Results and Modelling for ATC Experiment E1

Russell Boag

12 July 2017

Results

- Conventional statistical analyses are reported first in order to check whether our experimental manipulations had the expected effects on manifest RT and accuracy.
- Data from two particpants was excluded from the analyses; one who failed to complete all experimental blocks and one who made no PM responses at all for the entire experiment.
- We excluded trials with outlying RTs, defined as less than 0.2s or 3 times the interquartile range / 1.349 (a robust measure of standard deviation) above the mean (8.35% of responses overall). PM keypresses occurring during the control condition (which contained no PM stimuli) were also excluded (0% of responses overall).
- The following analyses compare mean accuracy and RT by stimulus type (conflict, nonconflict, PM) PM block (control, PM) and time pressure (Low, High) across 2 levels of trial load (2 vs 5 decision per trial).
- It should be noted that our time pressure factor was not crossed orthogonally with the trial load factor. Specifically, under low trial load (2 decisions per trial), low time pressure corresponded to a response deadline of 12 seconds (i.e., 6 seconds per decision on average) while high time pressure corresponded to a response deadline of 8 seconds (i.e., 4 seconds per decision on average). In contrast, under high trial load (5 decisions per trial), low time pressure corresponded to a response deadline of 20 seconds (i.e., 4 seconds per decision on average) while high time pressure corresponded to a response deadline of 10 seconds (i.e., 2 seconds per decision on average). As such, the following analyses compare the low and high time pressure levels separately for the low and high trial load conditions.
- Likewise, we can only sensibly compare low versus high trial load between the two conditions where time pressure is equivalent (i.e., the two blocks with time pressure of 4 seconds per decision). As such, all comparisons of trial load are done between the low load/high time pressure condition (i.e., 2 decisions per trial/4 seconds per decision) and high load/low time pressure condition (i.e., 5 decisions per trial/4 seconds per decision).
- In our omnibus significance testing for accuracy effects we used generalized linear mixed models with a probit link function. In our omnibus significance testing for mean correct RTs we used general linear mixed models.
- Significance was assessed with Wald's chi-square tests, and an alpha level of 0.05 was used in all analyses. The results of our omnibus analyses are tabulated in the supplementary materials. All standard errors reported in text and displayed in graphs were calculated using the bias corrected method.

Ongoing Task (Non-PM) Trials

- Accuracy was lower for conflicts (67.5%) compared to nonconflicts (80.6%), slightly lower under PM load compared to control (Control M = 74.9%, SE = 3.2%; PM M = 73.2%, SE = 3.4%), and decreased with higher levels of time pressure during both low trial load (Low TP M = 77.4%, SE = 2.7%; High TP M = 74.6%, SE = 2.7%) and high trial load conditions (Low TP M = 75.7%, SE = 2.5%; High TP M = 68.5%, SE = 2.5%).

- Mean RT was slower for conflicts (3.01s) compared to nonconflicts (2.65s), slower during PM blocks than control blocks (Control M = 2.62s, SE = 0.14s; PM M = 3.04s, SE = 0.15s), and faster across different levels of time pressure for both low trial load (Low TP M = 3.48s, SE = 0.12s; High TP M = 2.82s, SE = 0.08s) and high trial load conditions (Low TP M = 2.96s, SE = 0.09s; High TP M = 2.08s, SE = 0.07s).
- To summarise, the addition of PM load resulted in slower (*Mean Difference* = 0.43s) and slightly less accurate (*Mean Difference* = 1.7%) ongoing task performance, while increased time pressure led to faster but less accurate ongoing task performance.

PM Trials

- PM responses were scored correct if the participant pressed the PM-response key instead of an ongoing task (conflict/nonconflict) response key on the PM target trial.
- PM accuracy decreased across different levels of time pressure during both low trial load (Low TP M = 82.2%, SE = 1.8%; High TP M = 74.1%, SE = 2.1%) and high trial load conditions (Low TP M = 73.9%, SE = 1.9%; High TP M = 58.4%, SE = 2.6%).
- Mean RT for PM responses was significantly faster at higher levels of time pressure during both low trial load (Low TP M = 1.99s, SE = 0.05s; High TP M = 1.78s, SE = 0.05s) and high trial load conditions (Low TP M = 1.88s, SE = 0.05s; High TP M = 1.58s, SE = 0.05s).
- There were no significant differences in accuracy or RT between conflict PM targets and nonconflict PM target. This is expected since the PM cue (i.e., particular letters in an aircraft callsign) was completely non-focal, meaning the evidence used to make PM decisions was independent of evidence used to make ongoing task decisions.
- To summarise, as with the ongoing task, increased time pressure led to faster but less accurate PM performance.

Ongoing Task Responses on PM Trials compared with Non-PM Trials

- It is possible that reactive control over ongoing task decisions could lead to slower ongoing task RTs on PM trials in PM blocks, as compared with non-PM trials in PM blocks.
- To check whether reactive control was evident without the model-based analysis, we compared correct RTs on missed PM trials to correct ongoing task RTs. That is, RTs for 'conflict' responses to conflict PM targets and 'nonconflict' responses to nonconflict PM targets were compared with RTs for 'conflict' responses to non-PM conflicts and 'nonconflict' responses to non-PM nonconflicts (in the PM blocks).
- We ran a linear mixed effects model to examine the effects of stimulus type (Conflict, Nonconflict, PM (Conflict), PM (Nonconflict)) and time pressure on RTs for conflict and nonconflict responses.
- Planned comparisons revealed that conflict RTs were significantly faster on PM trials (2.96s) than on non-PM trials (3.21s). Likewise, nonconflict RTs were significantly faster on PM trials (2.63s) than on non-PM trials (2.88s).
- However, it should be noted that reactive control on PM trials is confounded in raw RT by statistical facilitation from the PM response. Specifically, on PM trials the accumulators for the ongoing task responses must compete with a much faster PM response accumulator. Overt ongoing task responses on PM trials are therefore more likely to be fast errors which outpace the PM accumulation process. As such, this comparison of mean RT is not sufficient to rule out reactive control. Rather, the critical test of reactive control is a comparison of accumulation rates, and is presented in the modelling section below.

