Simple questions on simple associations:
Regularity extraction in non-human primates

Jeremy Yeaton[1,3], Laure Tosatto[1,2], Joel Fagot[1,2,4], Jonathan Grainger[1,2], Arnaud Rey[1,2]

[1]Aix Marseille Univ, CNRS, LPC, Marseille, France
[2]Aix Marseille Univ, ILCB, Aix-en-Provence, France
[3]Department of Language Science, University of California, Irvine
4 Station de Primatologie, CNRS-Celphedia, UPS 846, Rousset-sur-Arc, France

Running title: SIMPLE ASSOCIATIONS

Word count (main text): 3849

Correspondence concerning this article should be addressed to:
Jeremy Yeaton
Department of Language Science, University of California – Irvine
2243 Social Sciences Plaza
Irvine, CA 92617, USA
Email: jyeaton@uci.edu

Abstract

When human and non-human animals learn sequences, they manage to implicitly extract statistical regularities through associative learning mechanisms. In two experiments conducted with a non-human primate species (Guinea baboons, *Papio papio*), we addressed simple questions on the learning of simple AB associations appearing in longer noisy sequences. Using a serial response time task, we manipulated the position of AB within the sequence, such that it could be either fixed (by appearing always at the beginning, middle, or end of a 4-element sequence; Experiment 1) or variable (Experiment 2). We also tested the effect of sequence length in Experiment 2 by comparing the performance on AB when it was presented at a variable position within a sequence of 4 or 5 elements. The slope of RTs from A to B was taken for each condition as a measurement of learning rate. While all conditions differed significantly from a no-regularity baseline, we found strong evidence that the learning rate did not differ between the conditions. These results indicate that regularity extraction is not impacted by the position of the regularity within a sequence and by the length of the sequence. These data provide novel general empirical constraints for modeling associative mechanisms in sequence learning.

Keywords: statistical learning, sequence learning, associative learning, animal cognition

Introduction

A key building block of our cognitive life is the ability to detect regularities between two events A and B that cooccur frequently in the environment. By repeatedly processing and encoding these AB regularities, we progressively manage to predict or expect B when A is appearing. These fundamental statistical learning abilities help us learn and execute complex sequences of information more rapidly and fluidly (Christiansen, 2019; Frost et al., 2019; Perruchet & Pacton, 2006).

This crucial ability has been studied in human and non-human primates like tamarins (e.g., Hauser et al., 2001), macaques (e.g., Wilson et al., 2015), and baboons (e.g., Malassis et al., 2018; Minier et al., 2016; Rey et al., 2019, 2022; Tosatto et al., 2022), suggesting that regularity extraction is supported by common associative learning mechanisms across these species (Rey et al., 2019). The advantage of studying these mechanisms in non-human primates is the absence of any interference related to language recoding processes that may blur the study of associative learning mechanisms.

Previous studies with Guinea baboons (*Papio papio*) have revealed several general properties of these associative mechanisms. These studies used a visuo-motor pointing task derived from the serial reaction time task (Nissen & Bullemer, 1987) in which baboons were expected to touch a moving target on a touch screen that could appear on nine equidistant possible positions. It was found that when baboons were repeatedly exposed to regularities composed of three successive positions ABC, RTs on the third position of the regular triplet (i.e., C) was found to decrease faster than RTs on the second position (i.e., B) (Minier et al., 2016). Using sequences of three positions, it has also been reported that baboons were able to learn second-order regularities when first-order regularities were inconsistent (Rey et al., 2022). With longer sequences composed

of nine positions, it has been shown that baboons are segmenting these long sequences into chunks of 3 to 4 positions, revealing fundamental limits of associative learning mechanisms (Tosatto et al., 2022).

In the present study, our goal was to expand further our knowledge about the general properties of associative mechanisms in sequence learning and to address rather simple questions on simpler associations. Considering the simplest case of an AB regularity, we tested if the learning of this regularity would vary as a function of its position within a longer sequence. In the first experiment, baboons were exposed to sequences composed of a fixed length of 4 elements. An AB regularity systematically appeared at the same position within the sequence on each trial. The sequence of 4 positions was therefore composed of the AB regularity and two additional random elements (X) that were drawn from the 7 remaining possible positions. Three conditions were tested: AB was either presented first, followed by the two random elements (ABXX condition), after two random elements (XXAB condition) or between the two random elements (XABX condition). Baboons were repeatedly administered one of the three conditions at a time each for 500 trials in order to compare the learning rates of AB in each condition. If the position of the regularity in the sequence had an effect on its learning, we expect differences in the decrease in RTs for the predicted B position as a function of its position in the sequence.

In a second experiment, we tested if the learning of an AB regularity would vary as a function of the sequence's length. Baboons were either exposed to 4-element sequences composed of the AB regularity and 2 random elements or to 5-element sequences composed of the AB regularity and 3 random elements. Here, contrary to Experiment 1, AB was not associated to a specific position within the longer sequence and could appear at any position on each trial. We can therefore contrast the learning rates obtained in Experiment 1 in which AB appeared at a fixed

position in sequences of 4 elements with the learning of AB when it appeared randomly at any position in the sequences of 4 (in Experiment 2). We can also contrast the learning rate of AB when it appeared in sequences of 4 or 5 elements.

EXPERIMENT 1

Method

Participants

We tested 20 Guinea baboons (*Papio papio*, 16 females, age range 2.92 – 25 years) living in a social group at the CNRS primate facility in Rousset, France. The baboons were members of a social group of 25 individuals living in a 700-m^2 outdoor enclosure containing climbing structures connected to two indoor experimental areas containing the test equipment. Water was provided *ad libitum* during the test, and the monkeys received their normal food ration of fruits every day at 5 PM.

Apparatus

The baboons had free access to fourteen Automated Learning Devices for Monkeys (ALDM, Fagot & Bonté, 2010; Fagot & Paleressompoulle, 2009) equipped with tactile screens and a food dispenser. Whenever a monkey entered a test chamber, it was identified by its microchip, and the system was prompted to resume the trial list at the place at which the subject left it at its previous visit. The experiment was controlled by E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA).

