Felix's experiment

March 27, 2024

1 Methods

To estimate the effect of condition (solo, human-computer dyad, human-human dyad), random dot pattern coherence, stimulus duration, stimulus number, block number, and day number on the performance of participants we fitted a series if six models, three of them with data from the solo condition and the human-computer dyad condition, and three of them with data from the human-human dyad condition. The reason for splitting the data was that we needed to control for the identity of two humans in the human-human dyad condition but only for the identity of one human in the other conditions. In the following we first describe the analysis of the data from the solo and the human-computer dyad condition, and then the analysis of the data from the human-human dyad condition.

1.1 Solo and human-computer dyad condition; model 1

We fitted three Generalized Linear Mixed Models (GLMM; Baayen, 2008) which differed in their response variable and the size of the data set analyzed but had identical fixed effects structures and largely identical random effects structures. We fitted one model for each of the probability of a target hit (model 1a), joystick eccentricity (model 1b), and joystick accuracy (model 1c) as the response. All three aimed at estimating the extent to which the respective response variable was affected by the fixed effects of condition (solo or human-computer dyad), random dot pattern coherence, stimulus duration, stimulus number, block number, and day number. As we hypothesized that the effect of coherence could depend on the condition we also included the interaction between these two predictors into the fixed affects part of the model. To avoid pseudo-replication and account for the possibility that the response was influenced by several layers of non-independence we included three random intercepts effects, namely those of the ID of the participant, the ID of test day (nested in participant; thereafter 'day ID'), and the ID of the block (nested in participant and day; thereafter 'block ID'). The reason for including the latter two was that it could be reasonably assumed that the performance of participants varied between test days and also between blocks tested on the same day. To avoid an 'overconfident model' and keep type I error rate at the nominal level of 0.05 we included all theoretically random slopes (Schielzeth and Forstmeier, 2009; Barr et al., 2013). These were those of condition, coherence, their interaction, stimulus duration, stimulus number, block number, and day number within participant, coherence, stimulus duration, stimulus number, and block number within day ID, and finally coherence, stimulus duration, and stimulus number within block ID. Originally we also included estimates of the correlations among random intercepts and slopes into each model, but do to convergence and identifiability problems (recognizable by absolute correlation parameters being close to 1; Matuschek et al., 2017) we had to exclude all or several of these estimates from the full models (see Table SI 1 for detailed information).

For each model we conducted a full-null model comparison which aims at avoiding 'cryptic multiple testing' and keep type one error rate at the nominal level of 0.05 (Forstmeier and Schielzeth, 2011). As we had a genuine interest in all predictors present in the fixed effects part of each model the null models comprised only the intercept in the fixed effects part but were otherwise identical to the respective full model. This full-null model comparison utilized a likelihood ratio test (Dobson, 2002). Tests of individual effects were also based on likelihood ratio test, comparing a full model with a each in a set of reduced models which lacked fixed effects one at a time.

1.2 Dyadic condition; model 2

To investigate performance in the dyadic condition in which two humans played simultaneously, we fitted three models roughly corresponding to models 1a, b, and c. The main difference between models 2 and 1 was that they did not comprise a fixed effect nor random slopes of condition. This was because after subsetting the data two the human-human dyad trials, only one condition was left. Furthermore, we adjusted the random effects structure to account for the fact that there were always two humans playing simultaneously that could influence their behaviour. To this end we included additional random intercepts effects for the ID of the other player, and the dyad of the two players (which was identical, irrespective of which of the two players was considered the 'player' and which the 'other player'). Since for each particular stage we had measured the behavior of two players, each stage was represented by two rows in the respective data sets. To account for the fact that the respective measures were taken simultaneously from two players which could influence one another, we included an additional random intercepts effect for the ID of the trial (nested in the player-other-player dyad). Finally, we this time also included a random intercepts effect for the ID of the trial nested in player[1]. Again, we included all theoretically identifiable random slopes. However, due to convergence problems, we had to exclude some of the random slopes from model 2 b and c. We excluded estimation of parameters for the correlations among random intercepts and slopes when convergence problems appeared or when were seemingly unidentifiable (see above). The full model structure is depicted in Table .

1.3 Implementation

We fitted all models in R (version 4.3.3; R Core Team, 2023). In model 1a and 2a we included the response as a two columns matrix with the number of targets hit and not hit in the first and second column respectively (Baayen, 2008). The models were then fitted with a binomial error structure and logit link function (McCullagh and Nelder, 1989). In essence, such models model the proportion of targets hit. We are aware that in principle one would need an 'observation level random effect' which would link the number of targets hit and not hit in a given stage. However, in a relatively large proportion of stages (19.7%) there was only a single target that appeared and in the majority (47.0%) of stages only two targets appeared, making it unlikely that a respective random effect can be fitted successfully. Model 1b, 1c, 2b, and 2c we fitted with a beta error distribution and logit link function (Bolker, 2008). Models fitted with a beta error distribution cannot cope with values in the response being exactly 0 or 1. Hence, when such values were present in a given response variable we transformed them as suggested by Smithson and Verkuilen (2006).

Models 1a and 2a we fitted using the function glmer of the package lme4 (version 1.1-35.1; Bates et al., 2015), and models 1b, 1c, 2b, and 2c we fitted using the function glmmTMB of the equally named package (version 1.1.8; Brooks et al., 2017). We determined model stability by dropping individual levels of the random effects factors from the data, one at a time, fitting the full model to each of the subsets, and finally comparing the range of fixed effects estimates obtained for the subsets with those obtained for the model fitted on the respective full data set. This revealed all models to be of good stability. We estimated 95% confidence limits of model estimates and fitted values by means of parametric bootstraps (N=1000 bootstraps; function bootMer of the package lme4 for model 1 and function simulate of package glmmTMB for models 2 and 3). In none of the models the response was overdispersed (maximum dispersion parameter: 1.0). Sample sizes for all models are indicated in Table . The number of targets hit and missed were 130941 and 74151 (model 1a) and 120256 and 72875 (model 2a), respectively.

[1]Felix, this is a huge inconsisteny between model 1a and 2a as we didn't do so in model 1... We need to discuss how to proceed. I feel we should maybe drop the respective term from model 2a

Table SI 1. Random effects structure and sample size of the models.

model	random effects structure	sample size
1a (prob. of hit) (solo)	(1 + cond*coh + dur + stim + block + day|ID) + (1 + coh + dur + stim + block|dayID) + (1 + coh + dur + stim||blockID)	88125 stimulus states 38 participants 104 day IDs 220 block IDs
1b (eccentricity) (solo)	(1 + cond*coh + dur + stim + block + day||ID) + (1 + coh + dur + stim + block||dayID) + (1 + coh + dur + stim||blockID)	107155 stimulus states 38 participants 104 day IDs 220 block IDs
1c (accuracy) (solo)	(cond*coh + dur + stim + block + day||ID) + (1 + coh + dur + stim + block||date.in.ID) + (1|blockID)	24944 stimulus states 38 participants 104 day IDs 220 block IDs
2a (prob. of hit) (dyadic)	(1 + coh + stim + dur + block + day||ID) + (1 + coh + stim + dur + block + block||dyad) + (1 + coh + stim + dur + block + block||other_player) + (1 + coh + stim + dur||date.in.ID) + (1 + coh + stim + dur + block||block.in.date.in.ID) + (1|trial.ID) + (1|olre)	81612 stimulus states 34 participants 34 other players 50 dyads 99 day IDs 202 block IDs 42135 block IDs 81612 OLREs
2b (eccentricity) (dyadic)	(1 + coh + dur + day||ID) + (1 + coh + stim + block||dyad) + (1 + block||other_player) + (1 + coh + dur||date.in.ID) + (1 + coh + dur + stim. nr.||block.in.date.in.ID) + (1|trial.ID)	100800 stimulus states 34 participants 34 other players 50 dyads 99 day IDs 202 block IDs 51916 trial IDs
2c (accuracy) (dyadic)	(1 + coh + block + day||ID) + (1 + stim||dyad) + (1 + day||other_player) + (1|date.in.ID) + (1 + dur||block.in.date.in.ID) + (1|trial.ID)	23208 stimulus states 34 participants 34 other players 50 dyads 99 day IDs 202 block IDs 12062 trial IDs

| indicates that estimates for the correlations among random intercepts were included, || that they were not; abbreviations: cond: condition; coh: coherence; dur: stimulus duration; stim: stimulus number; block: block number; day: day number; cond*coh means condition, coherence, and their interaction; OLRE abbreviates observation level random effect.

2 Results

2.1 Solo and human-computer dyad condition; model 1

In case of target hit probability model (model 1a) we found a clearly significant full-null model comparison (likelihood ratio test: $\chi^2 = 139.972, df = 7, P<0.001$), and also the interaction between condition and coherence was clearly significant (Table SI 2). Plotting the model and the data, revealed the target hit probability increased steeper in the solo condition as compared to the computer condition (Fig. SI 1). Furthermore, target hit probability increased with stimulus duration (Fig. 2), stimulus number Fig. SI 3, block number (Fig. SI 4), and over the days the experiment commenced (Fig. SI 5).

Also for the model with eccentricity being the response (model 1b) we found a clearly significant full-null model comparison ($\chi^2 = 158.883, df = 7, P<0.001$). Again also the interaction between coherence and condition was clearly significant (Table SI 3), and, with the exception of block number, also all other terms in this model revealed significance. Plotting the data and the fitted model revealed that eccentricity was about the same in the in the solo and the computer condition when the coherence of the random dot pattern was low but increased a little steeper in the solo condition (Fig. SI 6). Furthermore, eccentricity slightly increased with increasing stimulus duration (Fig. SI 7), slightly decreased with stimulus number (Fig. SI 8), slightly but not significantly increased with block number (Fig. SI 9), and increased clearly over the course of the experiment (Fig. SI 10).