Model Analysis

- Following Strickland (2017), we modelled the current task with a 3-accumulator LBA with two accumulators for the ongoing task (conflict/nonconflict) repsonses and one accumulator for the PM response.
- Each accumulator is described by several model parameters. Each accumulator begins a decision with a starting amount of evidence drawn from a uniform distribution on the interval [0, A]. After a stimulus is presented, evidence accumulates linearly at a rate drawn from a normal distribution with mean v and standard deviation *sv*. Evidence continues accumulating until it reaches a response threshold b; the first accumulator to reach threshold determines the overt response. Note that since A is held constant across all conditions, we report threshold in terms of B ($B = b - A$), where differences in B between experimental conditions reflect pure threshold effects. Finally, the addtional components of RT which fall outside of the decision process such as stimulus encoding and motor response time are captured by the nondecision time paramter *t0*.
- Our design includes several factors over which model parameters can vary, including latent response (i.e., conflict, nonconflict, and PM accumulators), and three manifest factors, stimulus type, time pressure/trial load, and PM demand. The latent response factor refers to the accumulators that can lead to each response (i.e., 'conflict', 'nonconflict', and 'PM'). It is important to be clear that the latent response factor corresponds to the accumulators, and not the response that was actually observed; the observed response is predicted by, not included in, the model.
- The stimulus type factor had four levels: *non-PM conflict*, *non-PM nonconflict*, *PM conflict*, and *PM nonconflict*. Since the time pressure (low, high) and trial load (2 vs. 5 decisions per trial) factors were not crossed orthogonally, they were captured by a four-level composite factor with the following levels: *low TP/low load*, *high TP/low load*, *low TP/high load*, *high TP/high load*. Lastly the PM demand factors had two levels: *control (i.e., no PM demand)* and *PM*.
- In order to reduce model complexity, we applied several theoretically sensible *a priori* constraints on which factors each paramter could vary over. First, we estimated only one A parameter for each participant, as is common practice in LBA modelling.
- Second, we allowed the *sv* parameter to vary by stimulus and latent response factors but not over the different PM block or time pressure conditions. This is more flexible than most previous LBA modelling, which only allows *sv* to vary as a function of whether the latent accumulator matches or does not match the stimulus. We used this more flexible approach because in our current model there are two types of 'correct' response for PM trials (i.e., correct PM and correct ongoing task decision). Also in line with Strickland (2017), we fixed the *sv* parameter for PM false alarms (i.e., 'PM' responses to non-PM stimuli) at 0.5. This is standard practice in LBA (and other parametric modelling) where one parameter must be fixed as a scaling parameter.
- Third, as with A, we estimated only one nondecision time (*t0*) parameter for each participant. This was done because our design minimized any potential differences in the motor movement required to make each response (i.e., participants kept their fingers positioned above the response key which were all located on one keyboard row). In addition, previous research has shown nondecision time does not appear to play a role in PM cost and is mostly negligible in explaining speed-accuracy trade-off effects. We also follow previous research in assuming that nondecision time is constant across trials.
- Finally, due to very low numbers of PM false alarms (i.e., PM responses to non-PM stimuli) we pooled estimates of both accumulation rate and variance (v and *sv*) across all experiment factors to give one PM false alarm accumulation rate and one corresponding *sv* parameter (which was used as a fixed scaling parameter as mentioned above).
- These *a priori* restrictions resulted in an 89 parameter most flexible 'top' model with one A, one *t0*, 20 B, 57 v, and 10 *sv* parameters. We compared this flexible top model against several simpler, more constrained variants outlined in the Model Selection section below.

Sampling

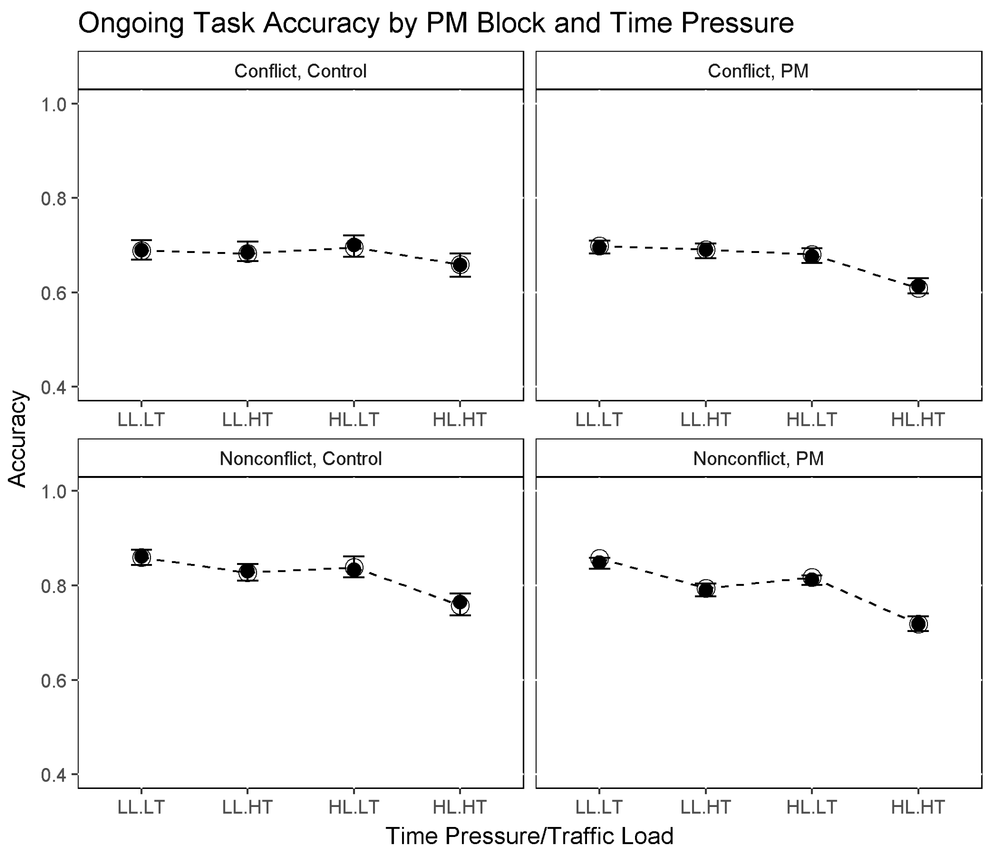
- Most evidence accumulation modelling to-date has relied on maximum-likelihood techniques to obtain point estimates of model parameter values. Here we instead use Bayesian techniques to estimate entire probability distributions of parameters rather than single point estimates.
- Although we could have fit a hierarchical model to estimate the common population distributions (hyperparameters) of each parameter, we opted to estimate parameters separately for each participant. This was due to several reasons. First, since this is the first time that such a model has been fit to this kind of task, we did not have adequate knowledge of the appropriate form of the population-level distributions. Because inappropriate population-level assumptions can introduce biases and shrinkage effects in hierarchical models, fitting to individual participants was the more conservative option. Second, because of the large number of participants in our sample and the complexity of our models, hierarchical methods proved too computationally expensive (estimated at several months of server time per fit) at the present time.
- Bayesian analysis requires the reseacher specify prior beliefs about the probabilities of parameters and the form of their distrubutions. However, note that because of our large sample sizes and use of inference based on posterior probability distributions, the influence of our particular choice of priors on the final parameter estimates was negligible. Since these analysis techniques have not been used on a dynamic applied task this complex, we did not have strong reasons to prefer any particular set of priors over others. We therefore used the modelling results of Strickland's (2017) PM task as a guide, but otherwise specified fairly uninformative priors (Table). All prior values were the same over control/PM blocks and the different levels of time pressure.
- Posterior parameter distributions were estimated using the differential evolution Markov-chain Monte-Carlo (DE-MCMC) algorithm. DE-MCMC is more adept at handling the high parameter correlations such as those common to evidence accumulation models. The number of chains was three times the number of parameters (e.g., for an 84 parameter model there were 252 chains per parameter). Chains were thinned by 20, meaning that one iteration in every 20 was kept. Sampling continued for each participant until a small Gelman's multivariate potential scale reduction factor (<1.1) indicated convergence, stationarity, and mixing. This factor is calculated with the number of chains doubled, by considering the first and second halves of each chain as separate chains. Convergence, stationarity, and mixing were verified by visual inspection. We retained the same number of samples for each participant: each of the 252 chains was 120 iterations long, producing 30,240 samples of each parameter's posterior distribution for each participant.