Materials and procedure

To initiate a trial, the baboon had to touch a yellow cross presented at the bottom of the screen. After the baboon touched it, the yellow cross disappeared, and nine white crosses were displayed, with one of them being replaced by the target, a red circle. When the target circle was

touched, it disappeared and was immediately replaced by the cross. The next position in the sequence was then replaced by a second red target circle until the end of the sequence was reached. When the baboon successfully completed the sequence of touches, it was automatically delivered a reward (grains of dry wheat). If the baboon touched an incorrect location or failed to complete the trial within 5000 ms, a green screen was displayed for 3000 ms to indicate the trial had been failed.

< insert Figure 1 here >

The task began with a familiarization phase during which baboons were presented with random sequences of four positions. For each touch, the response time (RT) between the appearance of the circle and the baboon's touch was recorded (Figure 1). After 500 of random trials, the baboon passed to the first block of experimental trials. They each saw 3 blocks of 500 trials each, one experimental condition being assigned to each block.

In all three conditions, each trial was composed of four touches: two forming the AB regularity which appeared on every trial (in the same position in the sequence), and two that were drawn uniformly from the 7 positions not used in the regularity. For example, if the regularity was 5-1, these two positions would appear in the same order, adjacent to one another, on every trial, and 5 or 1 would not appear again in the sequence.

The position of the regularity in the sequence varied across the three experimental conditions: it appeared in the first position (ABXX), the second position (XABX), or the third position (XXAB). The order of the conditions was counterbalanced across baboons. To avoid learning effect across conditions, each baboon had a different regularity for each condition. These regularities were matched for difficulty using the baseline RTs collected during a previous task where the baboons were presented with 1000 random sequences composed of six touches. RTs for

that task were averaged across all trials for each transition from one position to the next in the sequence. A baseline measure for all possible transitions from one position to another was obtained, yielding a 9 × 9 matrix of mean transition times (calculated over the entire group of monkeys, see Appendix A).

Three AB regularities were then assigned to each baboon with the following constraints. For each baboon, the three pairs could not have baseline RTs having a difference greater than 10ms. No position could be used twice in the three pairs for a given baboon (i.e., if a baboon was assigned the pair 5-1, neither 5 nor 1 could appear in any other pair). The three pairs used for each baboon are presented in Appendix B.

To measure the learning rate across repetitions of the AB regularity, we computed the slope of the regression line fitted to the RTs for the transition time from A to B (i.e., the RT on B) over the course of the 500 trials in each condition. Figure 2 provides an example of this procedure for one baboon and one experimental condition.

< insert Figure 2 here >

Analysis

We adopted a two-step trimming procedure. First, we excluded raw RTs greater than 800 ms. Second, RTs falling more than 2.5 standard deviations away from each baboon's mean for a given block of 100 trials were subsequently excluded (9.88% of data excluded)[1]. With the remaining data, we performed two main analyses which produced convergent results. The first analyses were Bayesian Repeated-Measures ANOVA and the procedure is explained in the next

[1] Inspection of the response times distribution revealed that a majority of responses were produced around 500ms. A smaller group of RTs appeared around 1,000 ms and was likely due to situations in which baboon's response was not recorded by the computer, because their hands were dirty. In this situation, they had to touch the screen again, and longer RTs were recorded (that are on average twice longer compared to the first responses). This is why we have adopted this two-step trimming procedure.

section. The second analyses were linear mixed-effects regression analyses and they are reported in Appendix D.

Bayesian Repeated-Measures ANOVA

For each baboon, the slope was taken of the linear regression fit to the RTs for the transition from A to B over the 500 trials for each condition (ABXX, XABX, XXAB). For the baseline, we used the slope of the RTs over all four touches from the 500 random trials in the first block. The slopes were estimated using the mldivide function in the pracma package in R (Borchers, 2021). Once the slopes had been extracted, they were submitted to a Bayesian repeated-measures ANOVA with condition as the within-subjects factor, followed by post-hoc pairwise Bayesian t-tests. We carried out another Bayesian repeated-measures ANOVA without the random baseline condition to examine whether there was any detectable difference between conditions. All Bayesian testing was carried out in the BayesFactor package for R (Morey & Rouder, 2018). We report Bayes Factors (BF), which quantify the odds of the hypothesis tested (difference of means = 0) compared with the alternative hypothesis (difference of means > 0). BFs of 1 to 3 are considered weak evidence, BFs > 3 positive evidence, BFs > 20 strong evidence, and BFs > 150 very strong evidence (Raftery, 1995). Such Bayesian testing has the advantage of being able to present evidence for either the H_1 (BF) or the H_0 by taking the inverse of the Bayes Factor (1/BF).

Results

Based on the results of the Bayesian comparison, there is decisive evidence that learning took place in all three of the regularity conditions relative to the random trials (BF = 97.18 ± 0.4%, Figure 3). When the random condition was excluded, there was positive evidence for the null (1/BF = 7.24 ± 0.77%)[2]. Each condition had a much steeper negative slope (m = -0.091, -0.086, and -

[2] By taking the inverse of the Bayes Factor (1/BF), we can examine the evidence for the null hypothesis ($\mu_1 = \mu_2$).

0.090 respectively) than the random condition (m = -0.004) . There is strong evidence that each of the conditions is different from the random baseline (indicating that learning took place; ABXX = 11.40 ± 0%, XABX = 15.05 ± 0%, XXAB = 126.07 ± 0%). We find, however, that the conditions do not differ from one another. We find that there is positive evidence for the null hypothesis between all of our learning conditions (ABXX vs XABX = 4.10 ± 0.02%, ABXX vs XABX = 4.29 ± 0.02%, XABX vs XXAB = 4.20 ± 0.02%).

< insert Figure 3 here >

Discussion

In this first experiment, we found that the position of a simple AB regularity in a four-element sequence did not significantly impact the rate at which it is extracted. There was neither an advantage to having the regularity appear at the beginning or end of the trial, nor was there a crowding effect for the middle position, when A and B always appeared at the same positions.

These results suggest that the extraction of an AB regularity that is repeated on every trial at the same position in a sequence is not dependent upon its relative position in the sequence. The presence of random positions before or after the AB regularity does not impact its learning. The present data therefore provide a novel general property concerning associative learning mechanisms in sequences: the position of the regularity in the sequence does not matter.