Finally, also in case of the accuracy model (model 1c), the full-null model comparison was clearly significant ($\chi^2 = 302.247, df = 7, P<0.001$). With the exception of stimulus number, all effects in this model revealed significance (Table SI 4). Plotting the data and the fitted model revealed that in the solo condition accuracy was lower than in the computer as compared to the solo condition when the coherence of the random dot pattern was low but that accuracy was about the same in both conditions when coherence was large (Fig SI 11). Furthermore, accuracy increased slightly with stimulus duration (Fig. SI. 12), did hardly change with increasing stimulus number (Fig. SI 13), increased slightly with block number (Fig. SI 14), and clearly over the course of the experiment (Fig. SI 15).

2.2 dyadic condition; model 2

In case of the dyadic models, two full-model comparisons were significant (hit rate, model 2a: $\chi^2 = 181.069, df = 5, P < 0.001$; eccentricity, model 2b: null model did not converge; accuracy, model 2c: $\chi^2 = 218.208, df = 19, P < 0.001$). Regarding hit rate, we found that it increased with stimulus coherence (Fig. SI 16; Table SI 5) and stimulus duration (Fig. SI 17), and it increased very slightly with day number (Fig. SI 18). Eccentricity clearly increased with stimulus Coherence (Fig. SI 19; Table SI 6) and slightly increased with stimulus duration (Fig. SI 20) and block number (Fig. SI 21). Accuracy, finally, clearly increased with stimulus coherence (Fig. SI 23; Table SI 7) and slightly increased with stimulus duration (Fig. SI 22) and block number (Fig. SI 24).

Table SI 2. Results of the full model with hit probability being the response (conditions: solo and computer, model 1a)

term	estimate	SE	CL_{lower}	CL_{upper}	χ^2	df	P	min	max
intercept	0.505	0.017	0.473	0.540				0.498	0.510
condition	0.238	0.015	0.208	0.266				0.230	0.244
coherence	0.135	0.009	0.118	0.154				0.131	0.139
stim. duration	0.076	0.005	0.066	0.085	91.707	1	<0.001	0.074	0.077
stim. nr.	0.008	0.005	-0.002	0.017	4.532	1	0.033	0.006	0.009
block nr.	0.016	0.006	0.005	0.028	6.466	1	0.011	0.013	0.020
day nr.	0.017	0.008	0.001	0.032	3.862	1	0.049	0.011	0.020
condition:coherence	-0.064	0.011	-0.087	-0.041	26.472	1	<0.001	-0.067	-0.061

indicated are estimates, together with their standard errors, confidence limits, significance tests, and the range of estimates when excluding individuals one at a time; all covariates were z-transformed, mean and standard of the original variables were 0.338 and 0.320 (coherence), 246.129 and 142.150 (stim. nr.), 1907.127 and 358.333 (stim. dur.), 1.669 and 0.754 (block nr.), and 2.982 and 1.526 (day nr.) respectively; condition was dummy coded with CPRsolo being the reference level

Table SI 3. Results of the full model with eccentricity being the response (conditions: solo and computer, model 1b)

term	estimate	SE	CL_{lower}	CL_{upper}	χ^2	df	P	min	max
intercept	0.682	0.083	0.568	0.884				0.640	0.709
condition	-0.035	0.051	-0.147	0.018				-0.058	-0.003
coherence	0.431	0.024	0.385	0.461				0.421	0.443
stim. nr.	-0.023	0.009	-0.035	-0.010	5.521	1	0.019	-0.026	-0.018
stim. duration	0.028	0.004	0.019	0.037	31.216	1	<0.001	0.026	0.029
block nr.	0.032	0.023	0.019	0.072	1.762	1	0.184	0.018	0.046
day nr.	0.080	0.032	0.044	0.128	5.587	1	0.018	0.062	0.096
condition:coherence	-0.157	0.020	-0.200	-0.127	36.335	1	<0.001	-0.164	-0.149

indicated are estimates, together with their standard errors, confidence limits, significance tests, and the range of estimates when excluding individuals one at a time; all covariates were z-transformed, mean and standard of the original variables were 0.338 and 0.320 (coherence), 245.704 and 142.078 (stim. nr.), 1878.920 and 360.448 (stim. dur.), 1.669 and 0.754 (block nr.), and 2.978 and 1.525 (day nr.) respectively; condition was dummy coded with CPRsolo being the reference level

Table SI 4. Results of the full model with accuracy being the response (conditions: solo and computer, model 1c)

term	estimate	SE	CL_{lower}	CL_{upper}	χ^2	df	P	min	max
intercept	1.354	0.035	1.283	1.418				1.345	1.373
condition	0.208	0.028	0.157	0.262				0.200	0.219
coherence	0.519	0.013	0.493	0.541				0.515	0.529
stim. nr.	0.010	0.007	-0.004	0.023	2.249	1	0.134	0.009	0.015
stim. duration	0.026	0.006	0.013	0.039	16.033	1	<0.001	0.024	0.029
block nr.	0.040	0.007	0.026	0.055	22.652	1	<0.001	0.038	0.044
day nr.	0.060	0.012	0.036	0.084	14.045	1	<0.001	0.053	0.066
condition:coherence	-0.251	0.024	-0.294	-0.203	51.005	1	<0.001	-0.260	-0.244

indicated are estimates, together with their standard errors, confidence limits, significance tests, and the range of estimates when excluding individuals one at a time; all covariates were z-transformed, mean and standard of the original variables were 0.339 and 0.321 (coherence), 245.815 and 142.924 (stim. nr.), 1979.819 and 334.555 (stim. dur.), 1.671 and 0.753 (block nr.), and 3.007 and 1.534 (day nr.) respectively; condition was dummy coded with CPRsolo being the reference level

Table SI 5. Results of the full model with hit probability being the response (condition: dyadic, model 2a)

term	estimate	SE	χ^2	df	P
intercept	0.505	0.014			
coherence	0.119	0.006	85.632	1	<0.001
stim. nr.	0.006	0.005	1.431	1	0.232
stim. duration	0.068	0.005	85.006	1	<0.001
block nr.	0.005	0.005	0.981	1	0.322
day nr.	0.022	0.007	8.000	1	0.005

indicated are estimates, together with their standard errors, confidence limits, significance tests, and the range of estimates when excluding individuals one at a time; all covariates were z-transformed, mean and standard of the original variables were 0.341 and 0.320 (coherence), 251.508 and 145.935 (stim. nr.), 1909.733 and 357.622 (stim. dur.), 1.541 and 0.565 (block nr.), and 3.711 and 1.922 (day nr.) respectively. [confidence limits and stability not yet available].

Table SI 6. Results of the full model with eccentricity being the response (condition: dyadic, model 2b)

term	estimate	SE	*z or* χ^2	df	P
intercept	0.915	0.089			
coherence	0.405	0.027	14.996		<0.001
stim. nr.	0.010	0.009	1.105		0.269
stim. duration	0.026	0.005	20.439	1	<0.001
block nr.	0.074	0.022	10.165	1	0.001
day nr.	0.073	0.047	2.466	1	0.116

indicated are estimates, together with their standard errors, confidence limits, significance tests, and the range of estimates when excluding individuals one at a time; all covariates were z-transformed, mean and standard of the original variables were 0.340 and 0.320 (coherence), 251.445 and 145.706 (stim. nr.), 1883.088 and 359.591 (stim. dur.), 1.540 and 0.563 (block nr.), and 3.712 and 1.923 (day nr.) respectively. Due to convergence problems with two of the reduced models, we report Wald's z- instead of likelihood ratio tests for two effects (recognizable by *df* not indicated). [confidence limits and stability not yet available].

Table SI 7. Results of the full model with accuracy being the response (condition: dyadic, model 2c)

term	estimate	SE	χ^2	df	P
intercept	0.505	0.014			
coherence	0.119	0.006	85.632	1	<0.001
stim. nr.	0.006	0.005	1.431	1	0.232
stim. duration	0.068	0.005	85.006	1	<0.001
block nr.	0.005	0.005	0.981	1	0.322
day nr.	0.022	0.007	8.000	1	0.005

indicated are estimates, together with their standard errors, confidence limits, significance tests, and the range of estimates when excluding individuals one at a time; all covariates were z-transformed, mean and standard of the original variables were 0.341 and 0.321 (coherence), 253.005 and 146.759 (stim. nr.), 1985.164 and 335.175 (stim. dur.), 1.546 and 0.569 (block nr.), and 3.725 and 1.924 (day nr.) respectively. [confidence limits and stability not yet available].