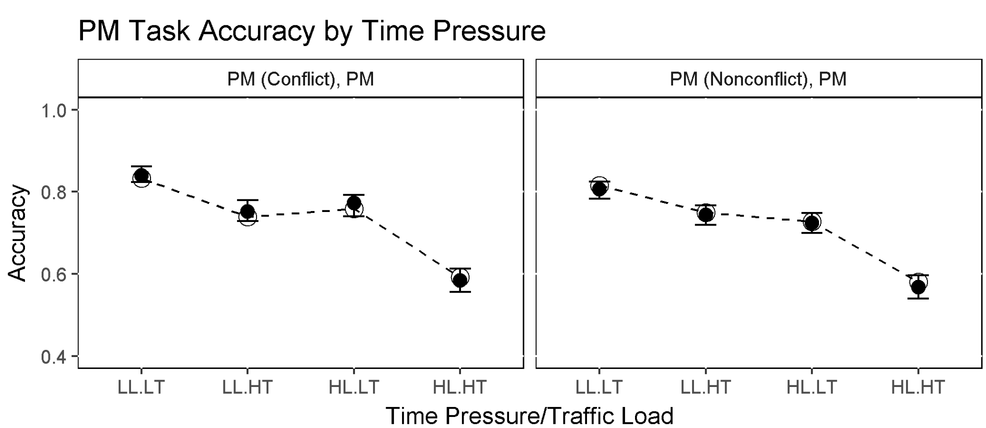
Model Parameter	Distribution	Mean	SD	Lower	Upper
A	Truncated Normal	3	1	0	10
B	Truncated Normal	2	1	0	None
v (Correct Ongoing Task Response)	Truncated Normal	1	2	0	None
v (Error Ongoing Task Response)	Truncated Normal	0	2	0	None
v (Correct PM Response)	Truncated Normal	1	2	0	None
v (PM False Alarm)	Truncated Normal	0	2	0	None
sv	Truncated Normal	0.5	1	0	None
t0	Uniform	0.3	1	0.1	1

Model Results

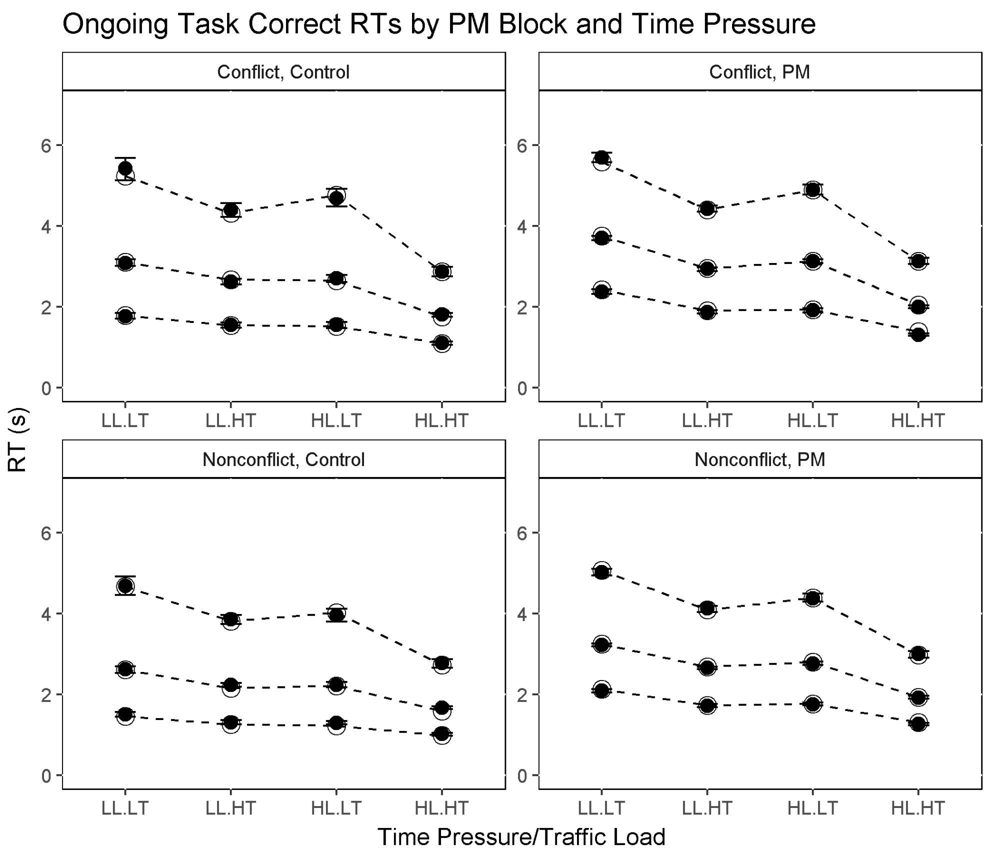
Model Fits: Accuracy and RT

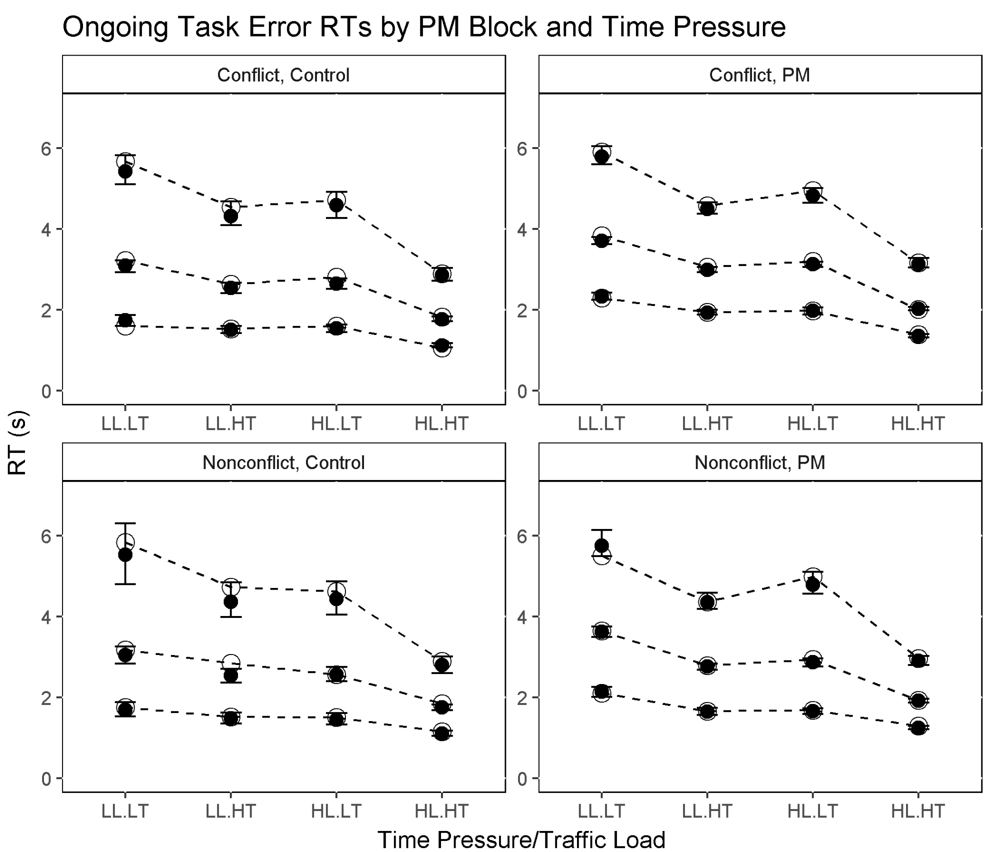
- To evaluate fit, we sampled 100 posterior predictions for each participant and then averaged over participants. The model provided good fits to both ongoing task and PM accuracy (Figure), and gave a good account of the entire distribution of response times (Figure). The model provided a close fit to the differences in manifest RT and accuracy observed across PM and control conditions and across different levels of time pressure. The next section explains how the model fit the data in terms of model parameters.
- Figure below shows model fits to ongoing task and PM accuracy by time pressure and PM block