If this general property is correct then changing the position of the regularity from trial to trial should not have an effect on the learning rate of the regularity. This was tested in Experiment 2 by using sequences of 4 elements including an AB regularity that appeared randomly at any possible position in the sequence on every trial. To test if the length of the sequence would impact the learning rate of the AB regularity, the same procedure was used with a sequence of 5 elements. Increasing the number of random elements within each trial may increase the interference produced by these random elements and slow down the extraction of the AB regularity.

EXPERIMENT 2

Method

Participants and apparatus

Twenty Guinea baboons (*Papio papio*, 13 females, age range 4.42 – 25.25 years) completed this experiment. Sixteen of these also completed Experiment 1. The general task and apparatus used was the same as in Experiment 1.

Materials and procedure

The trial format was the same as in Experiment 1, with the exception that two sequence lengths (4 and 5) were presented to the baboons in two different blocks of 500 trials (the order of the blocks was counterbalanced across baboons). For each sequence length, baboons first saw 200 random trials for (re)familiarization with the task, followed by the 500 experimental trials. Whether the trial was four or five touches, the general format was the same. It contained a single two-element regularity (AB) and either two or three other random touches (X) drawn uniformly from the positions not used in the regularity, as it was the case in Experiment 1.

To avoid learning effects across conditions, each baboon had different regularities for the two sequence lengths. These regularities were matched for difficulty based on the RTs collected during the random trials phase of Experiment 1. Two pairs were assigned to each baboon with the same following constraints as in Experiment 1 and the list of pairs is presented in Appendix C.

By contrast to Experiment 1, instead of a given regularity appearing at the same position in each trial, the regularity appeared in a random position of the sequence on each trial. This was done first by evenly distributing the regularity over the three or four possible positions in the trial (i.e.: ABXX, XABX, XXAB in the four-touch condition and ABXXX, XABXX, XXABX, XXXAB in the five-touch condition). This balanced list of trials was then shuffled such that the regularity could not appear in the same position for more than 4 trials in a row.

Analysis

We used the same exclusion criteria as in Experiment 1: Raw RTs greater than 800 ms were immediately excluded. RTs falling more than 2.5 standard deviations away from each baboon's mean were subsequently excluded (18.55% of data excluded). Given the results of Experiment 1 (i.e., no significant differences between learning conditions), we aggregated across the three conditions from Experiment 1 and treated them as a single "Fixed position" condition in our analysis here.

Bayesian Repeated-Measures ANOVA

For each baboon, the slope was taken of the linear regression fit to the RTs for the transition from A to B over the trials in each condition (Fixed position from Experiment 1, Variable position – length 4, Variable position – length 5). For the baseline, we used the slope of the RTs from the 500 random trials in the first block of Experiment 1. The slopes were estimated using the mldivide function in the pracma package in R (Borchers, 2021). Once the slopes had been extracted, they were submitted to a Bayesian repeated-measures ANOVA with condition as the within-subjects factor, followed by post-hoc pairwise Bayesian t-tests. We carried out another Bayesian repeated-measures ANOVA without the random baseline condition to examine whether there was any detectable difference between conditions. All testing was carried out in the BayesFactor package for R (Morey & Rouder, 2018).

Results

As in Experiment 1, we found strong evidence that the means of all of our conditions were not equal in our omnibus test including the random baseline (BF = 733.61 ± 0.4%; Figure 4). When we conducted our test without the random baseline, we instead found positive evidence for H_0: the conditions do not differ in learning rate (1/BF = 3.42 ± 0.61%). In the pairwise tests, we found

strong evidence that each of the learning conditions (Fixed, Variable-4, and Variable-5) differed from the baseline (BF = 104.56 ± 0%, 111.27 ± 0%, and 53.94 ± 0% respectively). In the pairwise tests between the learning conditions, we found evidence that the two length-four conditions do not differ (1/BF = 3.79 ± 0%), as well as that the two variable position conditions do not differ (1/BF = 3.94 ± 0.02%). The evidence for the relationship between the Fixed condition and the Variable-5 condition is inconclusive (1/BF = 2.04 ± 0%).

We note that the points included in the AB regularity occur at a much higher frequency (every trial) than the random points not included in the regularity (2/7 or 3/7 probability to occur in a given trial). To ensure that the learning effects observed in the RTs from A to B were not simply a function of the higher frequency of B relative to the randomly distributed points, we also examined the RT on A which had an equal frequency of appearance as B but had no (useful/predictable) transition information from any other point on the grid in the variable position conditions. We thus conducted another Bayesian repeated-measures ANOVA which included the random baseline condition and the learning slope from the RT on the first element in the regularity (A) for the variable position conditions in our analysis.

We found positive evidence for the null hypothesis in this analysis (i.e., the learning rate on A was the same as in the random baseline condition; 1/BF = 7.10 ± 0.54%). In follow-up pairwise testing, we found that the learning rate on A in the variable position conditions did not differ either from the random baseline (1/BF = 3.42 ± 0% and 3.17 ± 0% for Var-4 and Var-5 against baseline respectively), or from each other (1/BF = 4.26 ± 0.02%). We also found strong evidence that the learning rate on A is different from that on B in both of these conditions (BF = 91.77 ± 0% and 1685.68 ± 0% respectively). We can interpret this as confirmation that the learning

observed on B in the AB regularity is truly a function of its relationship to A, and not simply a question of relative frequency.

< insert Figure 4 here >

Discussion

In this second experiment, we first found that when the position of the regularity was variable from trial to trial, it was reliably extracted and, more importantly, the learning rate in this variable condition was similar to the fixed condition from Experiment 1. Second, increasing the length of the sequence (and therefore the number of random elements) did not have an effect on the learning rate of the AB regularity.

These results confirm the general property obtained in Experiment 1: the position of the AB regularity in the sequence does not matter. Its repeated occurrence on every trial determines its learning independently of its position in the sequence. The manipulation of sequence length also suggests another general property of associative learning mechanisms: length does not seem to affect regularity extraction. However, this second claim is restricted to the sequence length we manipulated here. It remains possible that with a longer sequence, learning rate could be adversely impacted.

We also confirmed that the learning observed on B was not merely an effect of increased frequency relative to other positions, as it had exactly the same appearance frequency as A, but A was not learned, while B was. We thus provide new empirical evidence that there is no effect of these manipulations on associative learning in this context.

General Discussion

The present study was designed to investigate the role of position and sequence length in the learning of a simple AB regularity inserted in random elements (i.e., noise). Three main results stem from our two experiments.

