420
## PROB_TYPESeparate[4]            2220
## PROB_TYPESeparate[5]            1798
## PROB_TYPESeparate[6]            2434
## PROB_TYPESeparate[7]            2285
## PROB_TYPESeparate[8]            2162
##
## Family Specific Parameters:
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## disc     1.00      0.00     1.00     1.00   NA       NA       NA
##
```

```
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```r
# increase iterations to investigate convergence (1000 to 4000 post-warmup)
m_1b<-brm(f_1, data=D_arith, prior=my_priors,
          family=sratio(),
          knots=list(time_sec = t_knots),
          control = list(adapt_delta=0.8), refresh=50,
          cores=3,iter=4500,warmup=500,chains=3)
```

```
## Warning: Rows containing NAs were excluded from the model.
```

```
## Compiling Stan program...
## Start sampling
```

```r
m_1b
```

```
##  Family: sratio
##   Links: mu = logit; disc = identity
## Formula: YC ~ cs(scale(SESSION)) + cs(scale(Attempt)) + cs(PROB_CAT_C) + cs(PROB_TYPE) + (scale(SESSI
##    Data: D_arith (Number of observations: 8843)
##   Draws: 3 chains, each with iter = 4500; warmup = 500; thin = 1;
##          total post-warmup draws = 12000
##
## Smooth Terms:
##                 Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sds(stime_sec_1)     1.20      0.91     0.05     3.42 1.00     8725     4816
##
## Group-Level Effects:
## ~CHILD_ID (Number of levels: 40)
##                            Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)                  0.77      0.10     0.60     0.99 1.00     2023
## sd(scaleSESSION)               0.19      0.04     0.12     0.27 1.00     3354
## cor(Intercept,scaleSESSION)    0.02      0.19    -0.35     0.38 1.00     4981
##                            Tail_ESS
## sd(Intercept)                  4004
## sd(scaleSESSION)               5790
## cor(Intercept,scaleSESSION)    7056
##
## ~Equation (Number of levels: 1255)
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.71      0.04     0.64     0.78 1.00     3136     5854
##
## ~GRA (Number of levels: 19)
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.41      0.11     0.24     0.66 1.00     2376     4100
##
## Population-Level Effects:
##                            Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## Intercept[1]                  -4.89      0.23    -5.34    -4.46 1.00     1589
## Intercept[2]                  -3.25      0.18    -3.61    -2.89 1.00     1147
## Intercept[3]                  -2.88      0.18    -3.23    -2.53 1.00     1119
## Intercept[4]                  -0.84      0.17    -1.16    -0.50 1.00     1002
## Intercept[5]                   0.30      0.17    -0.03     0.65 1.00     1049
## Intercept[6]                  -1.51      0.20    -1.90    -1.10 1.00     1406
```

```
## Intercept[7]                   1.41      0.19     1.04      1.78 1.00      1246
## Intercept[8]                   1.00      0.22     0.57      1.45 1.00      1657
## stime_sec_1                   -0.09      1.50    -3.06      2.85 1.00     21961
## scaleSESSION[1]                0.05      0.08    -0.11      0.22 1.00      7648
## scaleSESSION[2]                0.13      0.05     0.03      0.23 1.00      4949
## scaleSESSION[3]                0.24      0.05     0.13      0.34 1.00      5062
## scaleSESSION[4]                0.24      0.04     0.15      0.32 1.00      3713
## scaleSESSION[5]                0.32      0.05     0.23      0.42 1.00      4335
## scaleSESSION[6]                0.12      0.09    -0.05      0.30 1.00      9646
## scaleSESSION[7]                0.55      0.07     0.42      0.69 1.00      6237
## scaleSESSION[8]                0.25      0.11     0.02      0.47 1.00      9257
## scaleAttempt[1]                0.26      0.08     0.10      0.42 1.00     18388
## scaleAttempt[2]                0.28      0.04     0.19      0.36 1.00     15775
## scaleAttempt[3]               -0.02      0.04    -0.09      0.06 1.00     19725
## scaleAttempt[4]               -0.03      0.03    -0.09      0.02 1.00     15369
## scaleAttempt[5]               -0.07      0.04    -0.15      0.00 1.00     15767
## scaleAttempt[6]               -0.15      0.08    -0.31      0.02 1.00     19667
## scaleAttempt[7]                0.50      0.07     0.37      0.63 1.00     16656
## scaleAttempt[8]                0.24      0.11     0.03      0.45 1.00     15571
## PROB_CAT_CFindDifference[1]   -0.31      0.20    -0.72      0.08 1.00      6357
## PROB_CAT_CFindDifference[2]   -0.67      0.11    -0.90     -0.45 1.00      5334
## PROB_CAT_CFindDifference[3]   -0.19      0.11    -0.41      0.03 1.00      5286
## PROB_CAT_CFindDifference[4]   -0.24      0.08    -0.40     -0.08 1.00      3448
## PROB_CAT_CFindDifference[5]   -0.59      0.09    -0.78     -0.41 1.00      4452
## PROB_CAT_CFindDifference[6]    0.86      0.20     0.47      1.25 1.00      8639
## PROB_CAT_CFindDifference[7]   -0.48      0.14    -0.75     -0.20 1.00      8048
## PROB_CAT_CFindDifference[8]   -0.97      0.23    -1.42     -0.51 1.00     10152
## PROB_CAT_CStartUnknown[1]     -1.38      0.21    -1.80     -0.97 1.00      7383
## PROB_CAT_CStartUnknown[2]     -1.37      0.13    -1.64     -1.11 1.00      6415
## PROB_CAT_CStartUnknown[3]     -0.86      0.14    -1.13     -0.60 1.00      6888
## PROB_CAT_CStartUnknown[4]     -0.70      0.11    -0.90     -0.49 1.00      5484
## PROB_CAT_CStartUnknown[5]     -0.20      0.12    -0.43      0.03 1.00      6524
## PROB_CAT_CStartUnknown[6]      1.21      0.30     0.65      1.82 1.00     12306
## PROB_CAT_CStartUnknown[7]     -0.81      0.16    -1.12     -0.49 1.00      9194
## PROB_CAT_CStartUnknown[8]     -1.44      0.27    -1.96     -0.92 1.00     11402
## PROB_TYPEPartMPartMWhole[1]    0.18      0.37    -0.50      0.96 1.00     17876
## PROB_TYPEPartMPartMWhole[2]   -0.25      0.16    -0.56      0.07 1.00     15141
## PROB_TYPEPartMPartMWhole[3]   -0.29      0.17    -0.62      0.05 1.00     14270
## PROB_TYPEPartMPartMWhole[4]   -0.28      0.12    -0.51     -0.05 1.00     13635
## PROB_TYPEPartMPartMWhole[5]   -0.51      0.14    -0.79     -0.22 1.00     14846
## PROB_TYPEPartMPartMWhole[6]   -1.33      0.29    -1.90     -0.75 1.00     12352
## PROB_TYPEPartMPartMWhole[7]    0.27      0.24    -0.20      0.75 1.00     14745
## PROB_TYPEPartMPartMWhole[8]    1.03      0.37     0.30      1.77 1.00     13560
## PROB_TYPESeparate[1]          -0.30      0.17    -0.65      0.04 1.00     12464
## PROB_TYPESeparate[2]           0.09      0.10    -0.11      0.29 1.00      7881
## PROB_TYPESeparate[3]          -0.12      0.10    -0.32      0.08 1.00      7930
## PROB_TYPESeparate[4]          -0.52      0.08    -0.67     -0.37 1.00      5414
## PROB_TYPESeparate[5]          -0.21      0.09    -0.38     -0.03 1.00      7045
## PROB_TYPESeparate[6]          -0.38      0.19    -0.76     -0.00 1.00      9713
## PROB_TYPESeparate[7]           0.12      0.14    -0.16      0.40 1.00     10856
## PROB_TYPESeparate[8]           0.12      0.23    -0.32      0.57 1.00     11679
##                             Tail_ESS
## Intercept[1]                    3721
## Intercept[2]                    2318
```