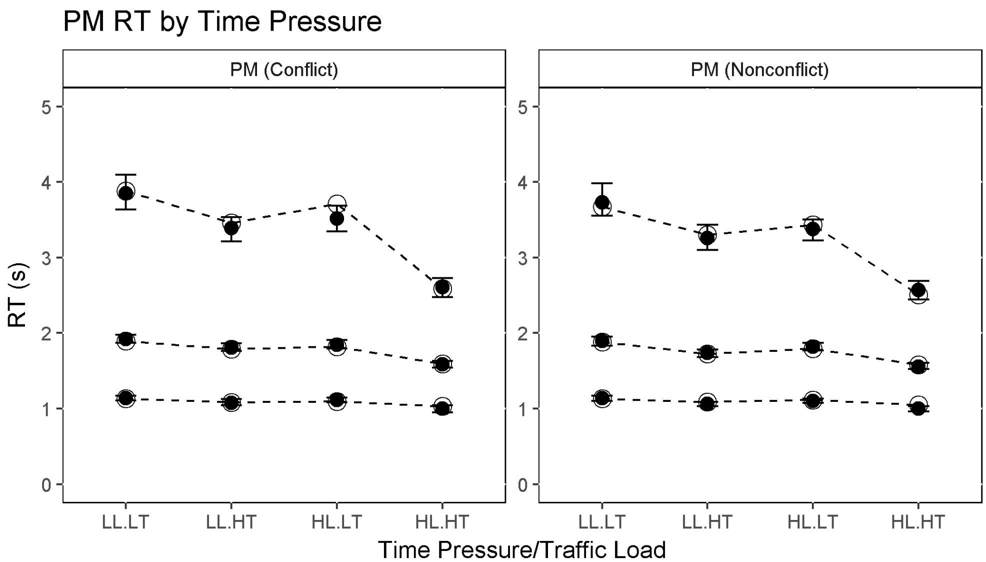


- Figure below shows model fits to ongoing task correct and error RT by time pressure and PM block


Ongoing Task Correct RTs by PM Block and Time Pressure
Conflict, Control
Conflict, PM
Nonconflict, Control
Nonconflict, PM
RT (s)
0
2
4
6
LL.LT
LL.HT
HL.LT
HL.HT
Time Pressure/Traffic Load

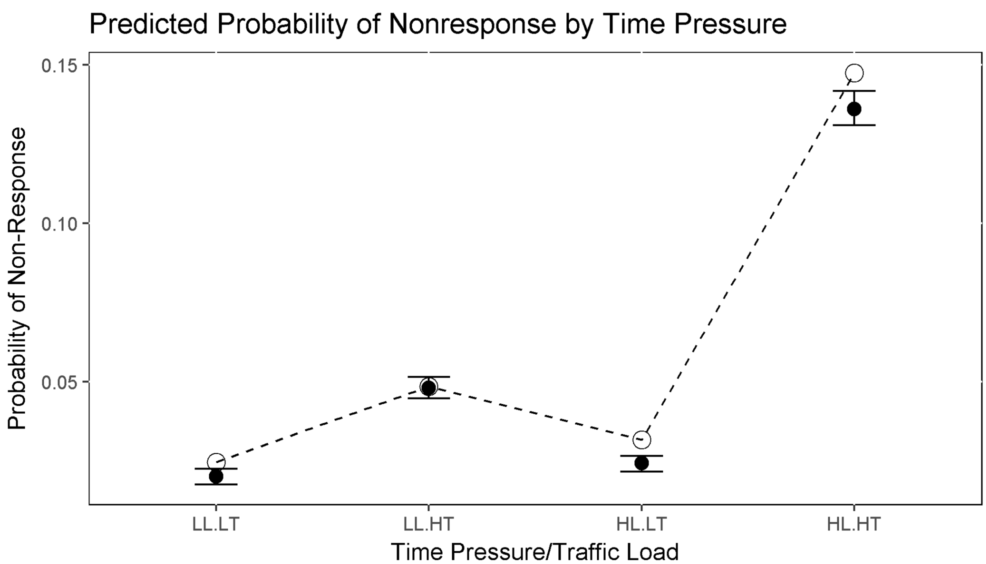

Ongoing Task Error RTs by PM Block and Time Pressure
Conflict, Control
Conflict, PM
Nonconflict, Control
Nonconflict, PM
RT (s)
0
2
4
6
LL.LT
LL.HT
HL.LT
HL.HT
Time Pressure/Traffic Load

- Figure below shows model fits to PM RT by time pressure and PM block



Model Fits: Nonresponse Proportions

- It should be noted that because the current model was fit to truncated data (i.e., with nonresponses removed), it is slightly misspecified. Specifically, because of the response deadline feature in our experimental design, a small proportion of our data were nonresponses which do not have associated RTs. As such, they do not contribute information to RT distributions fit by the LBA model.
- Since the model was fit without explicit information about nonresponses, we were interested in assessing the model's ability to predict nonresponses and how well those predictions fit empirical nonresponse proportion data. In order to do this, we simulated data out of the model and matched the order of the simulated stimuli and responses to the actual presentation order experienced by each participant. Whenever the cumulative sum of simulated RTs within a trial exceeded that trial's deadline, a nonresponse was predicted. Using this method, 100 posterior predictions for nonresponse proportions were sampled for each participant and predictions were then averaged over all participants. We then compared the predicted nonresponses with observed nonresponse proportions across the different levels of time pressure (i.e., different response deadlines) for both low and high trial load conditions.
- Figure below shows observed versus predicted nonresponse proportions.



- As shown in the figure, the model's predicted nonresponse proportions closely match the empirical nonresponse proportions. This gives us confidence that the slight model misspecification due to fitting to truncated data is not of concern in terms of the models predictive validity.

Model Selection

- We applied model selection to assess whether we could justify contstraining model parameters over blocked experimental conditions (e.g., Control/PM Block, Time Pressure) to obtain a simpler model with fewer parameters. To select between models, we used the Deviance Information Criterion (DIC), a measure which takes into account both goodness of fit and model complexity (number of parameters). In general, models with smaller DIC values are to be prefered as more parsimonious explanations of the data than models with larger DIC values. Table below shows each model we compared, its number of parameters, and its corresponding DIC value.
- Starting with the fully flexible top model, we built several simpler variants by systematically constrained threshold and rate parameters over PM and time pressure factors. This allowed us to assess whether it was necessary to vary thresholds and/or rates to account for observed PM demand and time pressure effects. We compared the following four constrained models to the top model:
- A model in which rates could vary by time pressure but thresholds could not
- A model in which thresholds could vary by time pressure but rates could not
- A model in which rates could vary across PM and control blocks but thresholds could not, and
- A model in which thresholds could vary across PM and control blocks but rates could not
- In each case the simpler model was rejected in favour of the fully flexible top model, suggesting that it is necessary to allow both rate and threshold parameters to vary over PM and time pressure (i.e., both parameters are influenced by PM and time pressure manipulations and are important in explaining the observed data).
- Finally, we tested an additional model (the selected model) which allowed both rates and thresholds to vary over both PM and time pressure, but included a slight simplification from the top model. The simplification involved constraining the PM rate parameter such that it was not allowed to vary over stimulus type (i.e., PM conflicts and PM nonconflicts had the same accumulation rate). This simplification makes theoretical sense, since the evidence used to make a PM decision (i.e., particular letters in an aircraft callsign) is independent of the evidence used to make either conflict or nonconflict

ongoing task decisions (i.e., speed, relative distance, and motion). This slightly simpler model produced the smallest DIC value and was thus selected as our preferred model.