First, in Experiment 1, AB was inserted at a fixed position in a 4-element sequence, either before, after or between 2 random elements. The AB dependency was progressively learned by baboons, but no differences were observed on the learning rates in these 3 conditions. This first result indicates that the absolute position of the dependency relative to the noise does not affect the extraction of the dependency itself, suggesting the following general property that, during serial learning, individuals learn the relationship between adjacent elements AB independently of the position of the regularity within the sequence.

A second main result stems from Experiment 2 and the manipulation of the position of AB within the sequence. In this second experiment, the position of AB (before, after or between random elements) varied across trials and its position relative to the noise was not a reliable information anymore. Under these conditions, the AB dependency was still extracted and learned by baboons, suggesting that coding for position was not necessary and that varying positions did not hinder learning. Furthermore, the learning rates between fixed and varying positions of AB were not different, suggesting that, even in a fixed design, coding of ordinal positions either does not facilitate learning of a simple AB dependency, or did not even occur at all.

Our third main result deals with the position and amount of noise in the sequence. As already stated, in a 1:1 ratio of signal to noise (i.e., when there are two random elements and two regular elements AB), the position of the regularity relative to the noise did not impact the learning of AB. But this absence of effect remained even with a decreased signal to noise ratio (i.e., three

random elements), indicating how much salient a simple AB dependency was in a random environment.

These results taken together suggest that baboons did not code the ordinal position of reliable elements (i.e., AB) in the sequence. Additionally, the number of randoms in the sequence, and consequently the signal to noise ratio and length of the sequence, did not hinder learning. These data therefore suggest a predominance of adjacency coding mechanisms over serial position coding ones during serial learning.

Acknowledgements

This work was supported by ERC Advanced grant 742141, the BLRI Labex (ANR-11-LABX-0036), Institut Convergence ILCB (ANR-16-CONV-0002), and the CHUNKED ANR project (#ANR-17-CE28-0013-02). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Open Practices Statements

Data, materials and analysis scripts from the experiments are available on Open Science Framework at https://osf.io/q9z2m/.

Conflict of Interest

The authors declare that they have no conflict of interest.

Animal rights

This research adhered to the applicable French rules for ethical treatment of research animals and received ethical approval from the French Ministry of Education (approval APAFIS#2717-2015111708173794 10 v3).

References

Borchers, H. W. (2021). pracma: Practical Numerical Math Functions. R package version 2.3.3. https://CRAN.R-project.org/package=pracma

Christiansen, M. H. (2019). Implicit Statistical Learning : A Tale of Two Literatures. *Topics in Cognitive Science*, *11*(3), 468-481. https://doi.org/10.1111/tops.12332

Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (*Papio papio*). *Behavior research methods*, *42*(2), 507-516.

Fagot, J., & Paleressompoulle, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, *41*(2), 396-404.

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research : A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128-1153. https://doi.org/10.1037/bul0000210

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate : Statistical learning in cotton-top tamarins. *Cognition*, *78*(3), B53-64. https://doi.org/10.1016/s0010-0277(00)00132-3

Malassis, R., Rey, A., & Fagot, J. (2018). Non-adjacent Dependencies Processing in Human and Non-human Primates. *Cognitive Science*, *42*(5), 1677-1699. https://doi.org/10.1111/cogs.12617

Minier, L., Fagot, J., & Rey, A. (2016). The Temporal Dynamics of Regularity Extraction in Non-Human Primates. *Cognitive Science*, *40*(4), 1019-1030. https://doi.org/10.1111/cogs.12279

Morey, R.D. & Rouder, J.N. (2018). BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-4.2. https://CRAN.R-project.org/package=BayesFactor

Nissen, MJ., & Bullemer, P. (1987). Attentional requirements of learning : Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1-32. https://doi.org/10.1016/0010-0285(87)90002-8

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in cognitive sciences*, *10*(5), 233-238.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2016). Retrieved from https://support.pstnet.com/.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological methodology*, 111-163.

Rey, A., Fagot, J., Mathy, F., Lazartigues, L., Tosatto, L., Bonafos, G., Freyermuth, J.-M., & Lavigne, F. (2022). Learning higher-order transitional probabilities in nonhuman primates. *Cognitive Science*, *46*(4).

Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (2019). Regularity Extraction Across Species : Associative Learning Mechanisms Shared by Human and Non-Human Primates. *Topics in Cognitive Science*, *11*(3), 573-586. https://doi.org/10.1111/tops.12343

Tosatto, L., Fagot, J., Nemeth, D., & Rey, A. (2022). The evolution of chunks in sequence learning. Cognitive Science. 46(4), e13124. Doi: 10.1111/cogs.13124

Wilson, B., Smith, K., & Petkov, C. I. (2015). Mixed-complexity artificial grammar learning in humans and macaque monkeys : Evaluating learning strategies. *The European Journal of Neuroscience*, *41*(5), 568-578. https://doi.org/10.1111/ejn.12834