```
## Intercept[3]                  2335
## Intercept[4]                  2147
## Intercept[5]                  2212
## Intercept[6]                  3409
## Intercept[7]                  2708
## Intercept[8]                  3881
## stime_sec_1                   8076
## scaleSESSION[1]               8464
## scaleSESSION[2]               7442
## scaleSESSION[3]               8303
## scaleSESSION[4]               6838
## scaleSESSION[5]               6632
## scaleSESSION[6]               7896
## scaleSESSION[7]               7303
## scaleSESSION[8]               8497
## scaleAttempt[1]               9039
## scaleAttempt[2]               8603
## scaleAttempt[3]               9472
## scaleAttempt[4]               8337
## scaleAttempt[5]               9400
## scaleAttempt[6]               9507
## scaleAttempt[7]               8655
## scaleAttempt[8]               9133
## PROB_CAT_CFindDifference[1]   8108
## PROB_CAT_CFindDifference[2]   8043
## PROB_CAT_CFindDifference[3]   7531
## PROB_CAT_CFindDifference[4]   7407
## PROB_CAT_CFindDifference[5]   8129
## PROB_CAT_CFindDifference[6]   9338
## PROB_CAT_CFindDifference[7]   9091
## PROB_CAT_CFindDifference[8]   9509
## PROB_CAT_CStartUnknown[1]     9106
## PROB_CAT_CStartUnknown[2]     8138
## PROB_CAT_CStartUnknown[3]     8732
## PROB_CAT_CStartUnknown[4]     7829
## PROB_CAT_CStartUnknown[5]     8368
## PROB_CAT_CStartUnknown[6]     7331
## PROB_CAT_CStartUnknown[7]     8313
## PROB_CAT_CStartUnknown[8]     9039
## PROB_TYPEPartMPartMWhole[1]   8476
## PROB_TYPEPartMPartMWhole[2]   8468
## PROB_TYPEPartMPartMWhole[3]   8697
## PROB_TYPEPartMPartMWhole[4]   9335
## PROB_TYPEPartMPartMWhole[5]   9316
## PROB_TYPEPartMPartMWhole[6]   8525
## PROB_TYPEPartMPartMWhole[7]   9348
## PROB_TYPEPartMPartMWhole[8]   9101
## PROB_TYPESeparate[1]          8932
## PROB_TYPESeparate[2]          8399
## PROB_TYPESeparate[3]          8433
## PROB_TYPESeparate[4]          7229
## PROB_TYPESeparate[5]          8698
## PROB_TYPESeparate[6]          9172
## PROB_TYPESeparate[7]          9565
```

```
## PROB_TYPESeparate[8]            8778
##
## Family Specific Parameters:
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## disc     1.00      0.00     1.00     1.00   NA       NA       NA
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
# increase prior uncertanity to investigate sensitivity to priors
# prior specification
my_priors_wider<-c(set_prior("normal(0, 3.0)", class = "b"),
             set_prior("normal(0, 5.0)", class = "Intercept"),
             set_prior("normal(0, 3.0)", class = "sds"),
             set_prior("lkj(1)", class = "cor"),
             set_prior("normal(0, 3.0)", class = "sd"))
m_1c<-brm(f_1, data=D_arith, prior=my_priors_wider,
          family=sratio(),
          knots=list(time_sec = t_knots),
          control = list(adapt_delta=0.8), refresh=50,
          cores=3,iter=1500,warmup=500,chains=3)
```

```
## Warning: Rows containing NAs were excluded from the model.

## Compiling Stan program...
## Start sampling

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be
## Running the chains for more iterations may help. See
## https://mc-stan.org/misc/warnings.html#bulk-ess
```

```
m_1c
```

```
##  Family: sratio
##   Links: mu = logit; disc = identity
## Formula: YC ~ cs(scale(SESSION)) + cs(scale(Attempt)) + cs(PROB_CAT_C) + cs(PROB_TYPE) + (scale(SESSI
##    Data: D_arith (Number of observations: 8843)
##   Draws: 3 chains, each with iter = 1500; warmup = 500; thin = 1;
##          total post-warmup draws = 3000
##
## Smooth Terms:
##                   Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sds(stime_sec_1)     2.45      1.92     0.09     6.97 1.00     2269     1339
##
## Group-Level Effects:
## ~CHILD_ID (Number of levels: 40)
##                              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)                    0.78      0.10     0.61     0.99 1.01      469
## sd(scaleSESSION)                 0.19      0.04     0.12     0.27 1.00      778
## cor(Intercept,scaleSESSION)      0.03      0.21    -0.38     0.42 1.00     1171
##                              Tail_ESS
## sd(Intercept)                    1135
## sd(scaleSESSION)                 1303
## cor(Intercept,scaleSESSION)      1714
##
## ~Equation (Number of levels: 1255)
```

```
##                Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.71      0.04     0.64     0.78 1.00      645     1631
##
## ~GRA (Number of levels: 19)
##                Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.41      0.11     0.24     0.67 1.01      437      662
##
## Population-Level Effects:
##                              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## Intercept[1]                    -4.96      0.23    -5.44    -4.53 1.01      361
## Intercept[2]                    -3.30      0.19    -3.69    -2.96 1.02      270
## Intercept[3]                    -2.93      0.19    -3.32    -2.58 1.02      271
## Intercept[4]                    -0.88      0.17    -1.24    -0.55 1.02      255
## Intercept[5]                     0.26      0.18    -0.09     0.59 1.02      263
## Intercept[6]                    -1.54      0.21    -1.96    -1.15 1.01      352
## Intercept[7]                     1.36      0.20     0.97     1.74 1.01      328
## Intercept[8]                     0.95      0.23     0.46     1.40 1.01      396
## stime_sec_1                     -0.36      2.99    -6.13     5.55 1.00     4723
## scaleSESSION[1]                  0.05      0.09    -0.11     0.23 1.00     2044
## scaleSESSION[2]                  0.13      0.05     0.03     0.24 1.00     1281
## scaleSESSION[3]                  0.24      0.05     0.13     0.35 1.00     1338
## scaleSESSION[4]                  0.24      0.04     0.15     0.32 1.00     1028
## scaleSESSION[5]                  0.32      0.05     0.22     0.41 1.00     1224
## scaleSESSION[6]                  0.12      0.09    -0.06     0.30 1.00     2378
## scaleSESSION[7]                  0.55      0.07     0.41     0.68 1.00     1684
## scaleSESSION[8]                  0.24      0.11     0.03     0.46 1.00     2488
## scaleAttempt[1]                  0.26      0.09     0.09     0.43 1.00     3423
## scaleAttempt[2]                  0.28      0.04     0.19     0.36 1.00     3280
## scaleAttempt[3]                 -0.02      0.04    -0.09     0.05 1.00     4779
## scaleAttempt[4]                 -0.03      0.03    -0.09     0.02 1.00     3000
## scaleAttempt[5]                 -0.07      0.04    -0.15     0.00 1.00     2931
## scaleAttempt[6]                 -0.15      0.08    -0.31     0.01 1.00     3195
## scaleAttempt[7]                  0.50      0.06     0.37     0.62 1.00     3852
## scaleAttempt[8]                  0.24      0.11     0.03     0.45 1.00     3258
## PROB_CAT_CFindDifference[1]     -0.34      0.21    -0.75     0.07 1.00     2024
## PROB_CAT_CFindDifference[2]     -0.69      0.12    -0.92    -0.46 1.00     1647
## PROB_CAT_CFindDifference[3]     -0.20      0.11    -0.42     0.01 1.00     1493
## PROB_CAT_CFindDifference[4]     -0.24      0.08    -0.40    -0.09 1.01     1091
## PROB_CAT_CFindDifference[5]     -0.60      0.09    -0.78    -0.41 1.00     1333
## PROB_CAT_CFindDifference[6]      0.87      0.20     0.49     1.27 1.00     2694
## PROB_CAT_CFindDifference[7]     -0.49      0.15    -0.77    -0.21 1.00     1943
## PROB_CAT_CFindDifference[8]     -1.00      0.23    -1.47    -0.56 1.00     2689
## PROB_CAT_CStartUnknown[1]       -1.42      0.22    -1.85    -0.99 1.00     1922
## PROB_CAT_CStartUnknown[2]       -1.40      0.13    -1.66    -1.14 1.00     1393
## PROB_CAT_CStartUnknown[3]       -0.88      0.13    -1.14    -0.61 1.00     1340
## PROB_CAT_CStartUnknown[4]       -0.71      0.10    -0.91    -0.50 1.01     1191
## PROB_CAT_CStartUnknown[5]       -0.20      0.12    -0.44     0.03 1.00     1542
## PROB_CAT_CStartUnknown[6]        1.24      0.30     0.66     1.85 1.00     2860
## PROB_CAT_CStartUnknown[7]       -0.82      0.16    -1.16    -0.52 1.00     2180
## PROB_CAT_CStartUnknown[8]       -1.48      0.27    -2.04    -0.93 1.00     3007
## PROB_TYPEPartMPartMWhole[1]      0.18      0.38    -0.53     0.98 1.00     3941
## PROB_TYPEPartMPartMWhole[2]     -0.25      0.16    -0.57     0.06 1.00     3130
## PROB_TYPEPartMPartMWhole[3]     -0.29      0.17    -0.63     0.05 1.00     2772
## PROB_TYPEPartMPartMWhole[4]     -0.28      0.12    -0.52    -0.06 1.00     2634
```