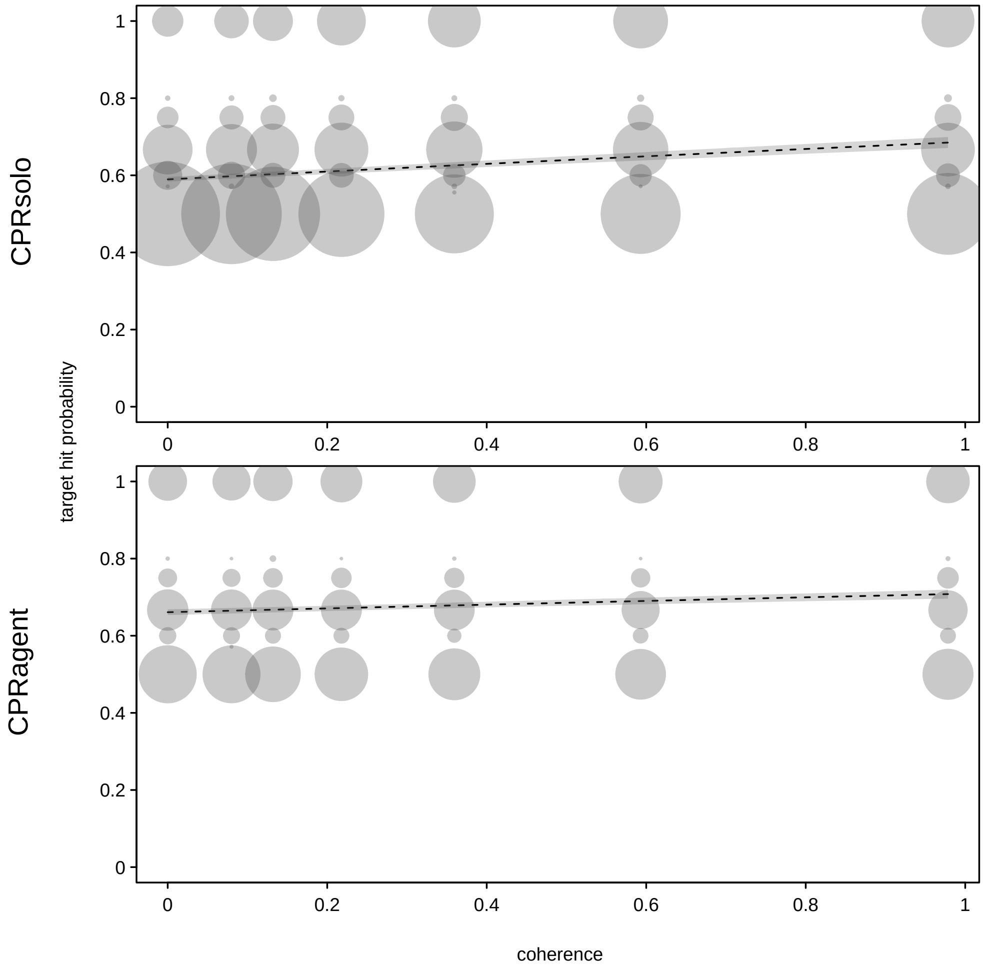

Figure SI 1. Target hit probability as a function of condition and coherence (model 1). Dashed lines and shaded areas depict the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 5 to 15602).

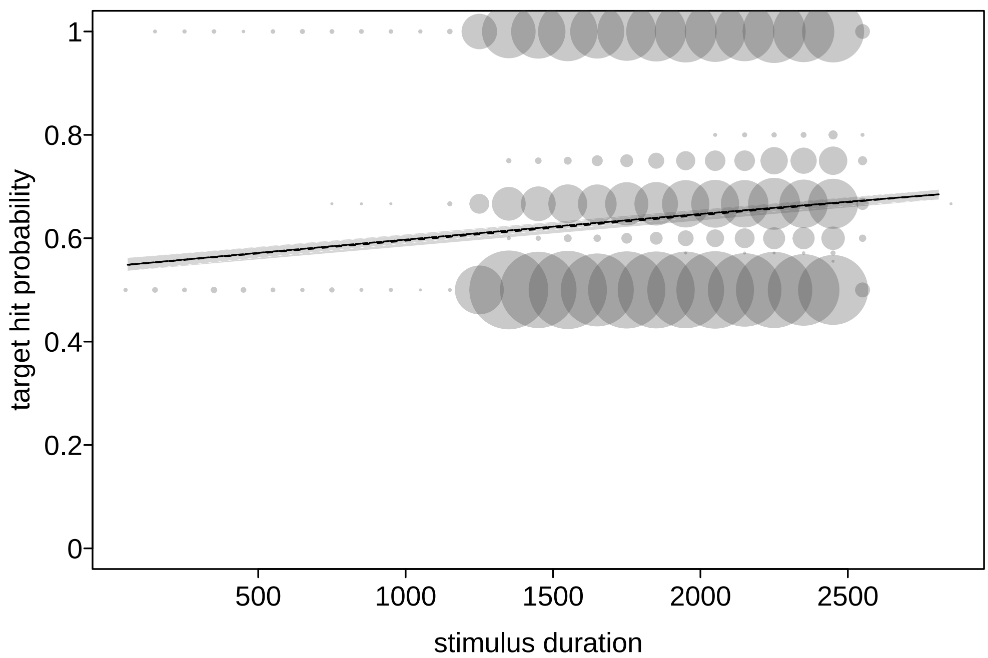

Figure SI 2. Target hit probability as a function of stimulus duration (model 1). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data (with stimulus duration being binned; bin width: 100 ms), whereby their area corresponds to the number of tied observations (range: 1 to 3824).

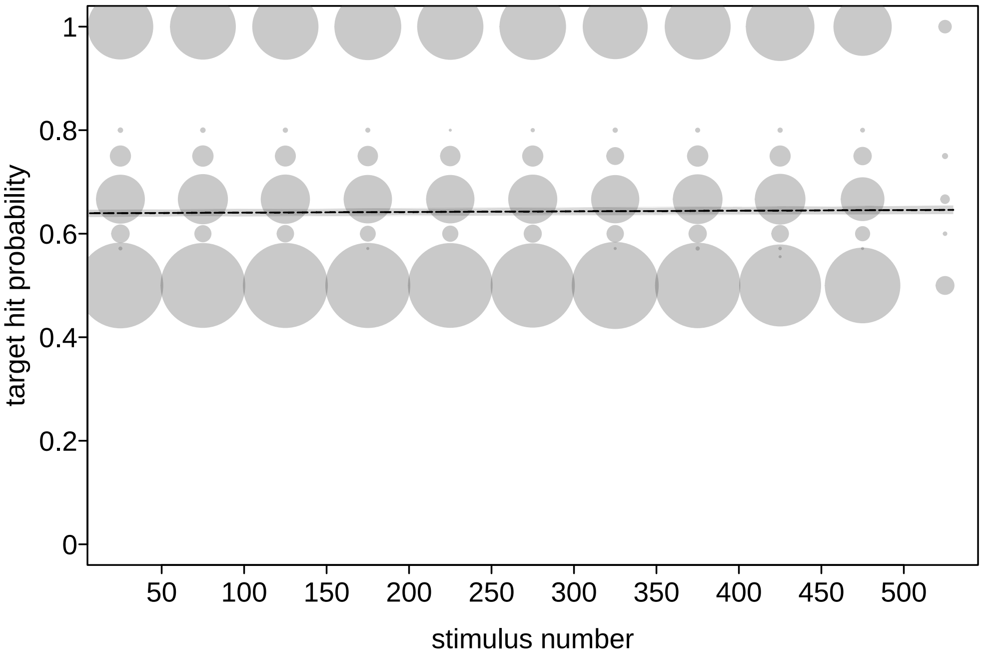

Figure SI 3. Target hit probability as a function of stimulus number (model 1). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data (with stimulus number being binned; bin width: 50 ms), whereby their area corresponds to the number of tied observations (range: 1 to 4677).

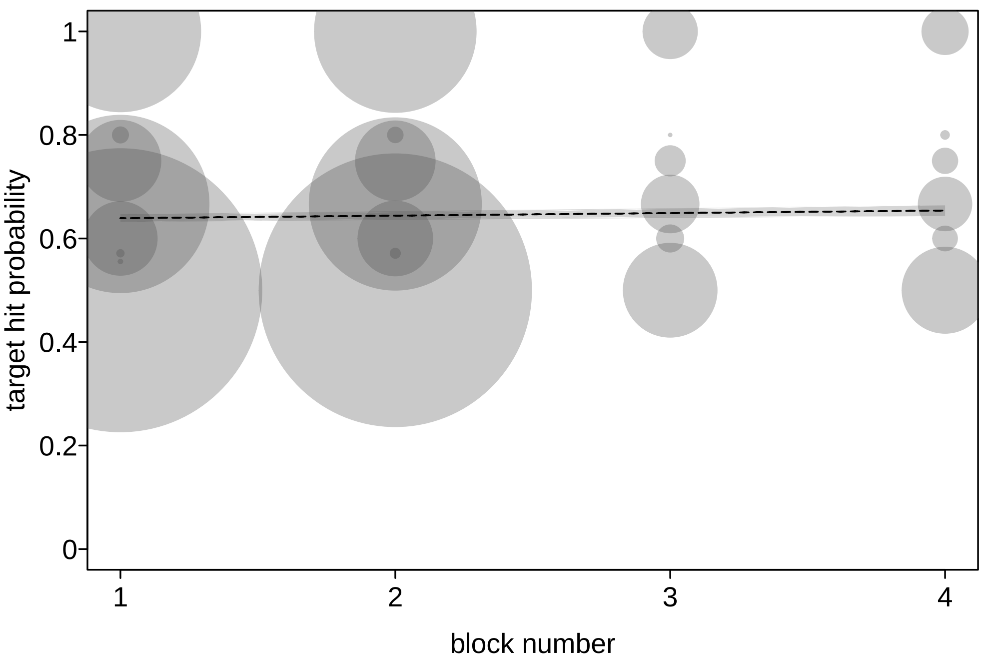

Figure SI 4. Target hit probability as a function of block number (model 1). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 5 to 51168).