Figure 1

Schematic of a single trial in Experiment 1

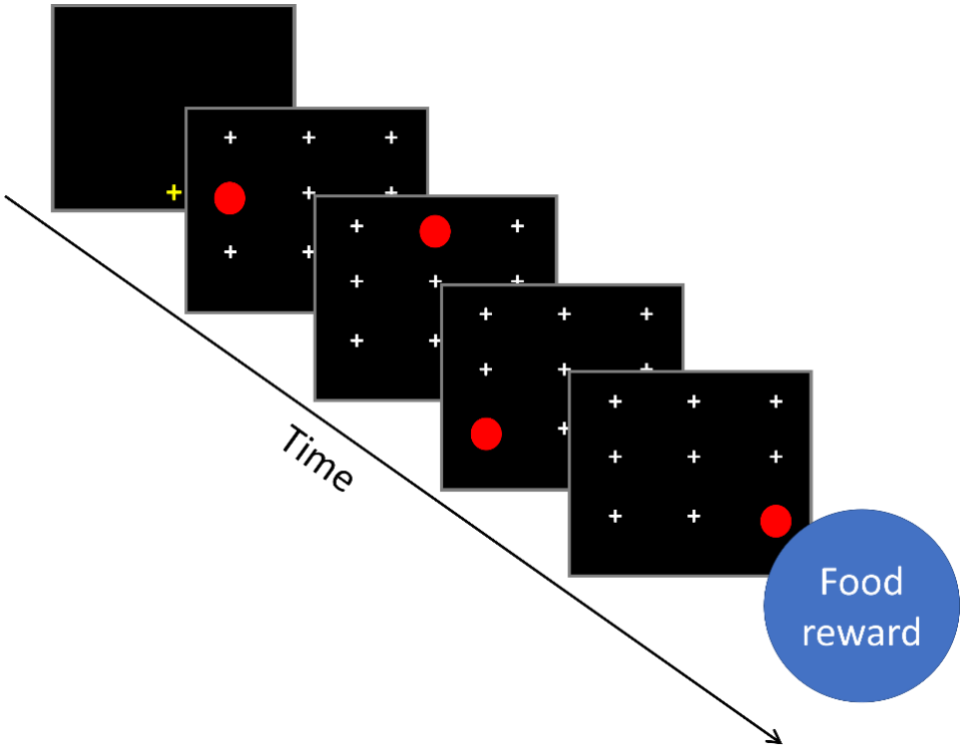


Figure 2

Example of experimental RTs with corresponding regression lines in the random (training) and XABX conditions for a single baboon.

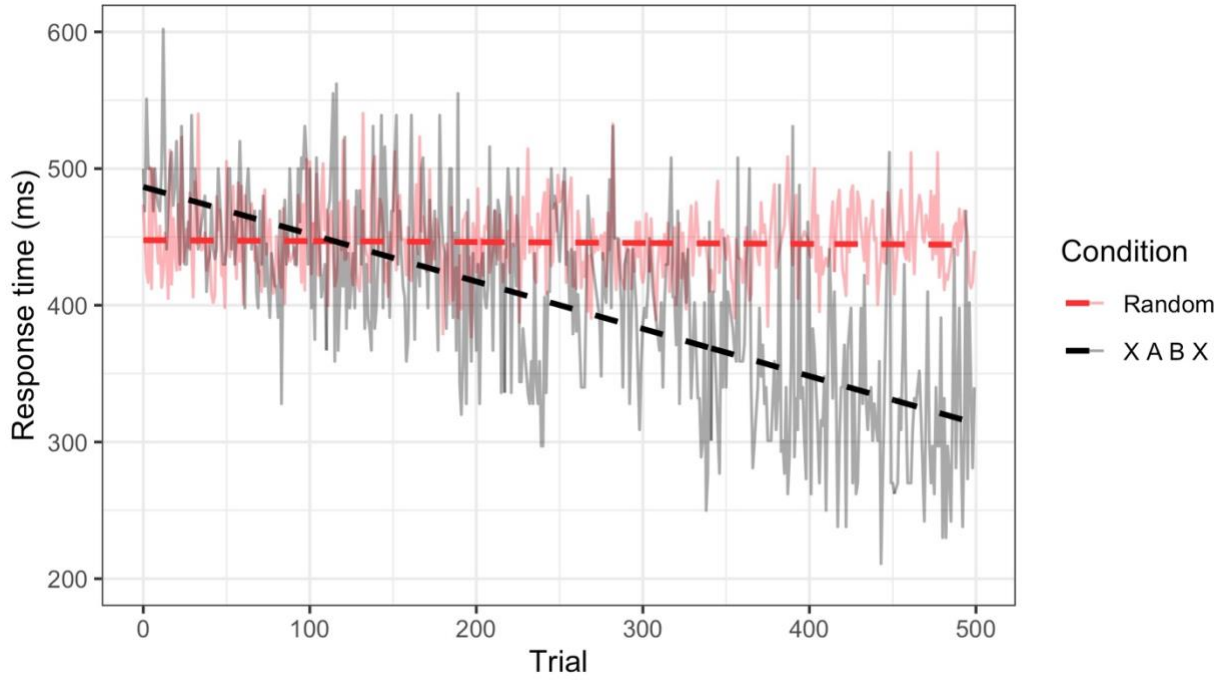
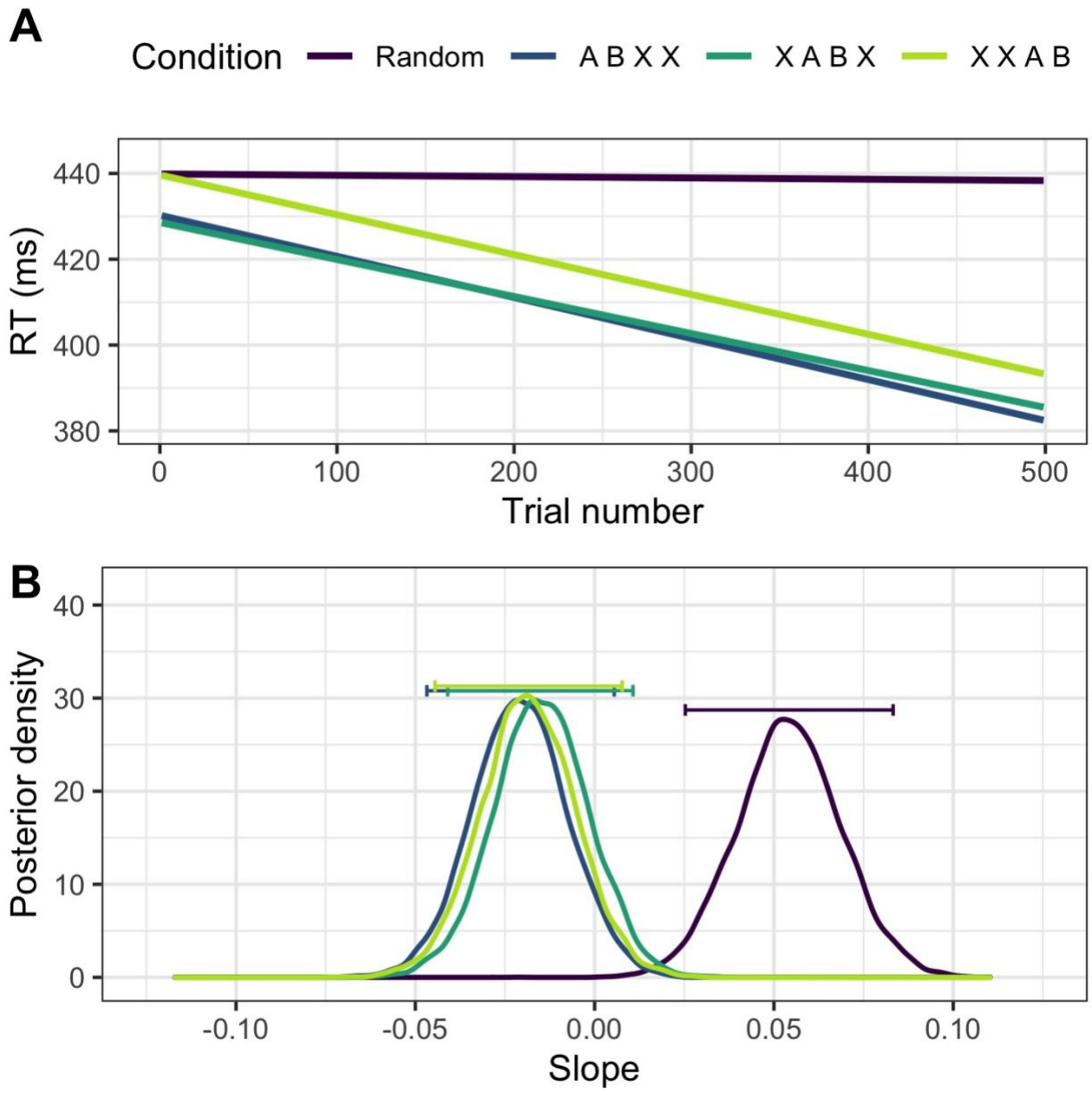