```
## PROB_TYPEPartMPartMWhole[5]    -0.51      0.14    -0.79    -0.25 1.00     4316
## PROB_TYPEPartMPartMWhole[6]    -1.37      0.29    -1.93    -0.80 1.00     3737
## PROB_TYPEPartMPartMWhole[7]     0.28      0.24    -0.20     0.75 1.00     3797
## PROB_TYPEPartMPartMWhole[8]     1.09      0.39     0.36     1.86 1.00     3669
## PROB_TYPESeparate[1]           -0.31      0.18    -0.64     0.03 1.00     3134
## PROB_TYPESeparate[2]            0.09      0.11    -0.12     0.29 1.00     2197
## PROB_TYPESeparate[3]           -0.12      0.10    -0.32     0.07 1.00     1781
## PROB_TYPESeparate[4]           -0.53      0.08    -0.67    -0.37 1.00     1495
## PROB_TYPESeparate[5]           -0.21      0.09    -0.40    -0.03 1.00     1904
## PROB_TYPESeparate[6]           -0.39      0.19    -0.77    -0.01 1.00     2879
## PROB_TYPESeparate[7]            0.12      0.15    -0.17     0.41 1.00     2749
## PROB_TYPESeparate[8]            0.13      0.23    -0.31     0.56 1.00     2952
##                              Tail_ESS
## Intercept[1]                      622
## Intercept[2]                      488
## Intercept[3]                      392
## Intercept[4]                      327
## Intercept[5]                      333
## Intercept[6]                      813
## Intercept[7]                      516
## Intercept[8]                      665
## stime_sec_1                      2078
## scaleSESSION[1]                  2108
## scaleSESSION[2]                  1554
## scaleSESSION[3]                  1730
## scaleSESSION[4]                  1789
## scaleSESSION[5]                  1948
## scaleSESSION[6]                  1776
## scaleSESSION[7]                  1773
## scaleSESSION[8]                  2354
## scaleAttempt[1]                  2246
## scaleAttempt[2]                  2201
## scaleAttempt[3]                  2187
## scaleAttempt[4]                  2388
## scaleAttempt[5]                  2014
## scaleAttempt[6]                  2348
## scaleAttempt[7]                  2382
## scaleAttempt[8]                  2195
## PROB_CAT_CFindDifference[1]      2223
## PROB_CAT_CFindDifference[2]      2329
## PROB_CAT_CFindDifference[3]      2140
## PROB_CAT_CFindDifference[4]      1979
## PROB_CAT_CFindDifference[5]      1856
## PROB_CAT_CFindDifference[6]      2200
## PROB_CAT_CFindDifference[7]      2027
## PROB_CAT_CFindDifference[8]      2218
## PROB_CAT_CStartUnknown[1]        2054
## PROB_CAT_CStartUnknown[2]        2180
## PROB_CAT_CStartUnknown[3]        1741
## PROB_CAT_CStartUnknown[4]        2041
## PROB_CAT_CStartUnknown[5]        2173
## PROB_CAT_CStartUnknown[6]        2281
## PROB_CAT_CStartUnknown[7]        2062
## PROB_CAT_CStartUnknown[8]        2185
```

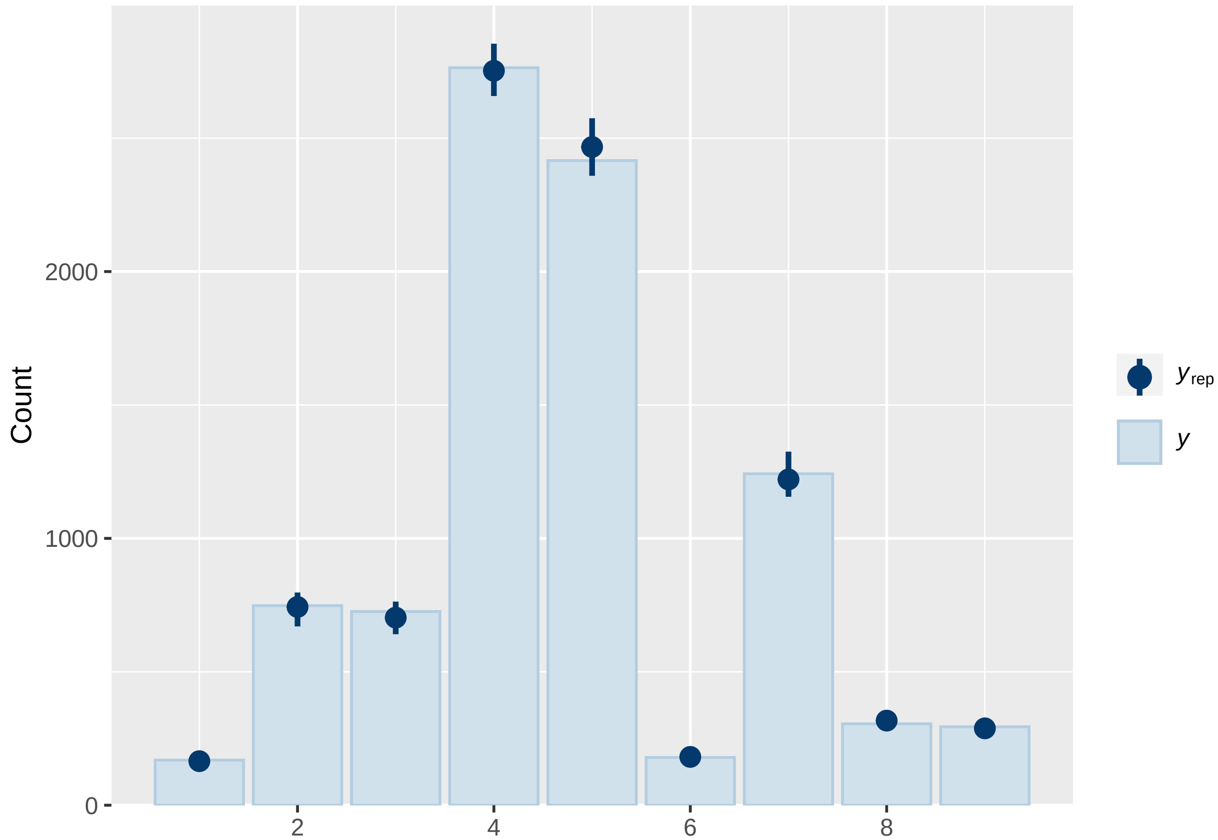
```
## PROB_TYPEPartMPartMWhole[1]     2087
## PROB_TYPEPartMPartMWhole[2]     2247
## PROB_TYPEPartMPartMWhole[3]     2087
## PROB_TYPEPartMPartMWhole[4]     2171
## PROB_TYPEPartMPartMWhole[5]     2268
## PROB_TYPEPartMPartMWhole[6]     2176
## PROB_TYPEPartMPartMWhole[7]     2319
## PROB_TYPEPartMPartMWhole[8]     2093
## PROB_TYPESeparate[1]            2065
## PROB_TYPESeparate[2]            2062
## PROB_TYPESeparate[3]            2047
## PROB_TYPESeparate[4]            1861
## PROB_TYPESeparate[5]            2169
## PROB_TYPESeparate[6]            2083
## PROB_TYPESeparate[7]            2337
## PROB_TYPESeparate[8]            2184
##
## Family Specific Parameters:
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## disc     1.00      0.00     1.00     1.00   NA       NA       NA
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Post-processing and Model Results