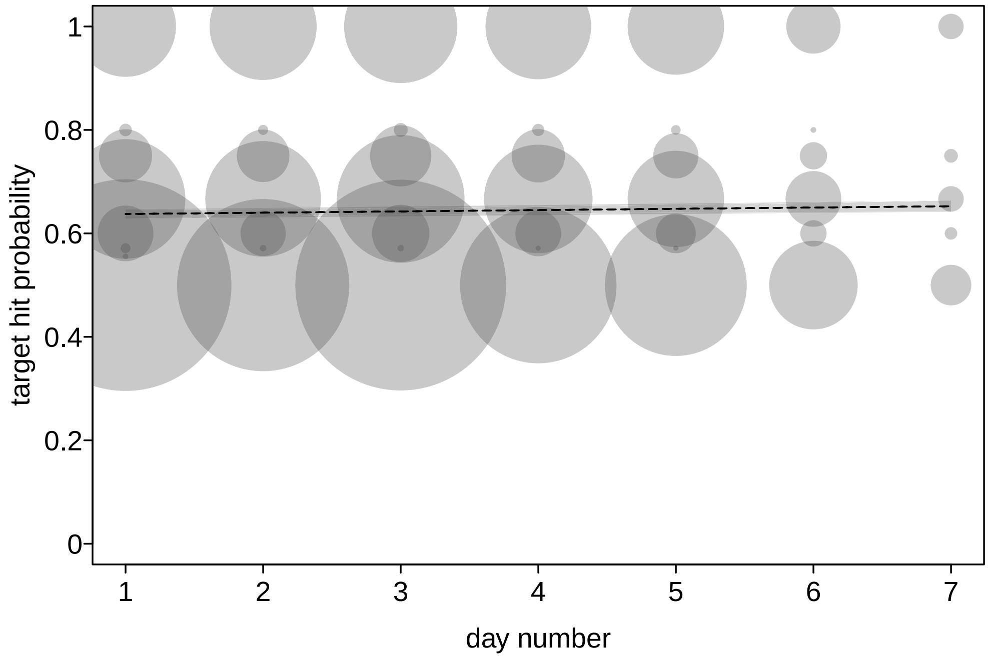

Figure SI 5. Target hit probability as a function of day number (model 1). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 7 to 28332).

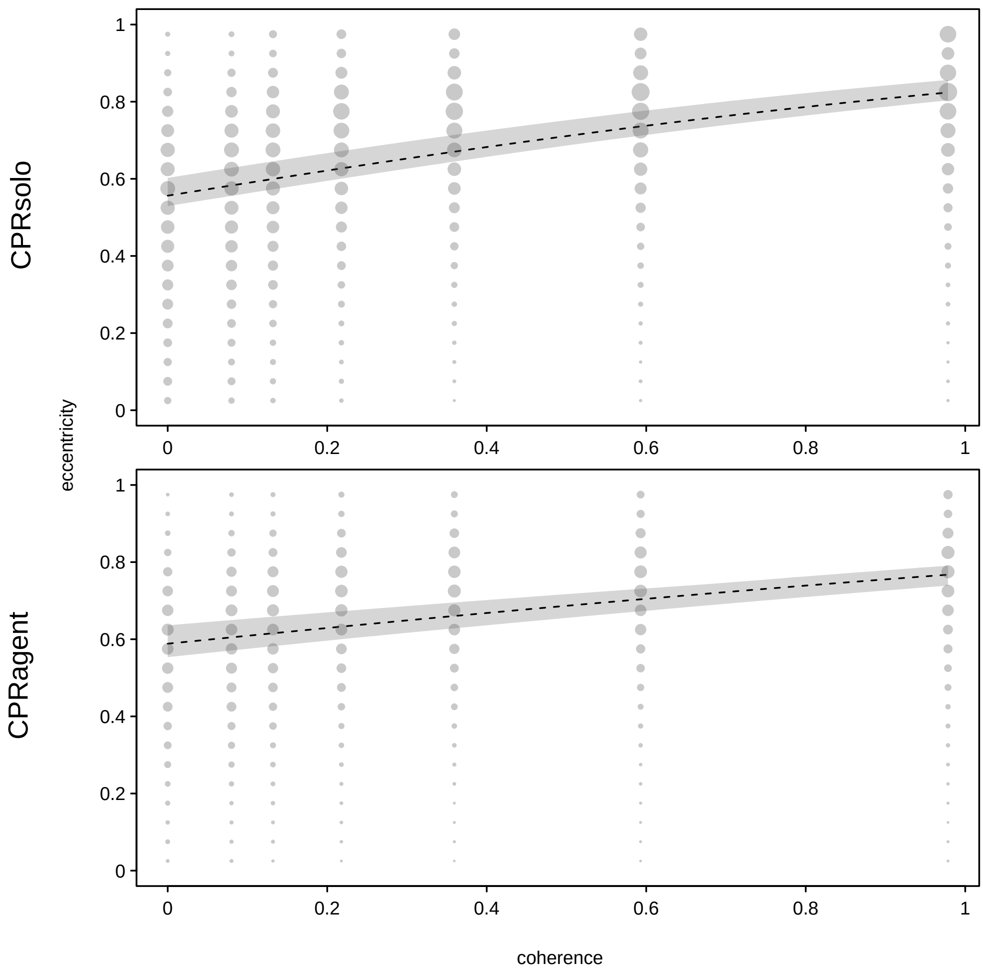

Figure SI 6. Eccentricity as a function of condition and coherence (model 1b). Dashed lines and shaded areas depict the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Eccentricity was binned with a bin width of 0.05. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 4 to 1623).

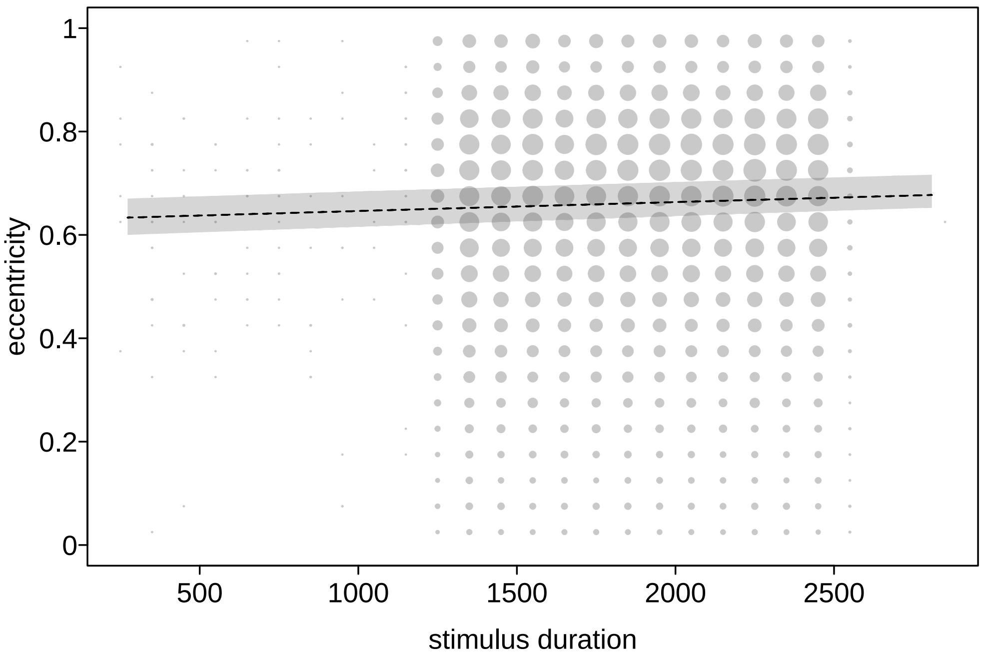

Figure SI 7. Eccentricity as a function of stimulus duration (model 1b). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Eccentricity was binned with a bin width of 0.05. Dots depict the raw data (with stimulus duration being binned; bin width: 100 ms), whereby their area corresponds to the number of tied observations (range: 1 to 1127).

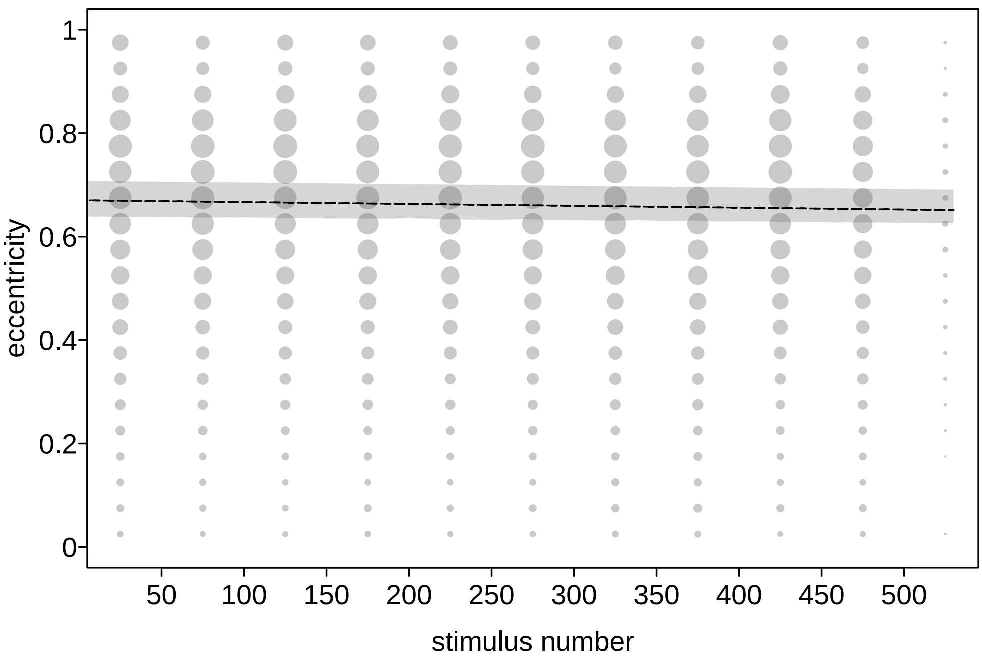

Figure SI 8. Eccentricity as a function of stimulus number (model 1b). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Eccentricity was binned with a bin width of 0.05. Dots depict the raw data (with stimulus number being binned; bin width: 50 ms), whereby their area corresponds to the number of tied observations (range: 1 to 1258).