- Although the results of model selection suggest that both rates and thresholds have some role in explaining PM cost and time pressure effects, we cannot say how important each parameter is or what proportion of a given effect is accounted for by each parameter. As such, in the next section we test the direction and magnitude of differences between conditions in the parameters of the selected model. Testing the direction of effects is important because it allows us to distinguish between competing theories of PM costs, whose predictions are also directional (e.g., capacity-sharing theories predict lower accumulation rates under PM load than control). Testing the magnitude of effects is similarly important, especially in applied settings, as it indicates which processes contribute the most to a given effect (such as PM costs) or are most affected by an experimental manipulation.

Model	Number.of.Parameters	DIC
Top Model	89	187068
Selected Model	84	186995
Selected Model with B fixed over Time Pressure	73	188370
Selected Model with V fixed over Time Pressure	81	190690
Selected Model with B fixed over PM Block	47	190742
Selected Model with V fixed over PM Block	74	190640

Model Summary

- To summarise the central tendency of model parameters over participants, we created a subject-average posterior distribution. This was obtained by computing the mean of each posterior sample over all participants for each parameter. In terms of answering theoretical questions, our primary interest is in mean threshold and accumulation rate parameters, which we explore in detail in the following sections. The other parameters all had reasonable mean values. The nondecision time mean of the subject-average posterior distribution was 0.35 (posterior $SD = 0.01$). The A posterior mean was 3.3 (posterior $SD = 0.05$). The sv posterior means and SDs are summarised in Table X. Consistent with other LBA modelling studies, sv parameters for the ongoing task are lower for correct response accumulators compared to error response accumulators.

Parameter	Mean	SD
sdv.ccC	0.3664	0.007816
sdv.nnC	0.551	0.0144
sdv.pcC	0.5419	0.02389
sdv.pnC	0.5911	0.0301
sdv.ccN	0.5319	0.01099
sdv.nnN	0.4613	0.008399
sdv.pcN	0.5403	0.03012
sdv.pnN	0.6254	0.0243
sdv.ppP	1.153	0.02907

- We next test the direction and magnitude of differences in threshold and accumulation rate parameters for the selected model across experimental conditions in order to assess how well they correspond to the theoretical predictions of capacity sharing, proactive control, reactive control, and effort/arousal.

- To this end, we calculated posterior distributions of the differences between experimental conditions. For example, to test the difference between response thresholds in control and PM conditions (i.e., testing the proactive control account of PM costs), we subtracted the control condition threshold from the PM condition threshold for every posterior sample, thus obtaining the posterior probability distribution of the difference between control and PM thresholds. Difference distributions were calculated independently for each participant before being averaged across participants to create a subject-averaged posterior difference distribution.
- For each subject-averaged difference distribution we report a Bayesan posterior-predictive p-value (Meng, 1994), which indicates the one-tailed probability that the difference between parameters is less than zero.
- Due to the power of our design, almost all of our observed parameter differences have $p = 0$, indicating a probability of 1 that an effect was present. However, some of our parameter differences were much larger in magnitude than others. As such, we illustrate the magnitude of the effect by reporting the standardised difference between parameters (i.e., M / SD of the posterior difference distribution). Because our posterior parameter distributions are approximately normal, this standardised statistic can be interpreted in a similar way to a Z-score. We therefore refer to this statistic as Z from here on.

Capacity Sharing (Non-PM Trial Accumulation)

- Capacity sharing theories of PM costs propose that holding PM intentions or monitoring for PM stimuli draws limited-capacity cognitive resources away from the ongoing task. As such, they predict that ongoing task evidence accumulation rates will be higher in control conditions (when more resources can be devoted to the ongoing task), and lower under PM load (when resources must be shared between the ongoing task and concurrent PM monitoring processes), which would lead to slower ongoing task RTs under PM load. Another prediction consistent with capacity sharing is increased accumulation rates for error responses under PM load relative to rates for correct responses. This would lead to lower accuracy under PM load relative to control.
- The figure below shows accumulation rates in control and PM blocks for non-PM ongoing task stimuli (i.e., conflicts and nonconflicts that were not also PM targets). Contrary to the predictions of capacity sharing theories, accumulation rates for correct responses were higher under PM load than in the control condition for both conflict and nonconflict stimuli. It should be noted however that accumulation rates for error responses were also greater under PM load, which is not inconsistent with a capacity sharing account.
- Taken together, this pattern of large increases in correct and error accumulation rates from control to PM conditions provides convincing evidence against a capacity sharing account of PM costs in this task. As will be discussed later, these results are more indicative of an overall increase in effort or arousal/task engagement during PM blocks, which may be because the ongoing task becomes subjectively more difficult and/or engaging with the addition of the concurrent PM task relative to control blocks.

Proactive Control (Thresholds)