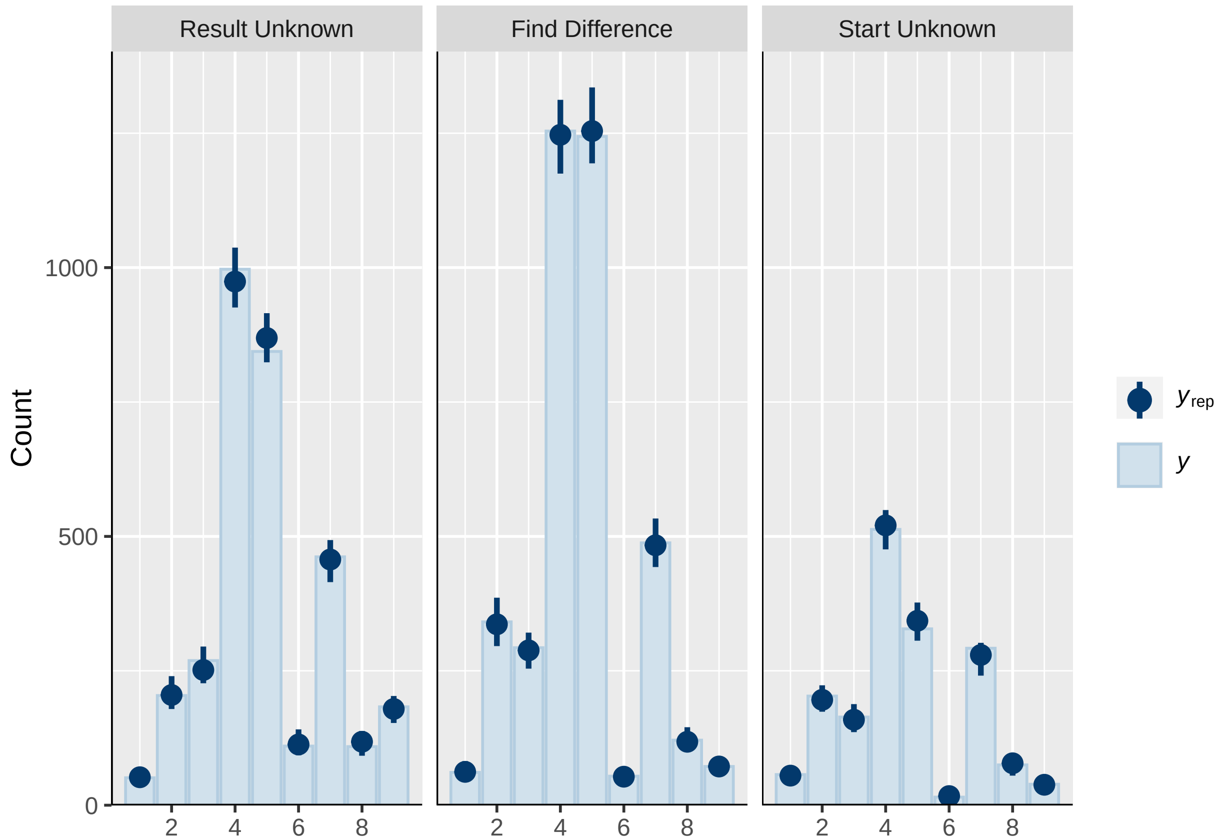
Here, we reproduce the figures that communicate the main results of the article. Although we showed that the baseline model presented in the article is robust, we adopt the variant with more MCMC samples (m_1b) in this section to take advantage of its greater Effective Sample Size.

Posterior predictive check

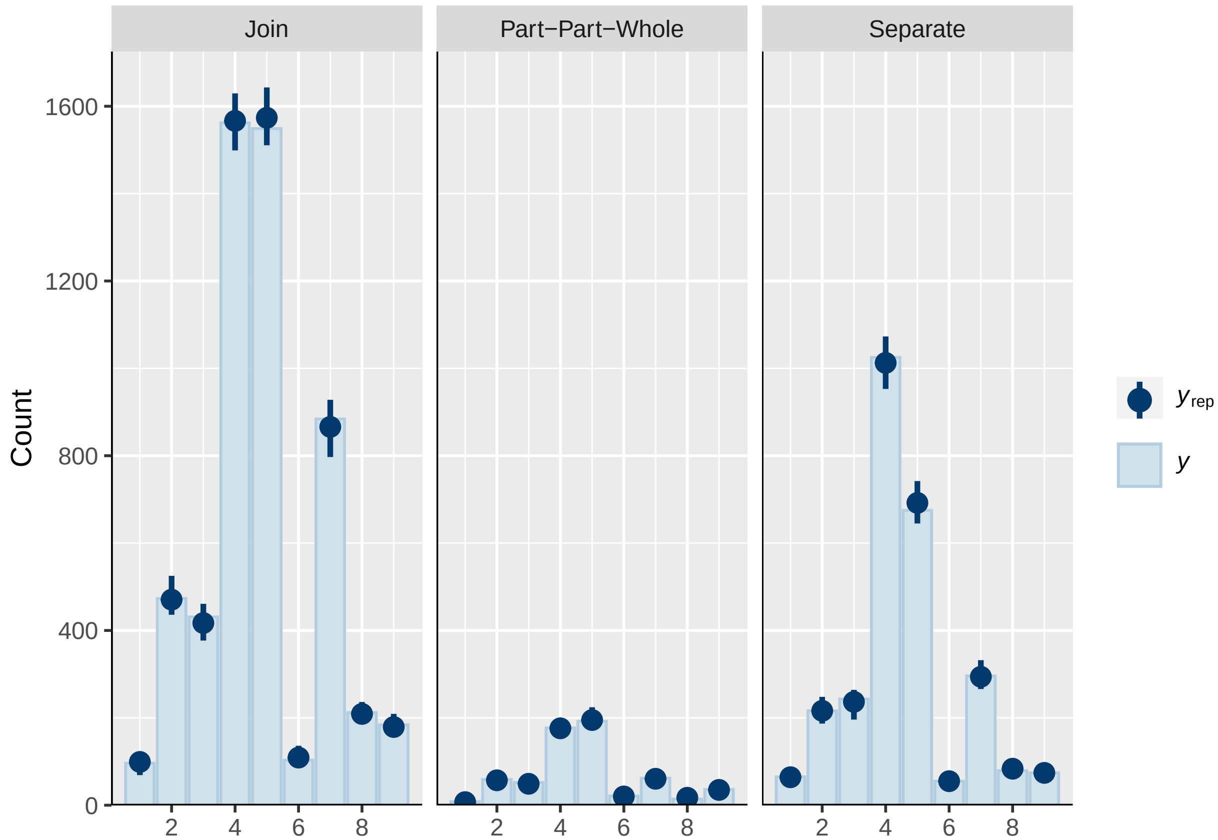
```
M<-m_1b
## posterior predictive check
# selected fixed effects
pp_check(M,type="bars",ndraws = 100)
```