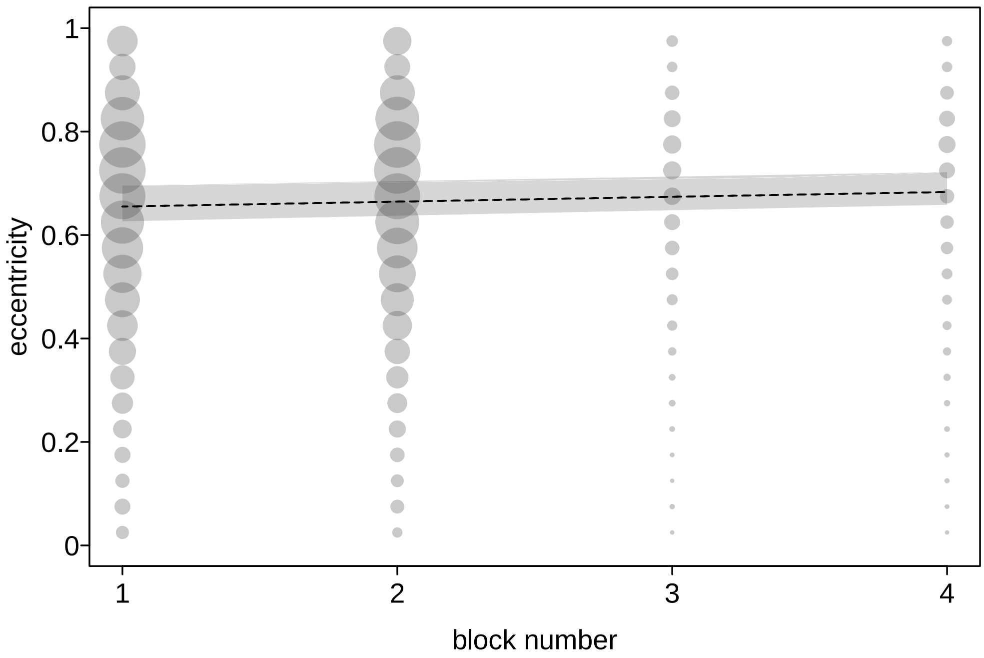

Figure SI 9. Eccentricity as a function of block number (model 1b). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Eccentricity was binned with a bin width of 0.05. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 18 to 5189). Note that block number was not significant in this model.

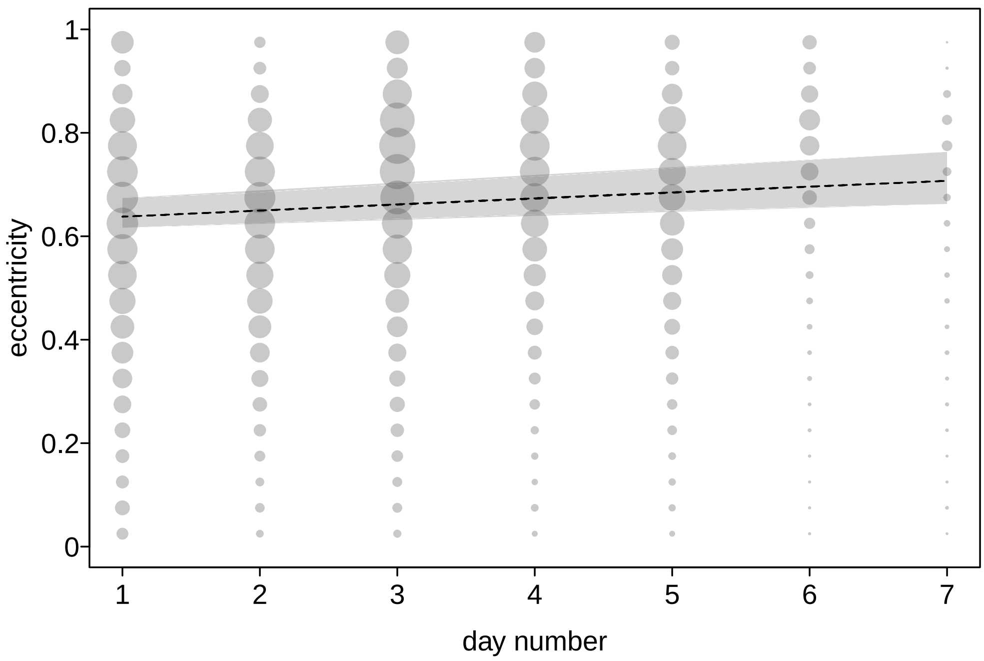

Figure SI 10. Eccentricity as a function of day number (model 1b). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Eccentricity was binned with a bin width of 0.05. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 3050).

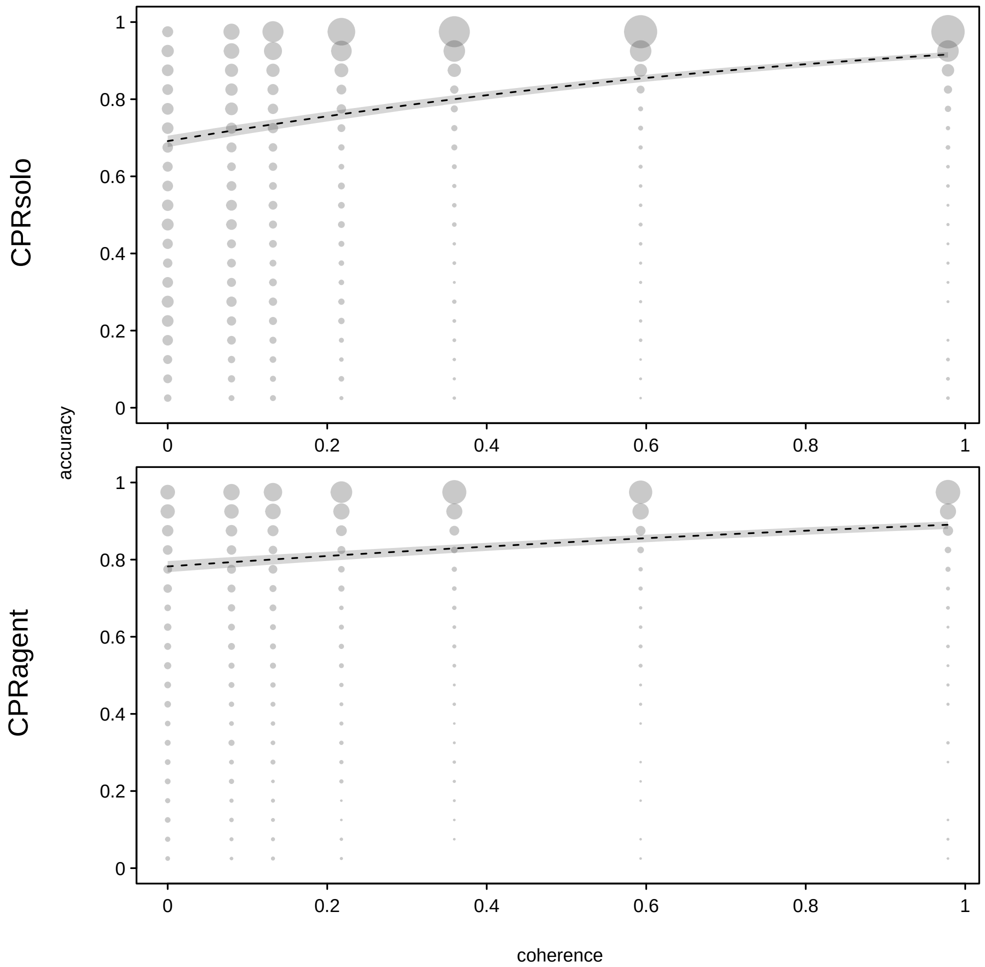

Figure SI 11. Accuracy as a function of condition and coherence (model 1b). Dashed lines and shaded areas depict the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Accuracy was binned with a bin width of 0.05. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 1478).

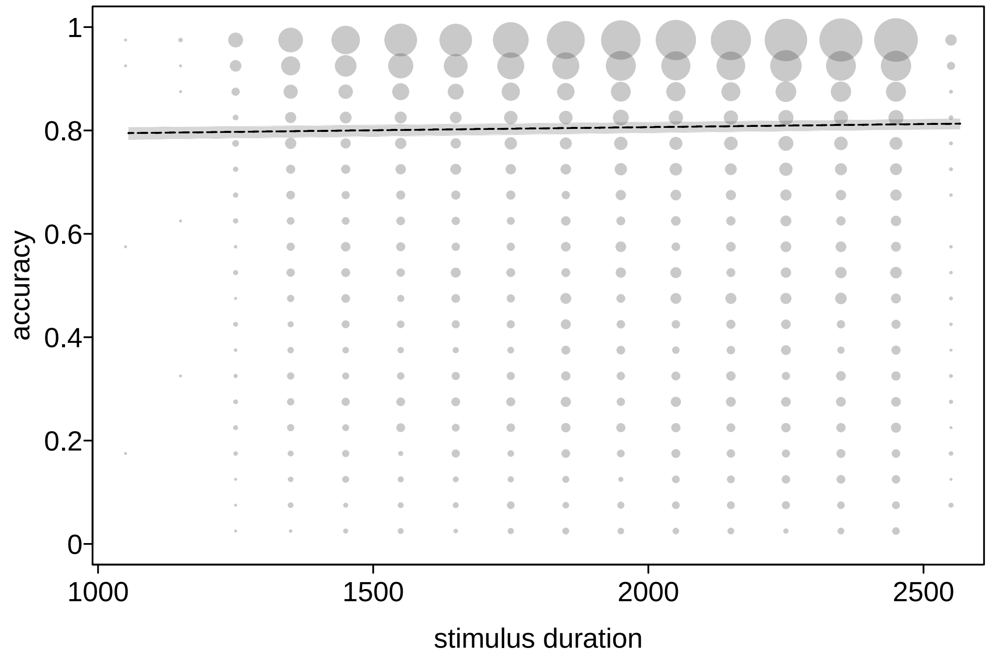

Figure SI 12. Accuracy as a function of stimulus duration (model 1c). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Accuracy was binned with a bin width of 0.05. Dots depict the raw data (with stimulus duration being binned; bin width: 100 ms), whereby their area corresponds to the number of tied observations (range: 1 to 1126).