Figure 3

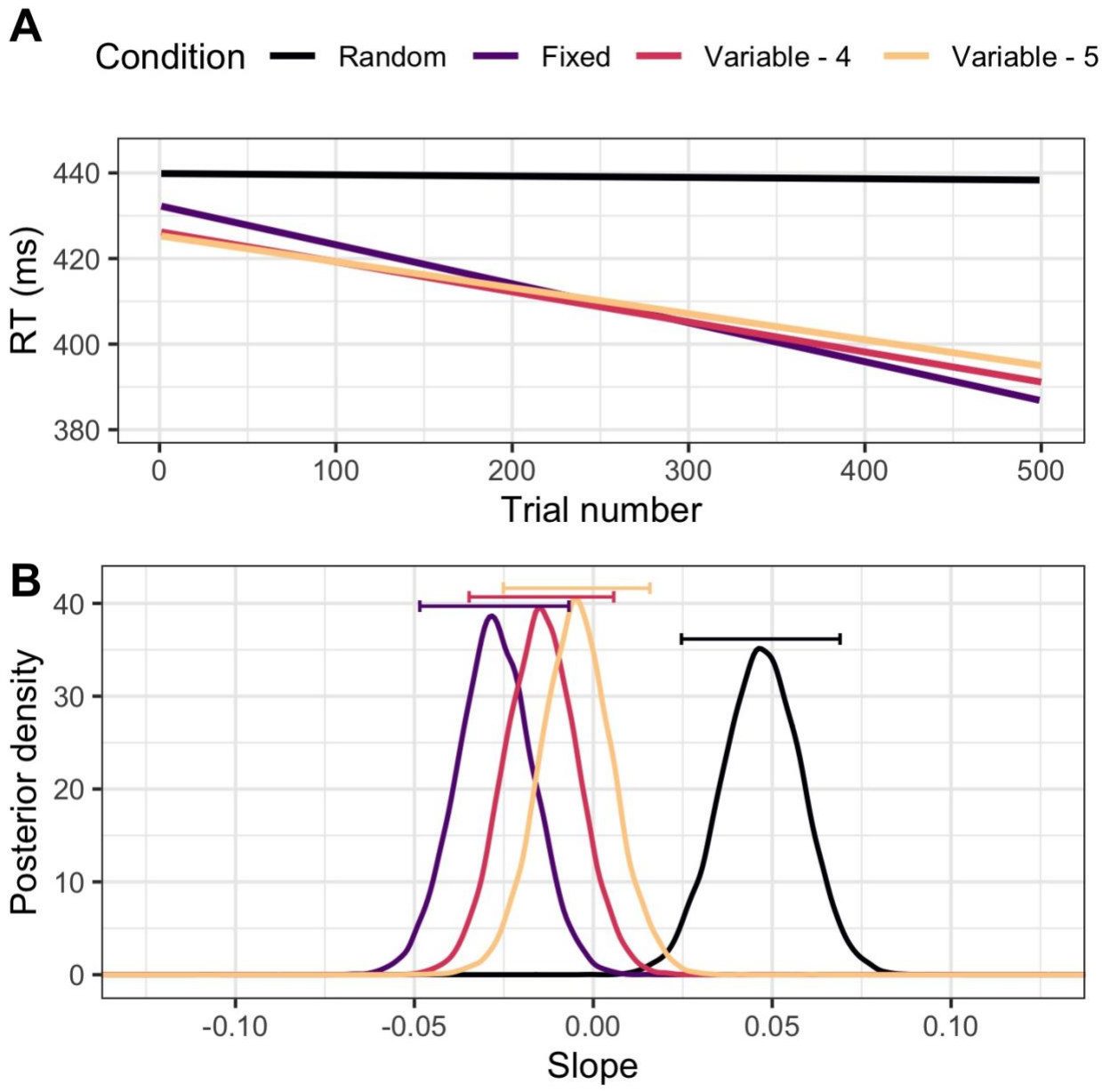
Results of Experiment 1



Note. A. Regression lines for the experimental conditions averaged over all participants. B. Posterior distributions for slopes in the Random and three positional conditions. Horizontal bars show 95% of posterior estimates.

Figure 4

Results of Experiment 2



Note. A. Regression lines for the experimental conditions averaged over all participants. Note that the Fixed condition represents the aggregate of the three experimental conditions from experiment 1. B. Posterior distributions for slopes in the Random and experimental conditions. Horizontal bars show 95% of posterior estimates.