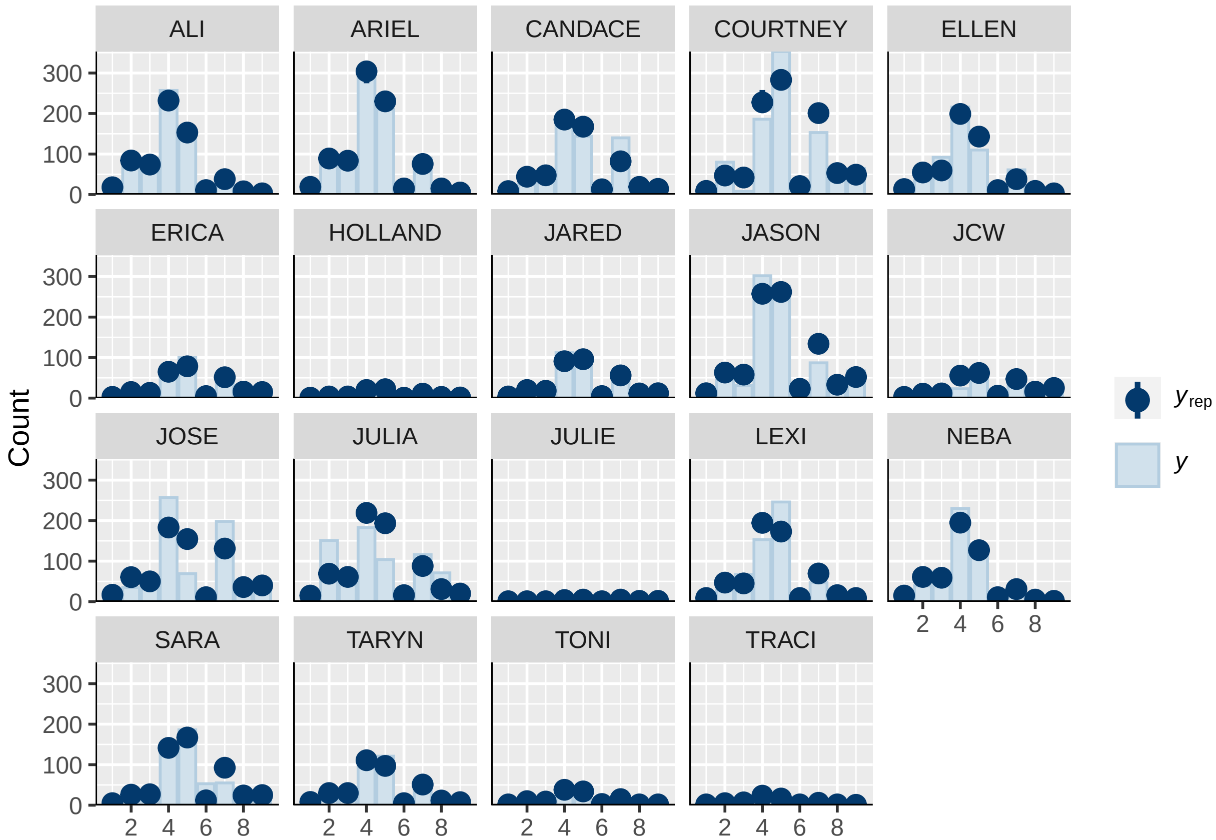

```r
pp_check(M,type="bars_grouped",group="PROB_CAT_C",ndraws = 100)
```




```
pp_check(M,type="bars_grouped",group="PROB_TYPE",ndraws = 100)
```




```r
# selected random effects
pp_check(M,type="bars_grouped",group="GRA",ndraws = 100)
```



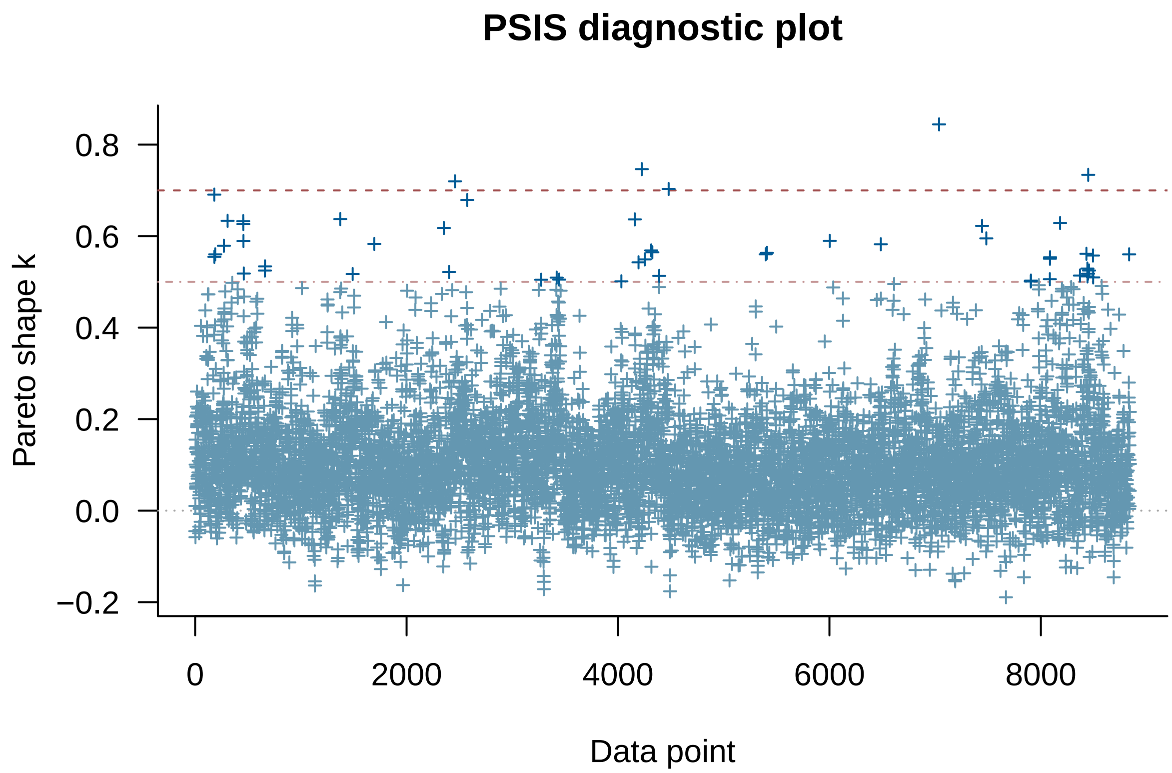

```
pp_check(M,type="bars_grouped",group="CHILD_ID",ndraws = 100)
```

The posterior predictive checks demonstrate the model fit is adequate: the observed strategy use patterns both overall and with respect to problem type and category are closely reproduced. Stratifying by the random effects of instructor (GRA) and child (CHILD_ID) reveals some lack-of-fit, as can be expected due with a random intercept. Overall, the model presented fits very well.