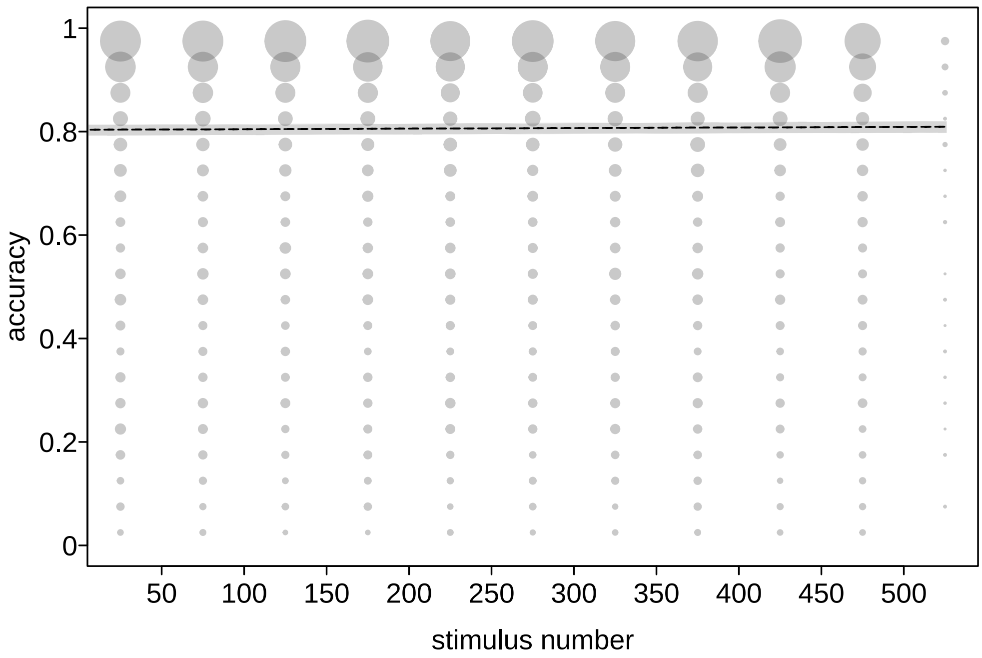

Figure SI 13. Accuracy as a function of stimulus number (model 1c). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Accuracy was binned with a bin width of 0.05. Dots depict the raw data (with stimulus number being binned; bin width: 50), whereby their area corresponds to the number of tied observations (range: 1 to 1130). Note that stimulus number was not significant in this model.

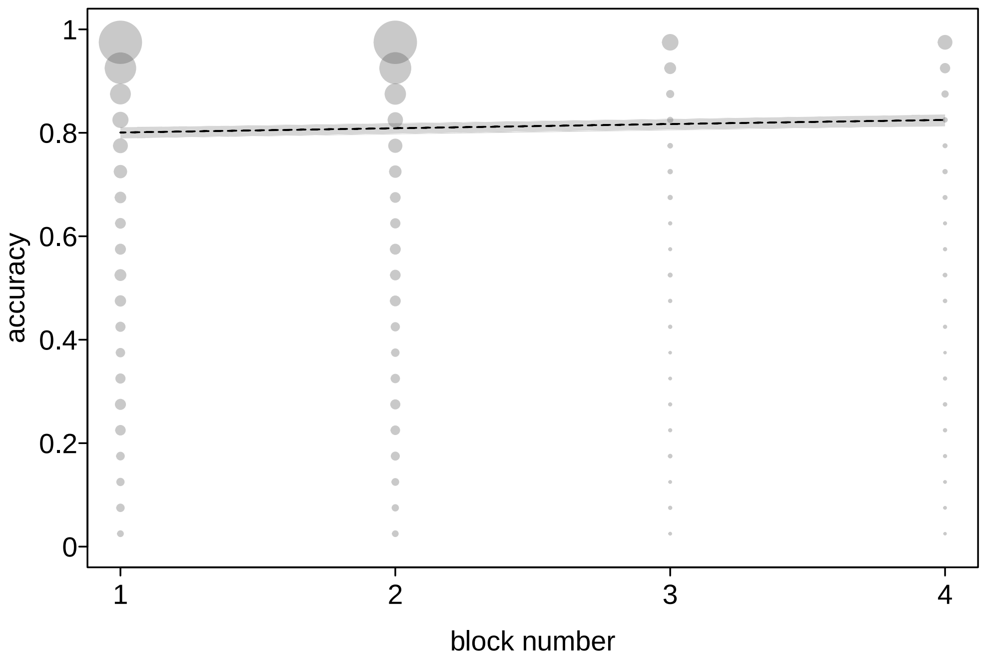

Figure SI 14. Accuracy as a function of block number (model 1c). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Accuracy was binned with a bin width of 0.05. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 8 to 4467).

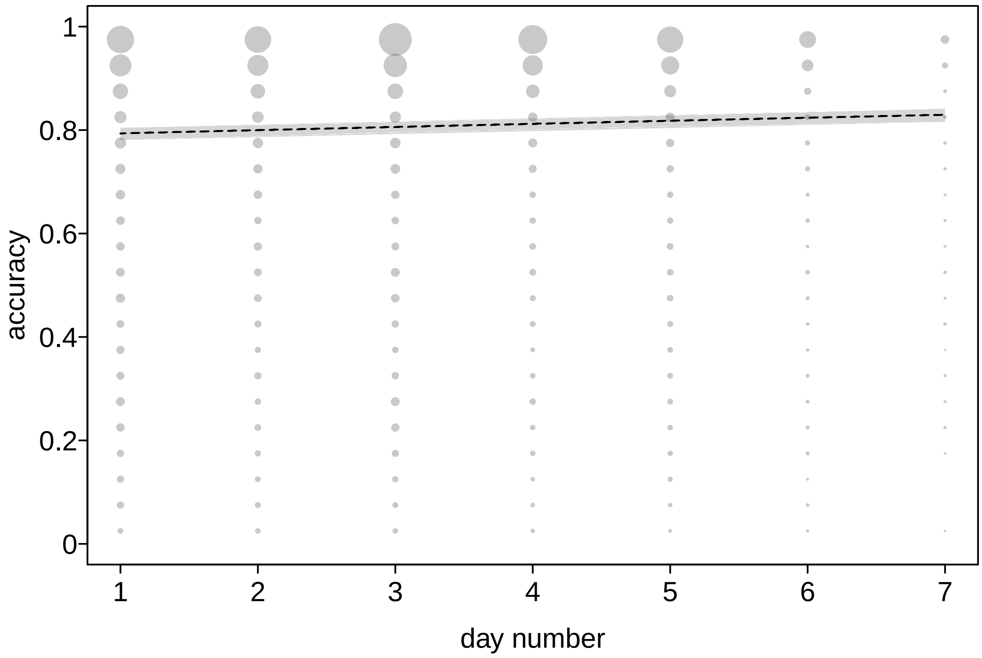

Figure SI 15. Accuracy as a function of day number (model 1c). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Accuracy was binned with a bin width of 0.05. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 2488).

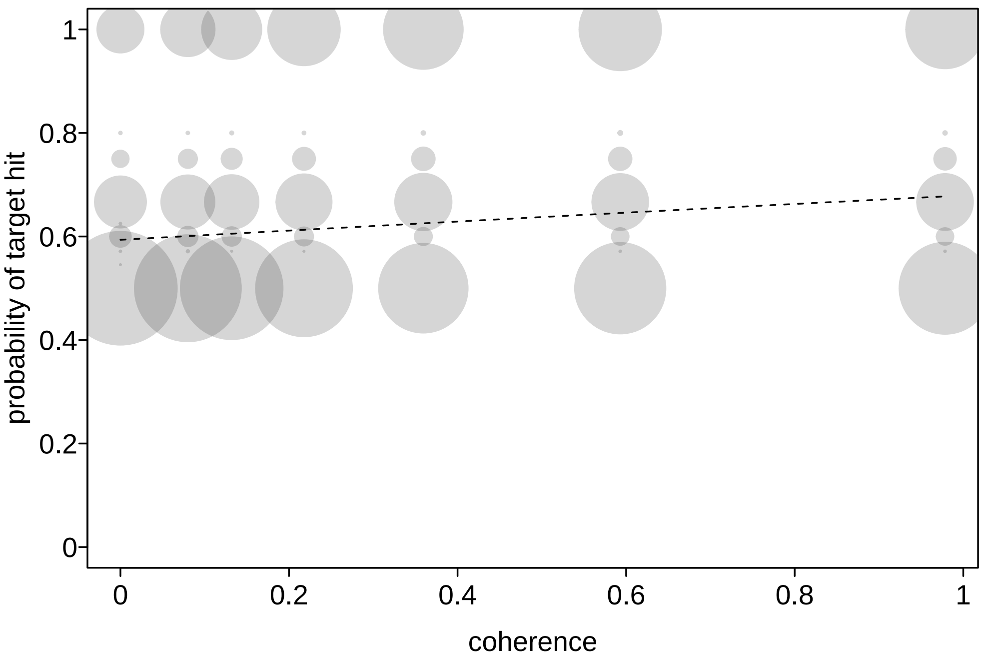

Figure SI 16. Target hit probability as a function of stimulus coherence in the dyadic condition (model 2a). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 8185). [confidence limits not yet available]