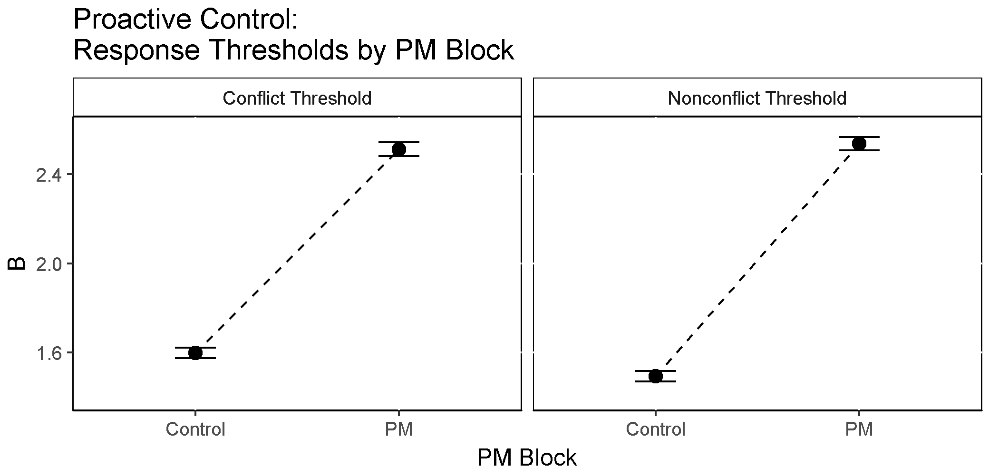
Proactive Control under PM Demand

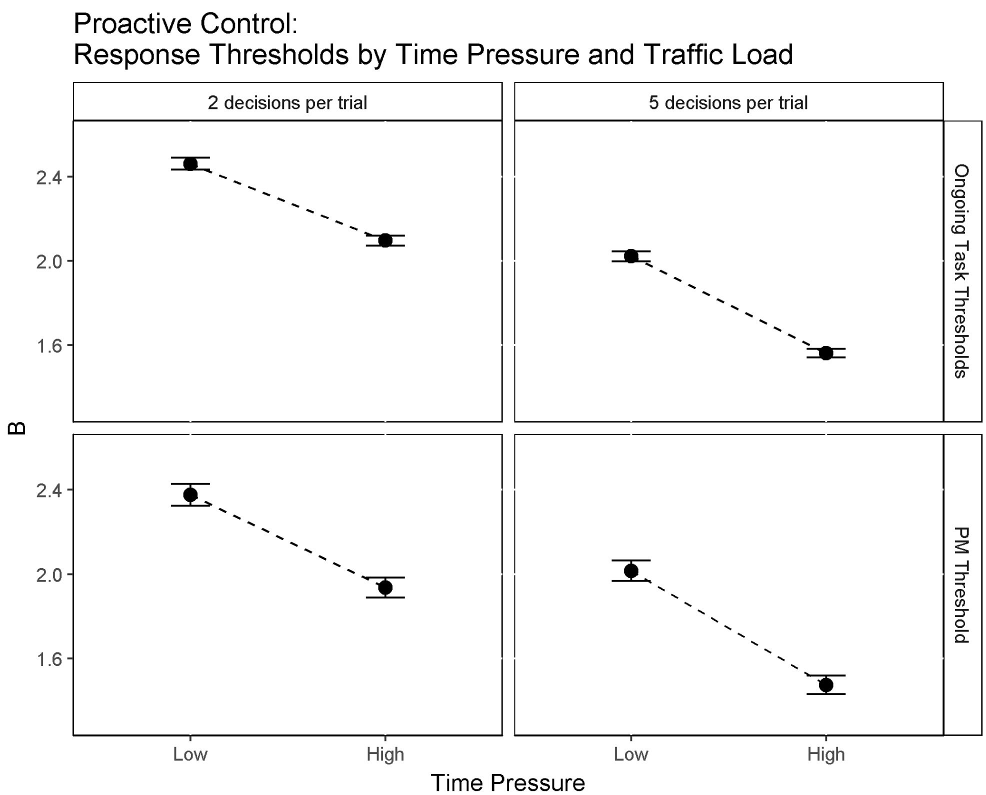
- Proactive control over ongoing task decisions predicts higher conflict and nonconflict response thresholds under PM load compared to during control blocks.
- Z-score effect sizes and p-values for threshold comparisons are shown in the table below.
- The figure below shows conflict and nonconflict response thresholds in the control and PM blocks.
- Ongoing task thresholds where much higher in PM than control blocks

- This is consistent with strategic delay theories on PM costs where responses to the ongoing task are deliberately delayed in order to avoid preempting PM targets.

Proactive Control under Time Pressure

- Spiel about speed-accuracy trade-off and adjusting threshold to increase speed at expense of accuracy
- The lower panels of the figure below also show ongoing task and PM response thresholds under low and high time pressure for both low trial load (2 decisions per trial) and high trial load (5 decisions per trial)
- In both cases thresholds decreased under high time pressure relative to low time pressure
- Consistent with much choice-RT modelling of the speed-accuracy trade-off
- Consistent with proactive control of decision processes whereby thresholds are strategically lowered in order to facilitate fast responding (at the expense of accuracy)

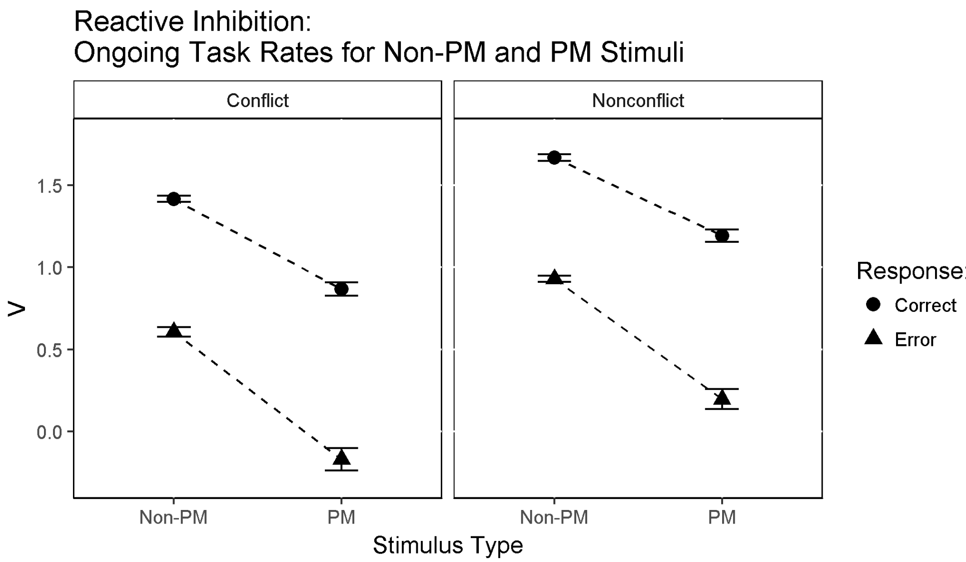




Reactive Inhibition (PM vs. Non-PM Trial Accumulation)

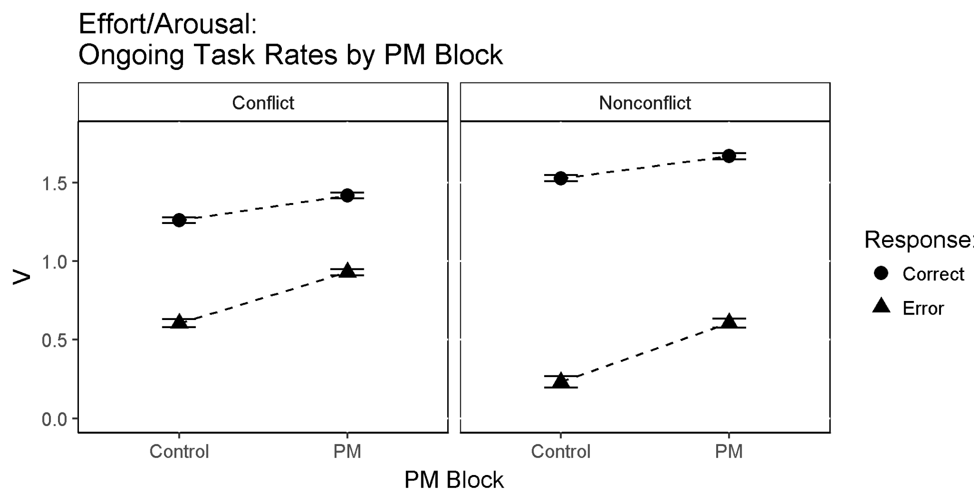
- A prediction of PMDC theory is that ongoing task (conflict/nonconflict) evidence accumulation rates will be lower on PM trials due to 'reactive' inhibitory control of the decision process by the PM stimulus detector.
- The figure below shows accumulation rates for ongoing task responses to non-PM conflict/nonconflict stimuli compared to PM conflict/nonconflict stimuli (i.e., ongoing task stimuli that also contained a PM target).
- Consistent with the predictions of PMDC's reactive inhibition mechanism, rates for ongoing task accumulators where much lower for stimuli containing a PM cue compared to when the same stimuli did not contain a PM cue.