Appendixes

Appendix A

Mean response times over the entire group of baboons for each of the 72 possible transitions calculated from the 1000 random trials

start	1	2	3	4	5	6	7	8	9
1	NA	519	573	495	482	509	521	497	543
2	569	NA	553	513	474	511	523	491	509
3	558	519	NA	513	472	488	544	493	512
4	551	517	560	NA	464	509	522	482	546
5	549	504	552	501	NA	483	535	479	527
6	567	515	546	507	484	NA	533	483	511
7	555	504	558	475	463	516	NA	484	541
8	554	512	540	485	448	472	512	NA	507
9	546	512	540	514	460	464	550	485	NA

Note. All transitions are in milliseconds (ms) and correspond to the time elapsed between the disappearance of the red circle from the 1st position of the Transition and the monkey's touch on the 2nd position of the Transition.

Appendix B

Repartition of AB pairs learned by each baboon in Experiment 1 and corresponding baseline mean transition times

Name	ID	Position order	ABXX	XABX	XXAB
ANGELE	1	2, 3, 1	8, 7	9, 2	4, 6
ARIELLE	2	3, 2, 1	1, 7	3, 9	6, 2
ATMOSPHERE	3	2, 3, 1	3, 6	1, 5	4, 8
DORA	4	3, 1, 2	6, 4	2, 9	8, 7
DREAM	5	1, 2, 3	9, 2	1, 6	3, 4
EWINE	6	2, 1, 3	6, 8	7, 4	2, 5
FANA	7	3, 1, 2	6, 3	4, 9	5, 1
FELIPE	8	3, 1, 2	8, 2	3, 9	4, 6
FEYA	9	1, 3, 2	4, 1	2, 3	9, 7
FLUTE	10	2, 3, 1	6, 9	3, 2	8, 7
HARLEM	11	2, 1, 3	3, 7	4, 9	8, 1
KALI	12	3, 2, 1	1, 7	3, 9	6, 2
LIPS	13	1, 2, 3	3, 2	8, 7	4, 6
LOME	14	2, 1, 3	5, 2	8, 9	3, 4
MAKO	15	1, 3, 2	6, 4	5, 2	1, 8
MALI	16	1, 3, 2	3, 4	6, 9	1, 7
MUSE	17	2, 1, 3	5, 4	1, 6	8, 9
NEKKE	18	1, 2, 3	6, 9	4, 2	1, 7
PETOULETTE	19	2, 3, 1	6, 9	8, 2	1, 7
VIOLETTE	20	3, 1, 2	9, 4	1, 6	3, 2

Appendix C

Repartition of AB pairs learned by each baboon in Experiment 2 and corresponding baseline mean transition times

Name	ID	Var-4	Var-5
ANGELE	1	6, 4	1, 7
ARIELLE	2	6, 9	3, 5
ATMOSPHERE	3	1, 2	4, 7
EWINE	6	9, 7	3, 6
FANA	7	1, 5	2, 8
FELIPE	8	4, 6	9, 2
FEYA	9	6, 4	8, 7
HARLEM	11	8, 4	5, 9
KALI	12	6, 7	8, 3
LIPS	13	4, 8	5, 6
LOME	14	1, 7	5, 4
MAKO	15	8, 9	4, 2
MALI	16	3, 4	8, 7
MUSE	17	9, 1	6, 2
PETOULETTE	19	2, 6	7, 3
VIOLETTE	20	7, 5	2, 8
ARTICHO	21	5, 6	4, 8
BOBO	22	8, 9	3, 7
HERMINE	23	9, 4	1, 6
PIPO	24	5, 3	1, 2

Appendix D

Results from Linear Mixed-Effects modeling analysis

The same model was fit with each of the four conditions (`cond`) from Experiment 1 as the baseline:

Variable name	Condition name
rnd	random baseline condition
pos1	ABXX
pos2	XABX
pos3	XXAB

Because models which used `trial number` as a variable did not converge, trials were grouped into groups of 20 and this is the slope term in the models (`trl20`).

For the other model terms, `rtVal` is the response time in milliseconds, `name` is the identifier for each baboon, and `ptPair` are the unique pairs of start and stop points on the screen.

In each case, we find that a) all conditions have a negative slope above chance except the random condition, and b) the regularity conditions (ABXX, XABX, XXAB) differ from the random baseline condition, but do not significantly differ from one another. While this presents promising evidence, the absence of an effect is not the same as evidence for the null, which is why we switched our analysis over to the Bayesian framework.

Random condition as baseline

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: rtVal ~ cond * trl20 + (trl20 + cond | name) * (1 | ptPair)
##    Data: .
## Control: lmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
## REML criterion at convergence: 647288.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8971 -0.6658 -0.1238  0.5115  6.0512
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  ptPair   (Intercept)  459.9602 21.4467
##  name     (Intercept)  356.7099 18.8868
##           trl20          0.4932  0.7023  -0.33
##           condpos1     729.7698 27.0143   0.15  0.33
##           condpos2     509.6477 22.5754  -0.08 -0.06  0.49
##           condpos3     923.9344 30.3963  -0.19 -0.06 -0.16 -0.28
##  Residual             2380.6093 48.7915
## Number of obs: 60916, groups:  ptPair, 81; name, 20
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     4.407e+02  4.878e+00  3.235e+01  90.349   <2e-16 ***
## condpos1       -1.530e+01  6.178e+00  2.010e+01  -2.477   0.0222 *
## condpos2       -1.076e+01  5.218e+00  2.075e+01  -2.062   0.0519 .
## condpos3       -6.420e+00  6.934e+00  1.993e+01  -0.926   0.3656
## trl20          -6.632e-02  1.614e-01  1.993e+01  -0.411   0.6854
## condpos1:trl20 -1.792e+00  7.984e-02  6.074e+04 -22.450   <2e-16 ***
## condpos2:trl20 -1.657e+00  8.062e-02  6.074e+04 -20.553   <2e-16 ***
## condpos3:trl20 -1.751e+00  8.212e-02  6.074e+04 -21.323   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) cndps1 cndps2 cndps3 trl20  cn1:20 cn2:20
## condpos1     0.120
## condpos2    -0.081  0.471
## condpos3    -0.169 -0.142 -0.254
## trl20       -0.298  0.329 -0.035 -0.045
## cndps1:tr20  0.043 -0.156 -0.040 -0.030 -0.106
## cndps2:tr20  0.042 -0.034 -0.186 -0.030 -0.105  0.213
## cndps3:tr20  0.041 -0.033 -0.039 -0.143 -0.103  0.208  0.206
```

ABXX condition as baseline

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: rtVal ~ cond * trl20 + (trl20 + cond | name) + (1 | ptPair)
##    Data: .
## Control: lmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
## REML criterion at convergence: 647288.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8971 -0.6658 -0.1238  0.5115  6.0512
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  ptPair   (Intercept)  459.9562 21.4466
##  name     (Intercept) 1241.8690 35.2402
##           trl20          0.4932  0.7023   0.08
##           condpos2     645.7056 25.4107  -0.61 -0.40
##           condpos3    1911.8915 43.7252  -0.68 -0.25  0.34
##           condrnd      729.5335 27.0099  -0.85 -0.33  0.63  0.73
##  Residual             2380.6095 48.7915
## Number of obs: 60916, groups:  ptPair, 81; name, 20
##
## Fixed effects:
##                  Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     4.254e+02  8.316e+00  2.306e+01  51.156  < 2e-16 ***
## condpos2        4.541e+00  5.920e+00  2.070e+01   0.767   0.4517
## condpos3        8.883e+00  9.919e+00  1.967e+01   0.896   0.3813
## condrnd         1.530e+01  6.177e+00  2.012e+01   2.478   0.0222 *
## trl20          -1.859e+00  1.722e-01  2.588e+01 -10.792 4.49e-11 ***
## condpos2:trl20  1.354e-01  1.007e-01  6.074e+04   1.345   0.1786
## condpos3:trl20  4.127e-02  1.019e-01  6.074e+04   0.405   0.6855
## condrnd:trl20   1.792e+00  7.984e-02  6.074e+04  22.450  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) cndps2 cndps3 cndrnd trl20  cn2:20 cn3:20
## condpos2    -0.582
## condpos3    -0.649  0.338
## condrnd     -0.813  0.629  0.722
## trl20        0.023 -0.290 -0.186 -0.235
## cndps2:tr20  0.072 -0.204 -0.060 -0.097 -0.289
## cndps3:tr20  0.071 -0.100 -0.124 -0.096 -0.285  0.488
## cndrnd:tr20  0.091 -0.128 -0.076 -0.156 -0.364  0.623  0.616
```