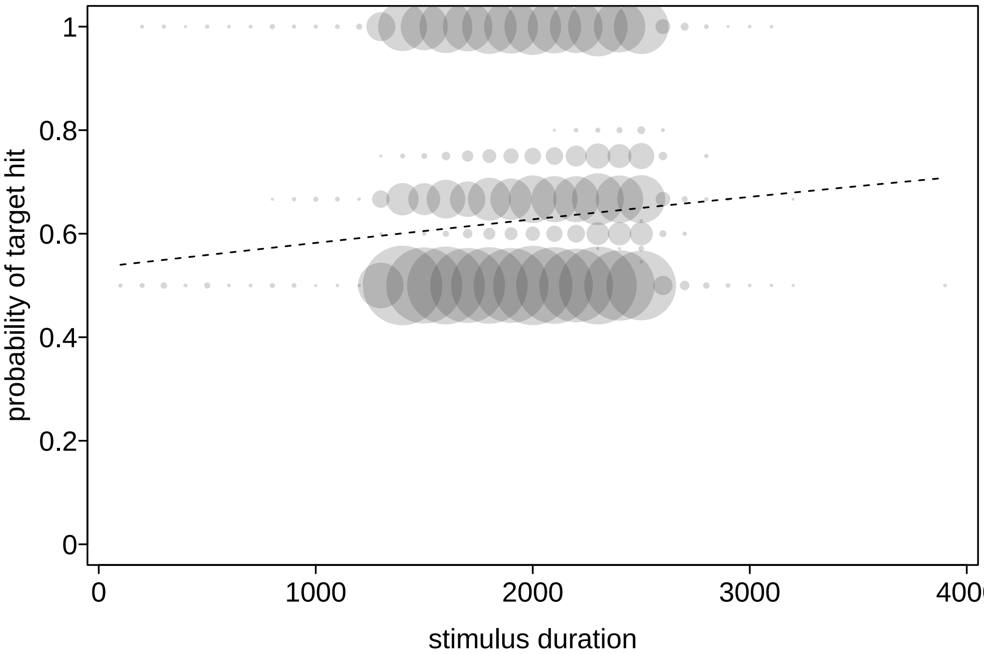

Figure SI 17. Target hit probability as a function of stimulus duration in the dyadic condition (model 2a). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 3887). [confidence limits not yet available]

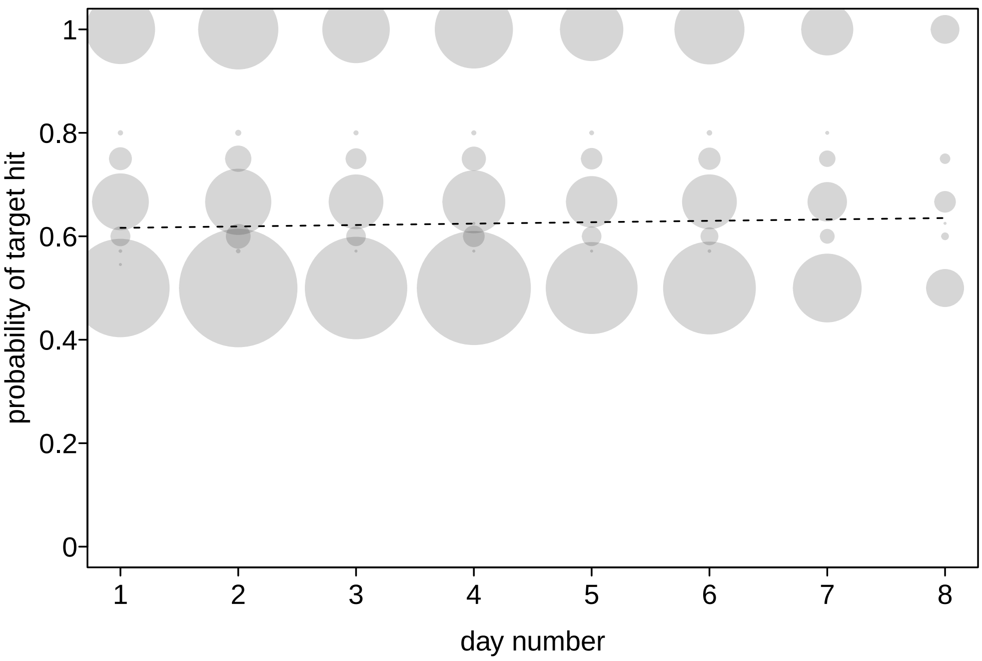

Figure SI 18. Target hit probability as a function of day number in the dyadic condition (model 2a). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 8767). [confidence limits not yet available]

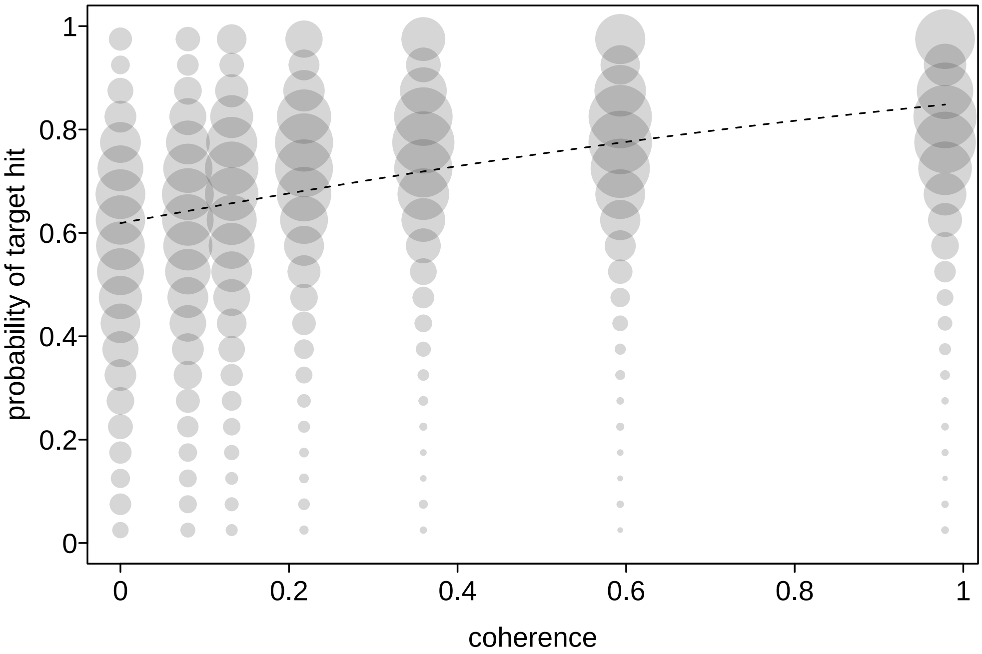

Figure SI 19. Eccentricity as a function of stimulus coherence in the dyadic condition (model 2b). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. For plotting we binned the response (bin width: 0.05). Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 9 to 2451). [confidence limits not yet available]

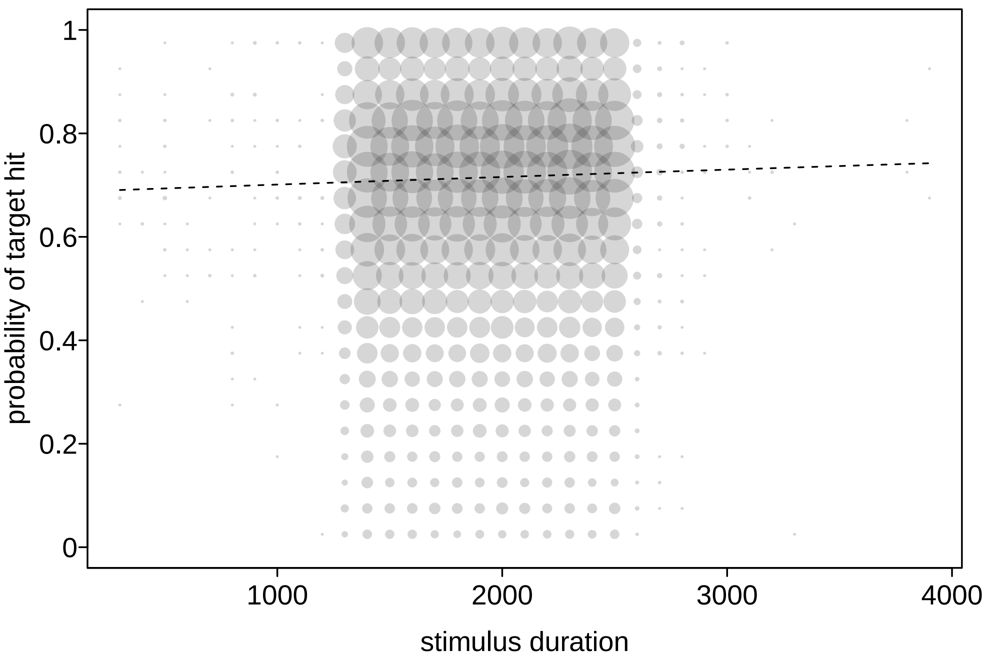

Figure SI 20. Eccentricity as a function of stimulus duration in the dyadic condition (model 2b). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. For plotting we binned the response (bin width: 0.05) and the predictor stimulus duration (bin width: 100). Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 1229). [confidence limits not yet available]