PSIS Diagnostic

```r
## checking for influential observations
# selected fixed effects
loo(M, save_psis = TRUE) %>% plot()
```

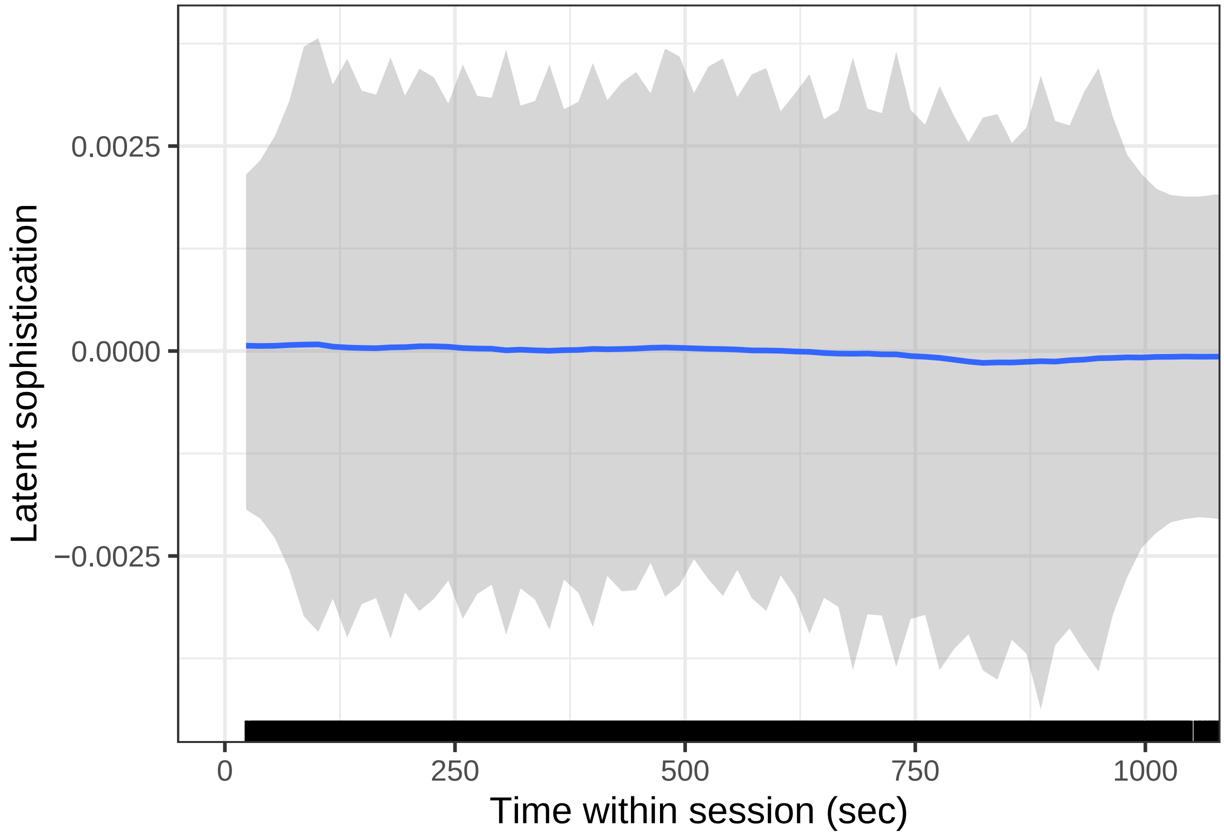
```
## Warning: Found 5 observations with a pareto_k > 0.7 in model 'M'. It is
## recommended to set 'moment_match = TRUE' in order to perform moment matching
## for problematic observations.
```



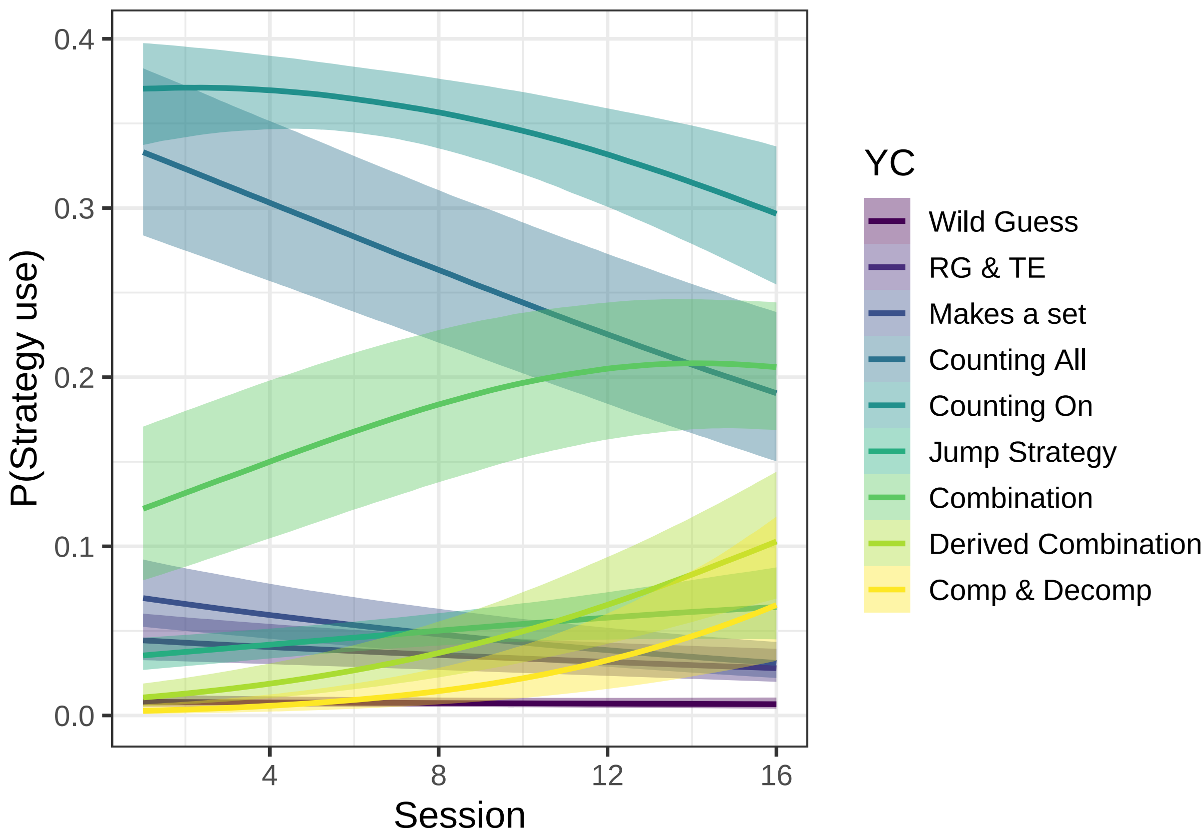
The PSIS diagnostic plot shows that very few observations may be deemed "influential" under the model. Remaining observations are well-supported by the model.

Model Results as Figures

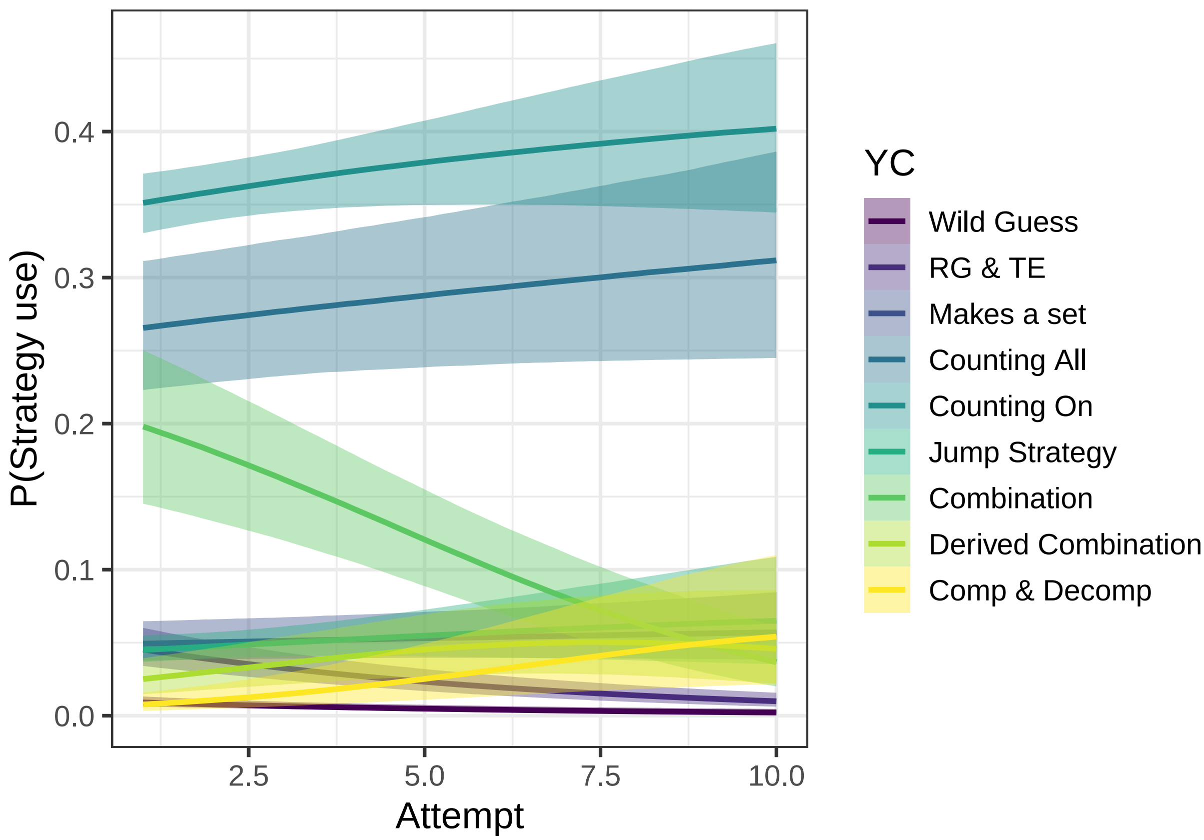
```
# Latent sophistication over time within session (Gaussian Process)
# 90% CrI for plot readability
plot(conditional_smooths(M, prob=0.9),rug=TRUE,plot=FALSE)[[1]] +
  xlab("Time within session (sec)") +
  ylab("Latent sophistication") +
  theme_bw(base_size = 14) +
  coord_cartesian(xlim=c(0,1030)) #only 2% of attempts > 1030secs
```