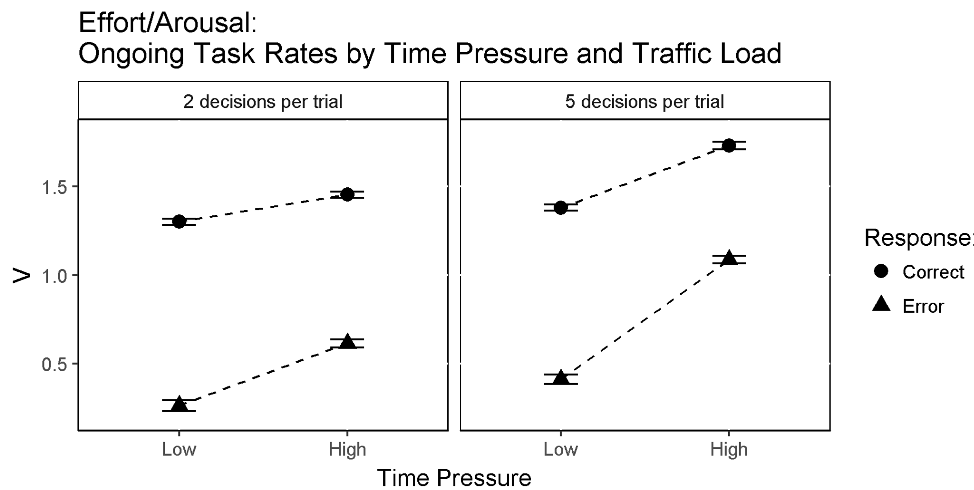
- This supports the idea that when the PM detector detects a PM target, the accumulation process for the competing ongoing task response is supressed or inhibited.
- The reactive inhibition of ongoing task responses was slightly stronger for the incorrect ongoing response accumulators
- These findings are consistent with the idea that response accumulators compete with each other based on their inputs; in the presence of a PM stimulus, evidence accumulation processes for conflict and nonconflict are inhibited relative to when a PM stimulus is absent.



Effort/Arousal (Accumulation Rate Increases with PM Load and Time Pressure)

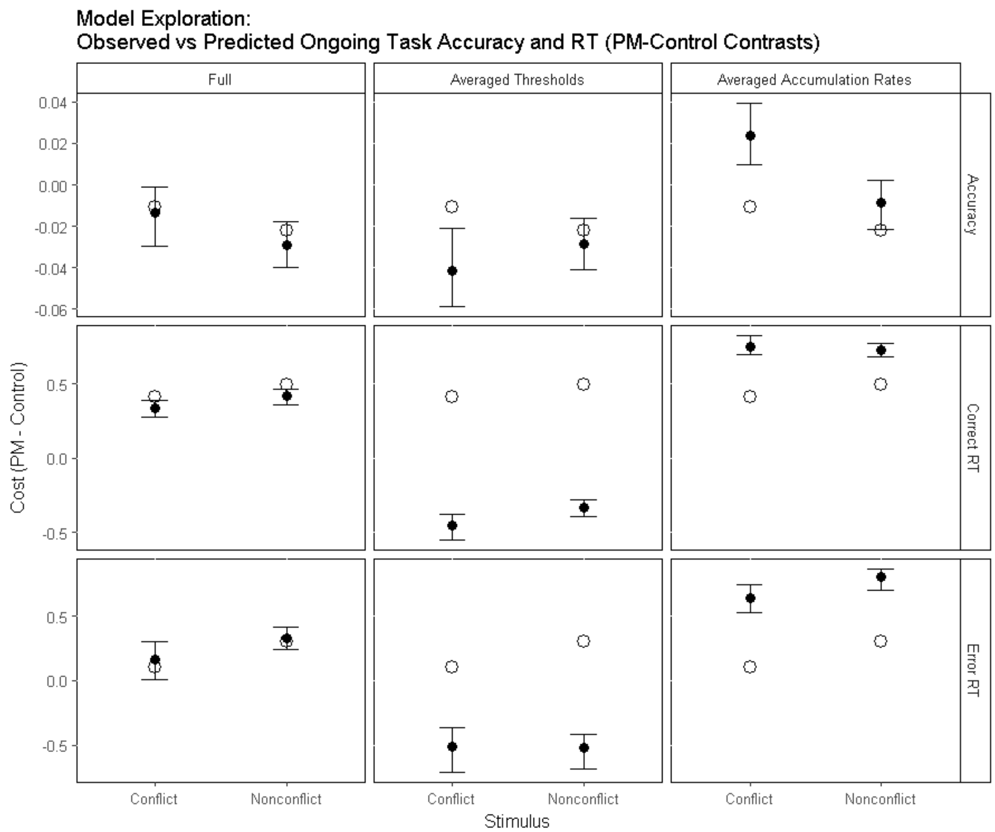
- As mentioned previously, in contrast with the predictions of capacity-sharing theories of PM costs, we found that evidence accumulation rates actually increased with the addition of PM load.
- The figure below also shows correct and error accumulation rates across the different levels of time pressure for both low trial load (2 decisions per trial) and high trial load (5 decisions per trial) conditions.
- As with the increases seen from control to PM blocks, both correct and error accumulation rates were higher under higher time pressure relative to low time pressure.
- Since rates for both correct and error responses increase, this effect is also suggestive of an overall increase in arousal/task engagement or the overall effort being invested in completing the task.
- Possible that the task becomes more difficult or engaging under high time pressure leading participants to deploy more resources
- Alternatively participants could be ‘satisficing’, that is, in the relatively easy control and low time pressure conditions participants may believe that they can disengage from the task somewhat, expending fewer cognitive resources while maintaining a satisfactory level of task performance.