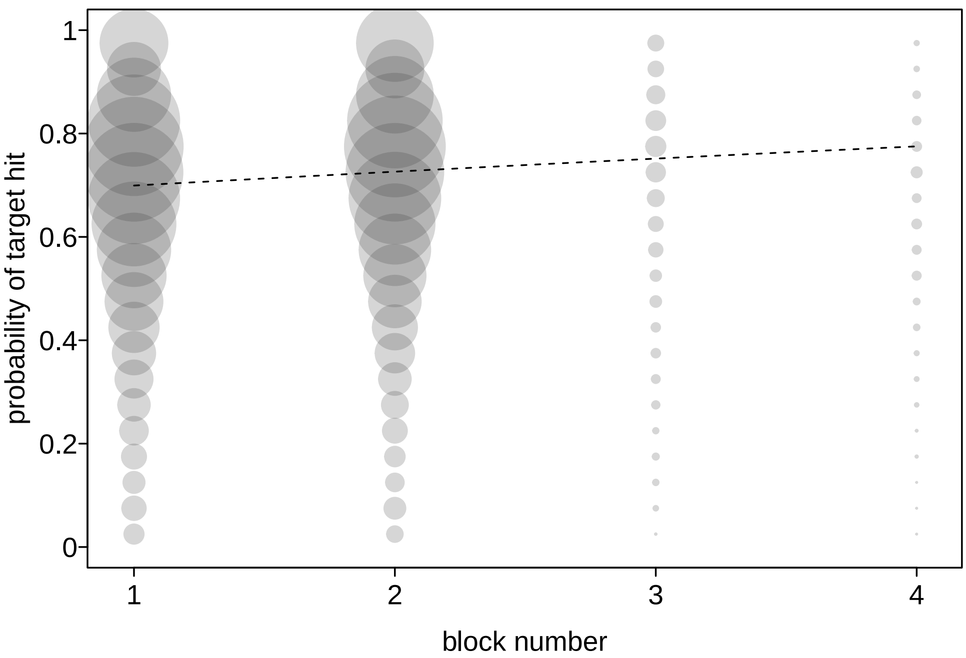

Figure SI 21. Eccentricity as a function of block number in the dyadic condition (model 2b). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. For plotting we binned the response (bin width: 0.05)). Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 6439). [confidence limits not yet available]

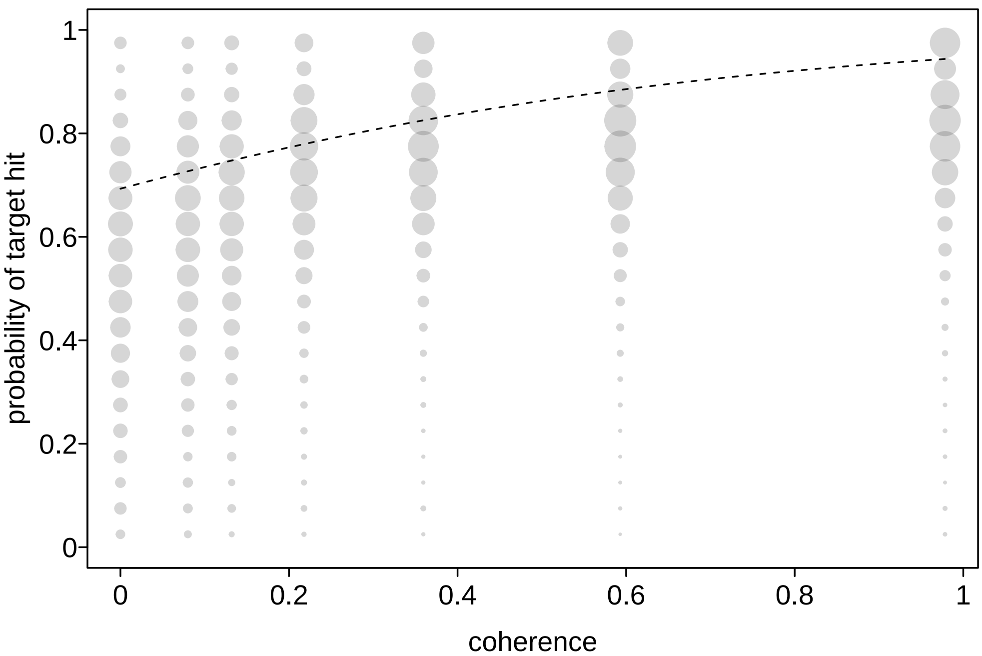

Figure SI 22. Accuracy as a function of stimulus coherence in the dyadic condition (model 2c). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. For plotting we binned the response (bin width: 0.05)). Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 2 to 594). [confidence limits not yet available]

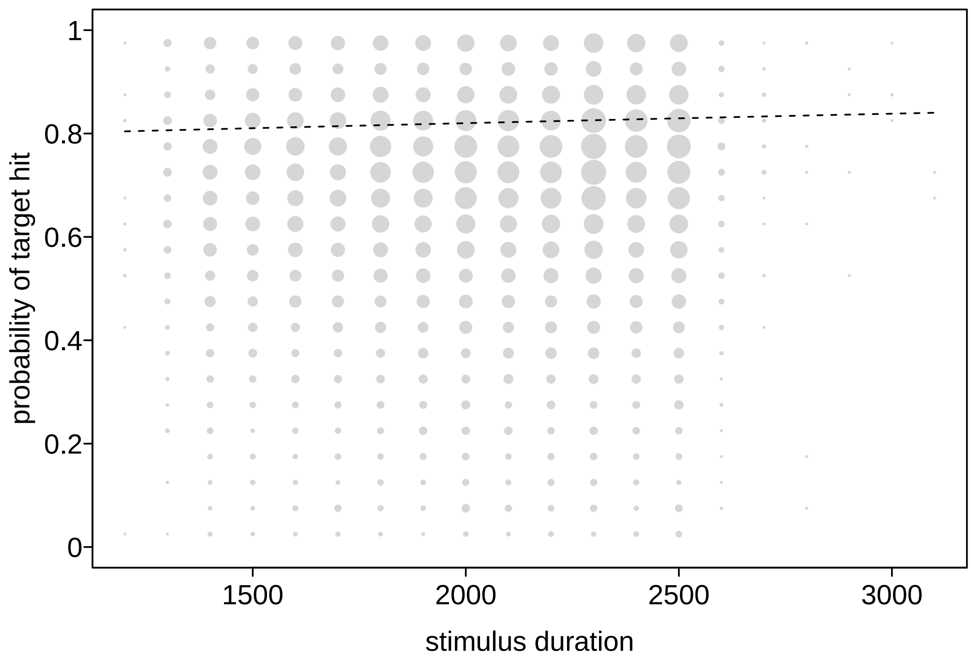

Figure SI 23. Accuracy as a function of stimulus duration in the dyadic condition (model 2c). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. For plotting we binned the response (bin width: 0.05)). Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 1 to 354). [confidence limits not yet available]

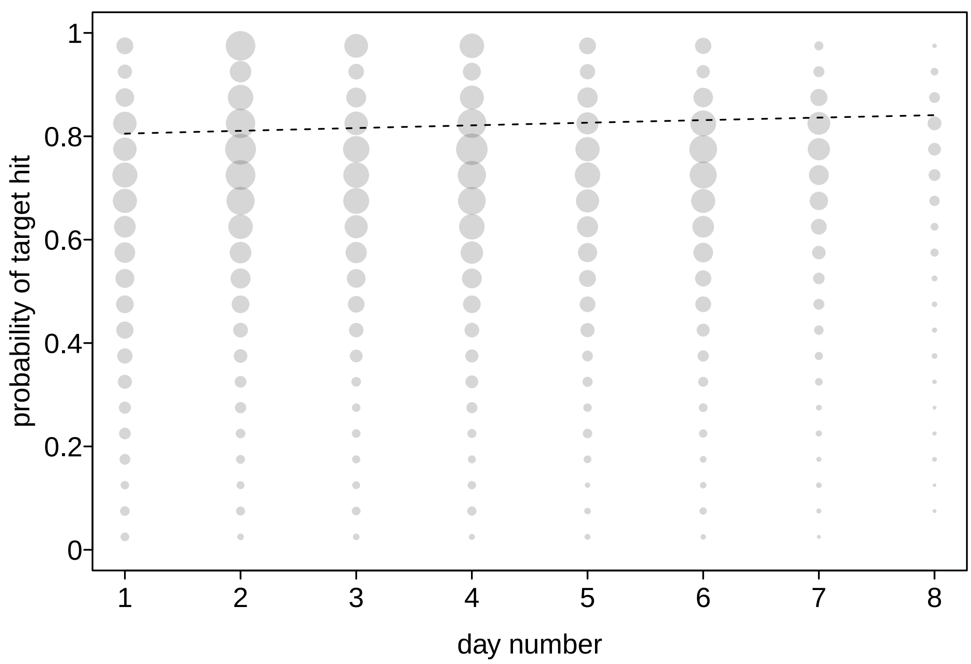

Figure SI 24. Accuracy as a function of day number in the dyadic condition (model 2c). The dashed line and shaded area depicts the fitted model and its 95% confidence limits for all other predictors in the model being at their average. For plotting we binned the response (bin width: 0.05)). Dots depict the raw data, whereby their area corresponds to the number of tied observations (range: 2 to 569). [confidence limits not yet available]

References

Baayen, R. H. (2008). *Analyzing Linguistic Data*. Cambridge University Press, Cambridge.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68:255–278.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bolker, B. M. (2008). *Ecological Models and Data in R*. Princeton University Press, Princeton, New Jersey.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., and Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9:378–400.

Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*. Chapman & Hall/CRC, Boca Raton.

Forstmeier, W. and Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behavioural Ecology and Sociobiology*, 65:47–55.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., and Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, 94:305–315.

McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*. Chapman and Hall, London.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schielzeth, H. and Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecololgy*, 20:416–420.

Smithson, M. and Verkuilen, J. (2006). A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11:54–71.