XABX condition as baseline

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: rtVal ~ cond * trl20 + (trl20 + cond | name) + (1 | ptPair)
##    Data: .
## Control: lmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
## REML criterion at convergence: 647288.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8971 -0.6658 -0.1238  0.5115  6.0512
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  ptPair   (Intercept)  459.9527 21.4465
##  name     (Intercept)  795.7343 28.2088
##           trl20          0.4932  0.7023  -0.27
##           condpos3    1813.0592 42.5800  -0.64 -0.01
##           condrnd      509.7522 22.5777  -0.74  0.06  0.73
##           condpos1     645.8306 25.4132  -0.14  0.40  0.25  0.37
##  Residual             2380.6090 48.7915
## Number of obs: 60916, groups:  ptPair, 81; name, 20
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     4.300e+02  6.851e+00  2.547e+01  62.762  < 2e-16 ***
## condpos3        4.342e+00  9.682e+00  1.974e+01   0.448   0.6587
## condrnd         1.076e+01  5.219e+00  2.074e+01   2.062   0.0519 .
## condpos1       -4.541e+00  5.921e+00  2.069e+01  -0.767   0.4517
## trl20          -1.723e+00  1.726e-01  2.610e+01  -9.983 2.11e-10 ***
## condpos3:trl20 -9.416e-02  1.025e-01  6.074e+04  -0.918   0.3584
## condrnd:trl20   1.657e+00  8.062e-02  6.074e+04  20.553  < 2e-16 ***
## condpos1:trl20 -1.354e-01  1.007e-01  6.074e+04  -1.345   0.1786
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) cndps3 cndrnd cndps1 trl20  cn3:20 cnd:20
## condpos3    -0.605
## condrnd     -0.704  0.721
## condpos1    -0.158  0.265  0.390
## trl20       -0.275  0.024  0.119  0.409
## cndps3:tr20  0.088 -0.127 -0.115 -0.101 -0.290
## cndrnd:tr20  0.111 -0.079 -0.186 -0.129 -0.369  0.621
## cndps1:tr20  0.089 -0.063 -0.117 -0.204 -0.295  0.497  0.632
```

XXAB condition as baseline

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: rtVal ~ cond * trl20 + (trl20 + cond | name) + (1 | ptPair)
##    Data: .
## Control: lmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
## REML criterion at convergence: 647288.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8971 -0.6658 -0.1238  0.5115  6.0512
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  ptPair   (Intercept)  459.9519 21.4465
##  name     (Intercept) 1062.9213 32.6025
##           trl20          0.4933  0.7023  -0.25
##           condrnd      923.8909 30.3956  -0.82  0.06
##           condpos1    1912.3018 43.7299  -0.61  0.25  0.79
##           condpos2    1812.5847 42.5745  -0.75  0.01  0.86  0.83
##  Residual             2380.6092 48.7915
## Number of obs: 60916, groups:  ptPair, 81; name, 20
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     4.343e+02  7.773e+00  2.376e+01  55.872  < 2e-16 ***
## condrnd         6.420e+00  6.934e+00  1.994e+01   0.926    0.366
## condpos1       -8.883e+00  9.920e+00  1.966e+01  -0.896    0.381
## condpos2       -4.342e+00  9.681e+00  1.975e+01  -0.448    0.659
## trl20          -1.817e+00  1.733e-01  2.654e+01 -10.485 6.21e-11 ***
## condrnd:trl20   1.751e+00  8.212e-02  6.074e+04  21.323  < 2e-16 ***
## condpos1:trl20 -4.127e-02  1.019e-01  6.074e+04  -0.405    0.686
## condpos2:trl20  9.416e-02  1.025e-01  6.074e+04   0.918    0.358
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) cndrnd cndps1 cndps2 trl20  cnd:20 cn1:20
## condrnd     -0.786
## condpos1    -0.581  0.787
## condpos2    -0.712  0.853  0.818
## trl20       -0.260  0.110  0.257  0.051
## cndrnd:tr20  0.101 -0.143 -0.079 -0.082 -0.378
## cndps1:tr20  0.082 -0.092 -0.124 -0.066 -0.305  0.643
## cndps2:tr20  0.081 -0.091 -0.064 -0.127 -0.303  0.639  0.515
```