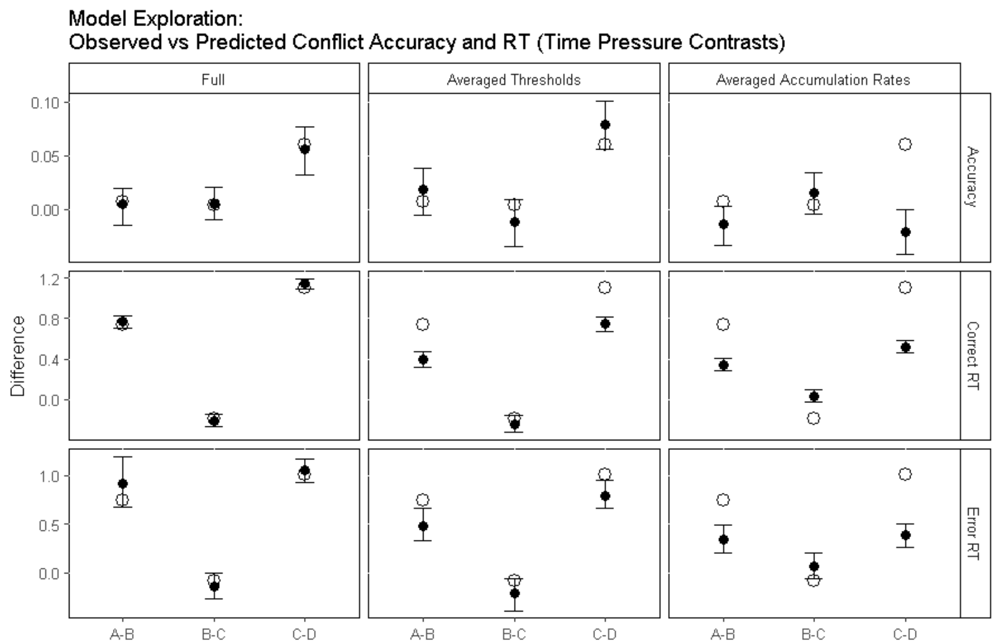


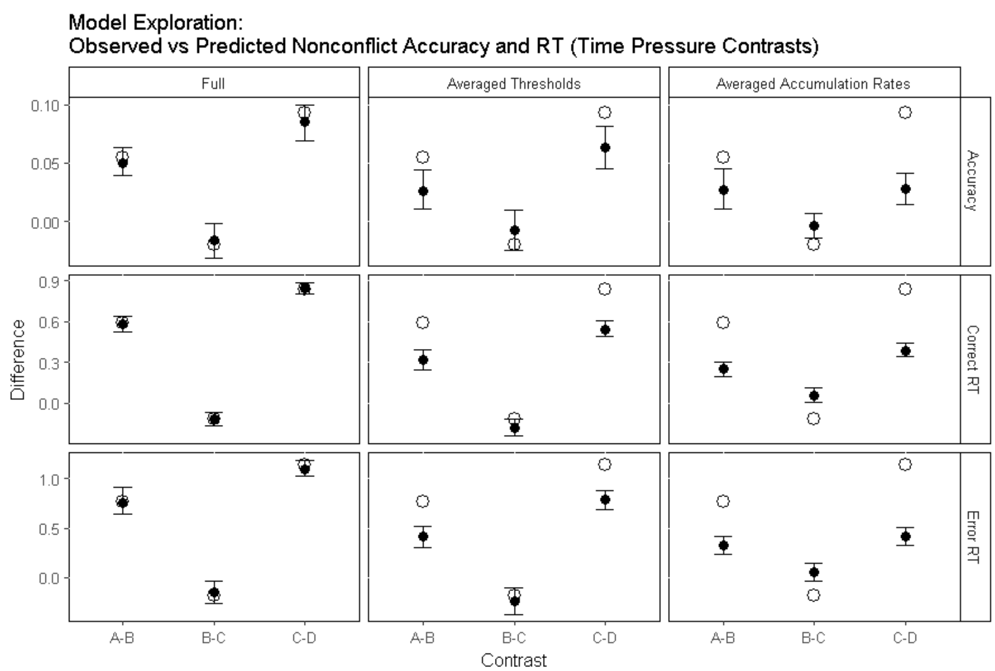
Model Exploration

- Given the complexity of our model, it is difficult to discern the overall contribution given parameters have on overall RT and/or accuracy. In this section we attempt to tease out the individual contribution to RT and accuracy provided by certain parameters and mechanisms in the model. It is our aim to give a clearer picture of the relative importance of key parameters in accounting for the observed effects.
- In order to evaluate a given parameter's contribution to the model, we first replace that parameter with the average either across control/PM blocks or across time pressure levels (e.g., replacing control and PM ongoing task thresholds with the average of the two). We can then examine the associated mis-fit of the model with the removed effect relative to the full model with the effect included.
- Compare fits to accuracy and RT for a model with thresholds averaged over control and PM blocks and a model with rates averaged over control and PM blocks.

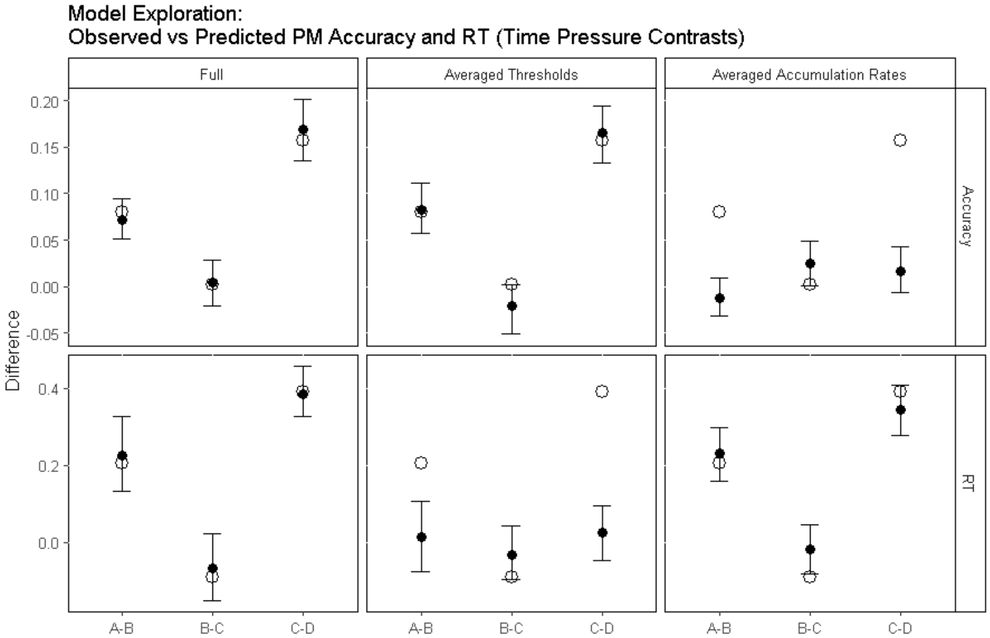


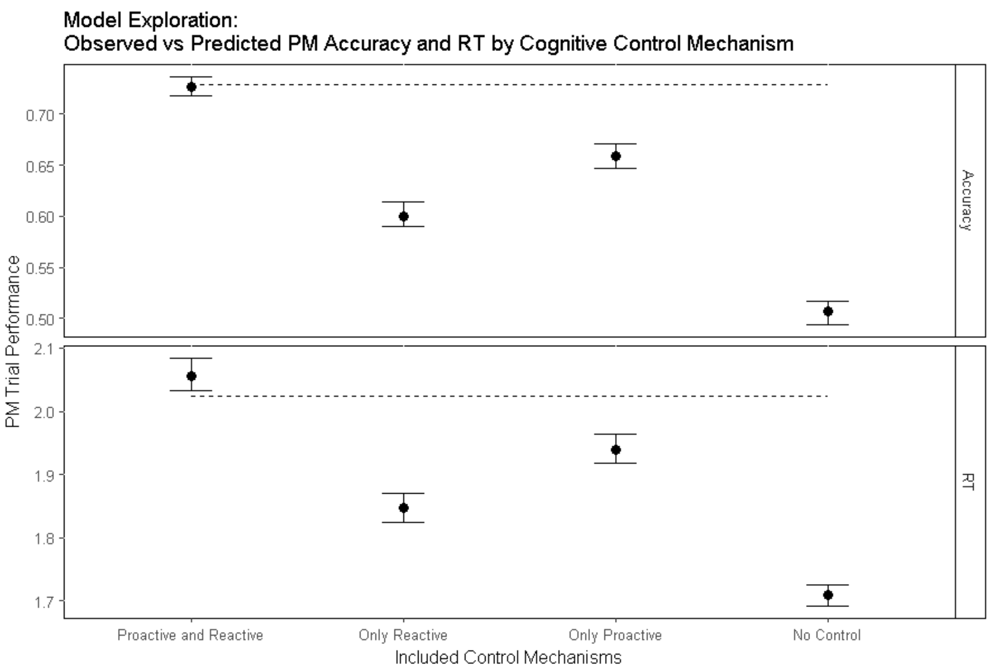
- Compare fits to accuracy and RT for a model with thesholds averaged over time pressure and a model with accumulation rates averaged over time pressure.





- Compare fits to overall RT and accuracy for a model which only includes reactive control mechanisms, a model which only includes proactive control mechanisms, and a model with neither control mechanism.


Model Exploration:
Observed vs Predicted PM Accuracy and RT (Time Pressure Contrasts)
Full
Averaged Thresholds
Averaged Accumulation Rates
Accuracy
RT
Difference
0.20
0.15
0.10
0.05
0.00
-0.05
0.4
0.2
0.0
A-B
B-C
C-D


Model Exploration:
Observed vs Predicted PM Accuracy and RT by Cognitive Control Mechanism
Accuracy
RT
PM Trial Performance
0.70
0.65
0.60
0.55
0.50
2.1
2.0
1.9
1.8
1.7
Proactive and Reactive
Only Reactive
Only Proactive
No Control
Included Control Mechanisms