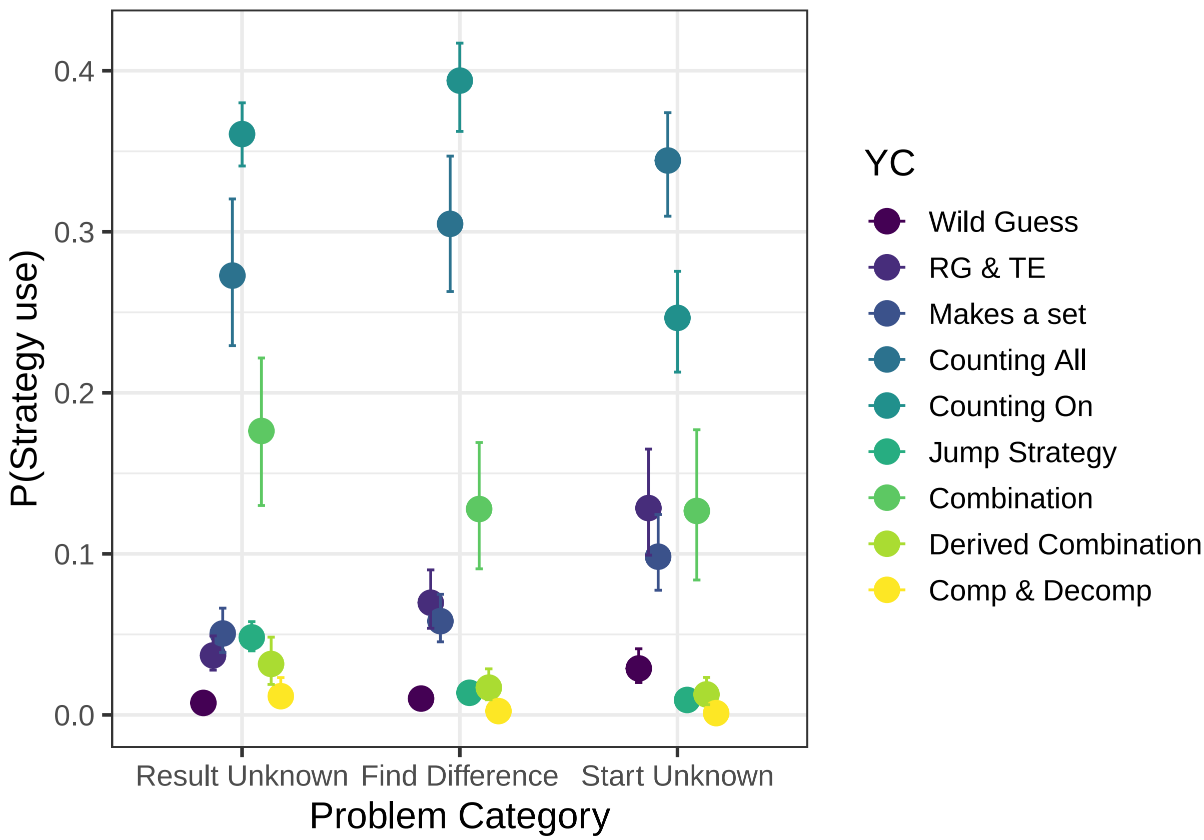

```
# Probability of strategy use across instructional sessions
# 90% CrI for plot readability
plot(conditional_effects(M,"SESSION",prob=0.9,
                          categorical=TRUE),plot=FALSE)[[1]] +
  scale_colour_viridis(alpha=1.0,direction=1,discrete = T) +
  scale_fill_viridis(alpha=1.0,direction=1,discrete = T) +
  xlab("Session") +
  ylab("P(Strategy use)") +
  theme_bw(base_size = 14)
```



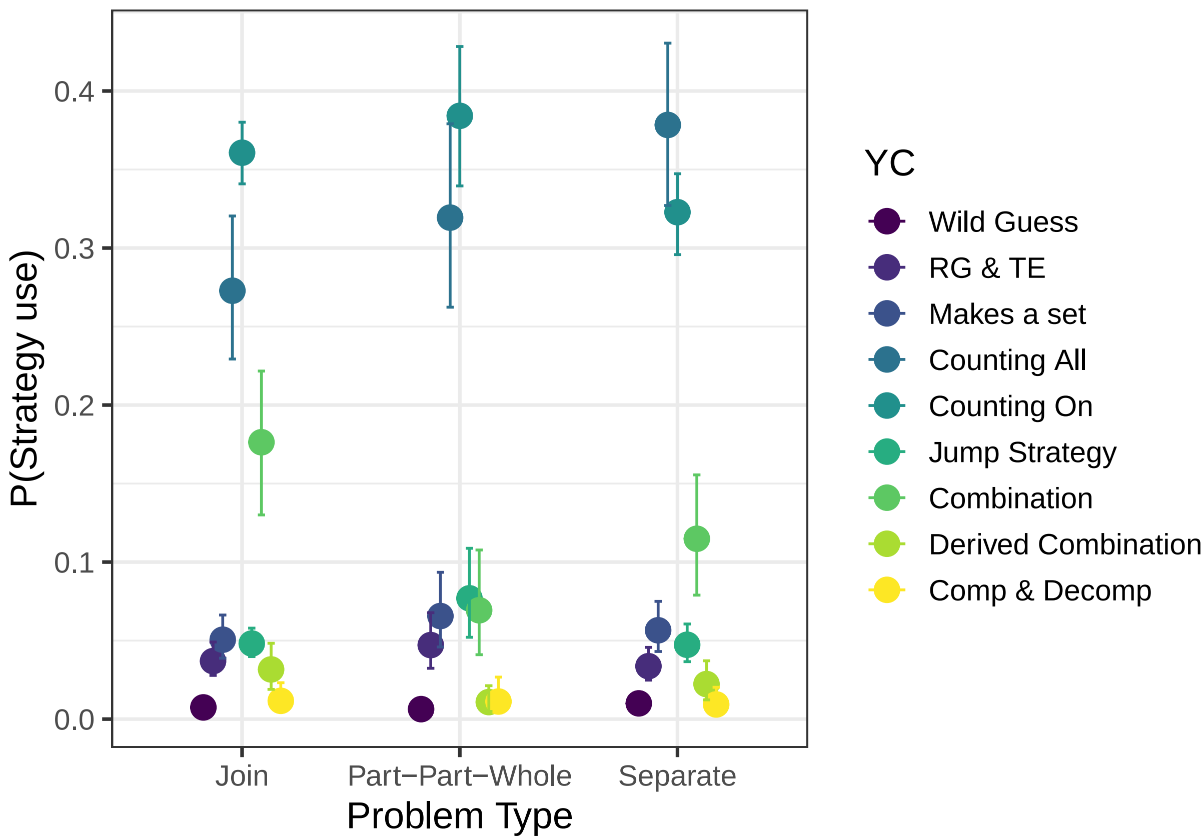

```
# Probability of strategy use across successive attempts
# 90% CrI for plot readability
plot(conditional_effects(M,"Attempt",prob=0.9,
                         categorical=TRUE),plot=FALSE)[[1]] +
  scale_colour_viridis(alpha=1.0,direction=1,discrete = T) +
  scale_fill_viridis(alpha=1.0,direction=1,discrete = T) +
  xlab("Attempt") +
  ylab("P(Strategy use)") +
  theme_bw(base_size = 14)
```




```r
# Probability of strategy use across problem categories
plot(conditional_effects(M,"PROB_CAT_C",prob=0.9,
                         categorical=TRUE),plot=FALSE)[[1]] +
  scale_colour_viridis(direction=1,discrete = T) +
  scale_fill_viridis(direction=1,discrete = T) +
  xlab("Problem Category") +
  ylab("P(Strategy use)") +
  theme_bw(base_size = 14)
```

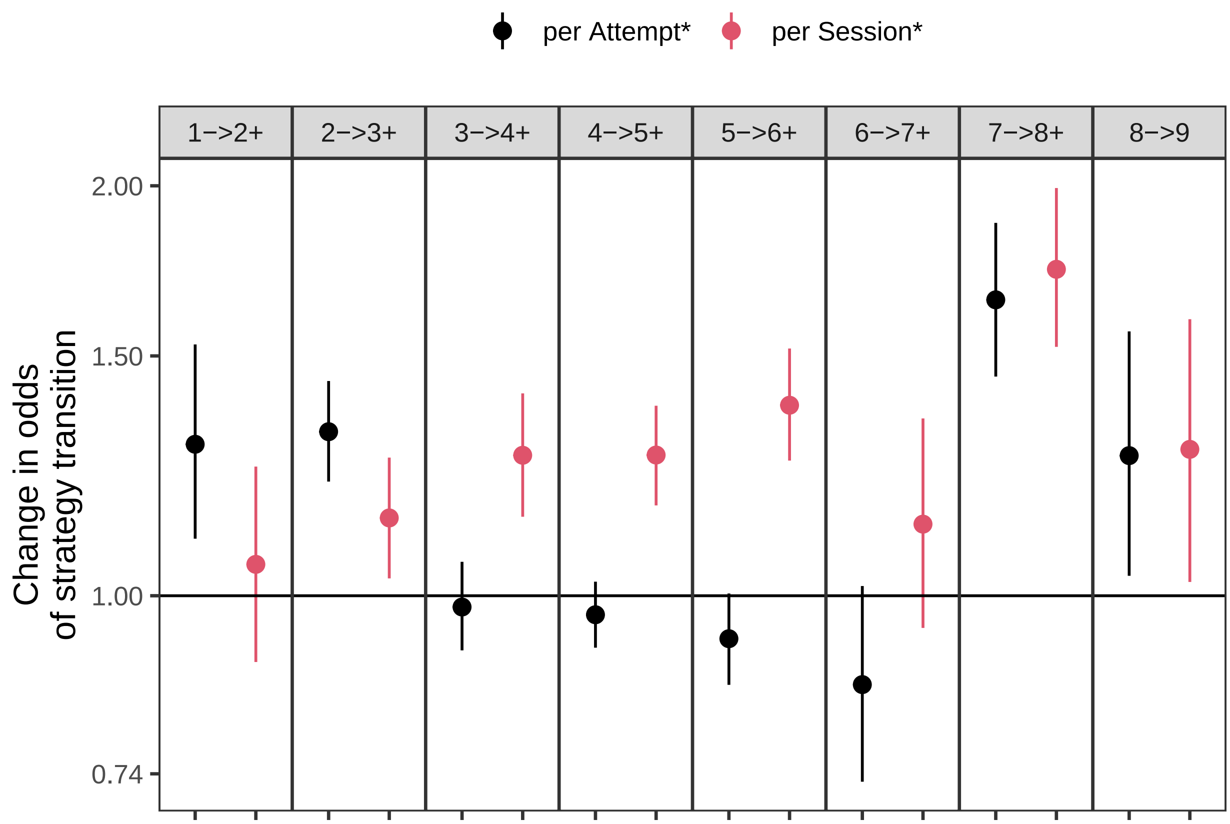



```
# Probability of strategy use across problem types
plot(conditional_effects(M,"PROB_TYPE",prob=0.9,
                                   categorical=TRUE),plot=FALSE)[[1]] +
  scale_colour_viridis(direction=1,discrete = T) +
  scale_fill_viridis(direction=1,discrete = T) +
  xlab("Problem Type") +
  ylab("P(Strategy use)") +
  theme_bw(base_size = 14)
```




```
# Probability of transition between strategies
# we edit the legend to omit the "SD" and will indicate in figure notes that we standardize Sessions an
strats<-fixef(M) %>% row.names() %>% str_extract("[([1-8])]") %>% as.factor()
levels(strats)<-c("1->2+","2->3+","3->4+","4->5+","5->6+","6->7+","7->8+","8->9")
names<-fixef(M) %>% row.names() %>% str_remove("[(1-8)]") %>% as.factor()
levels(names)<-c("Thres","Find Diff vs. Res Unk","Start Unk vs. Res Unk",
                 "P-P-Wh vs. Join","Sep vs. Join",
                 "per Attempt*","per Session*","time_sec")
##
fix_eff<-fixef(M) %>% as_tibble() %>%
         mutate(par_name = names,
                strat_num=strats) %>% slice(-c(1:9)) %>%
         filter(par_name %in% c("per Attempt*",
                                "per Session*"))
##
scaleFUN <- function(x) sprintf("%.2f", x)
ggplot(data=fix_eff,
             aes(x=par_name,
                 y=exp(Estimate),
                 ymin=exp(Q2.5),ymax=exp(Q97.5),
                 color=par_name)) +
     facet_wrap(~strat_num,nrow=1) +
     geom_pointrange() +
     geom_hline(yintercept=1)+
     scale_color_manual(values=c(1,2)) +
     scale_y_continuous(trans="log",
```

```
                         breaks=c(0.74,1.0,1.5,2.0),
                         labels=scaleFUN) +
     labs(x=NULL,y="Change in odds\nof strategy transition") +
     theme_bw(base_size=13) +
     theme(legend.position = "top",
           panel.grid.major = element_blank(),
           panel.grid.minor = element_blank(),
           axis.text.x=element_blank(),
           panel.spacing = unit(0,"cm")) +
    guides(color=guide_legend(title=NULL))
```




```
# Illustrating the variety of strategies the could be used by children to solve the same problem (Resul
pred_dat<-data.frame(CHILD_ID = unique(D_arith$CHILD_ID),
                     SESSION = 6, Attempt = 1, time_sec = 300,
                     PROB_CAT_C = "Result Unknown",
                     PROB_TYPE = "Join")
mod_pred<-predict(M,newdata = pred_dat,
                  re_formula = ~(scale(SESSION)|CHILD_ID)) %>%
          reshape::melt()
levels(mod_pred$X2)<-c(seq(1:9))
###
ggplot(data=mod_pred,aes(x=X2,y=X1,fill=value)) +
  geom_tile() +
  viridis::scale_fill_viridis(option="C")+
  labs(fill="P(Strategy\nuse)",x='Strategy',y='Student number',
       title="Result Unknown, Join\nSession 6, Attempt 1, 5mins in session") +
```

```
theme_bw(base_size=12)
```