Results

Results

Simulation A

The descriptive results for all 288 conditions included in Simulation A are reported in the appendix (TODO).

Outcome: Estimation bias (MSE)

ANOVA. We used a 4 × 2 × 3 × 4 × 3 factorial Type I ANOVA (with estimation bias as an outcome and number of participants, number of timepoints per participant, missingness type, compliance, and the simulated fixed autoregressive effect) to assess which of the manipulated factors had a considerable influence on estimation bias. The results from every simulation run (i. e., 1,000 results per condition = 288,000 rows) were combined into a single dataset for the analysis. Given the very large sample size (which would make any difference significant) and the exploratory character of the analysis, *p*-values and significance thresholds were not used make inferences. Instead, we used a threshold of 0.14 for the partial ω^2, indicating a large effect size (Field, Miles, and Field 2012). This cutoff will be used for all ANOVA results throughout the results section. The partial ω^2 was chosen as the less biased alternative to partial η^2 (Okada 2013). The results and effect sizes are reported in Table 1.

Four main effects above the effect size threshold of 0.14 were found: the main effect of missingness type (ω^2 = 0.73), compliance (ω_p^2 = 0.63), the number of timepoints per participant (ω_p^2 = 0.26), and the simulated fixed slope (ω_p^2 = 0.14). Furthermore, the interaction between the missingness type and compliance (ω_p^2 = 0.54) had an effect size above the cut-off.

The main effects of missingness type and compliance are visualised in Figure 1 and Figure 2 (respectively), while the interaction between missingness type and compliance are depicted in Figure 3.

Figure 1 shows that while the underestimation of the fixed slopes is fairly low (although still considerable) when the observations are missing completely at random or in block, it becomes severe when only the most extreme values (both at one side and at both sides) are missing. SOMETHING ABOUT JANNE'S PAPER! Additionally, the underestimation of the fixed slopes becomes more severe as the compliance gets lower.

The average estimation bias when compliance is 0.8 (which is very close to the average compliance of ESM studies in psychology) is -0.13. As a consequence, many estimates of inertia in psychological research could be seriously downward biased. Furthermore, the estimates are slightly biased even when compliance is 1 (i. e., there are no missing data; average bias: -0.04). This is in line with the findings about estimation biased caused by person-mean centering in multilevel autoregressive models (Hamaker and Grasman 2015).

Zooming in on the interaction between compliance and missingness type (Figure 3) suggests that the effect of compliance on estimation bias is dramatically more severe for the two conditions in which the most extreme values of the process were set as missing (as compared to the other two conditions, i. e., data MCAR and missing in blocks). In the worst-case scenario (low compliance of 0.4; the most extreme values at both sides missing), the average estimation bias was -0.48). Given that the average simulated fixed slope was 0.5, these results imply that even rather large autocorrelations can be estimated as close to 0 in studies with the combination of low compliance and a non-random missingness pattern. At the same time, the results about data MCAR and missing in blocks are encouraging. Even in a low-compliance (0.4) condition, the average estimation bias was -0.08 for the former and -0.09 for the latter.

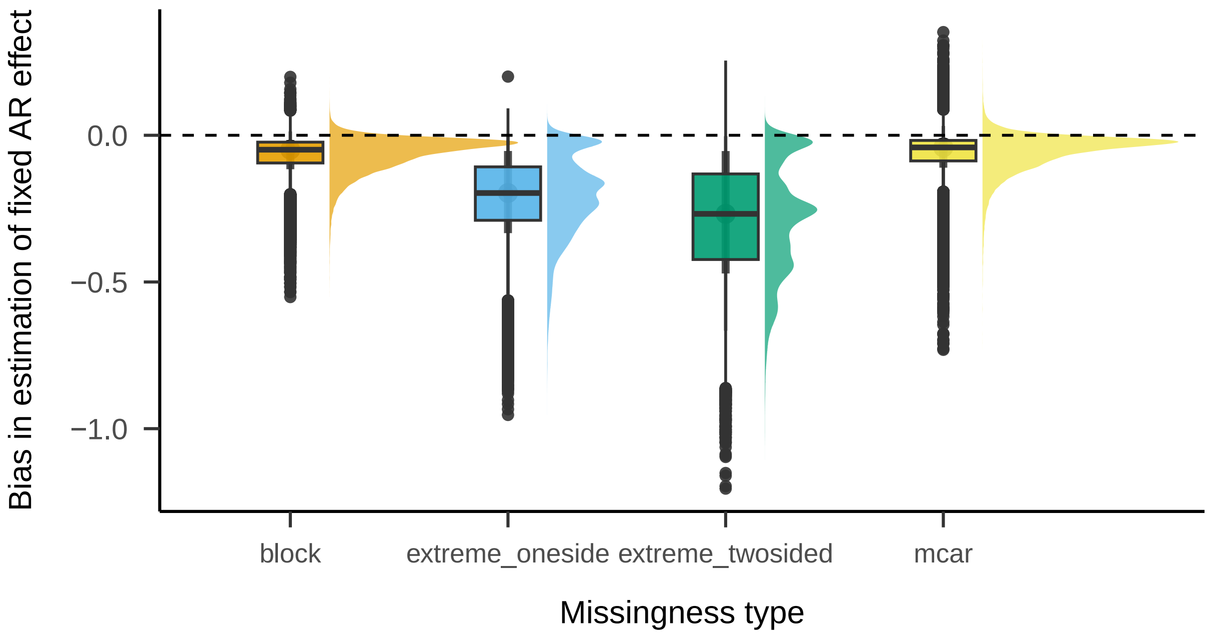

Figure 1: The effect of compliance on the bias in estimation of the fixed slopes.

The average estimation bias for all combinations of missingness type and compliance (averaged over the different values of the number of participants, timepoints per participant and simulated fixed slope) is reported in Table 2.

Outcome: Standard error

Descriptive statistics. The average standard errors for the different combinations of number of participants, timepoints per participant and compliance are reported in Table \ref{tab:tab_aov_se}.

ANOVA. To examine the effect of the manipulated parameters on the standard error of the estimation of the fixed slopes, we combined the results for each condition (1,000 simulation runs) into a single row. As such, the dataset used for the following analyses had one row per simulation condition (288 rows in total). A a $4 \times 2 \times 3 \times 4 \times 3$ factorial Type I ANOVA was used to analyse the data. The full ANOVA results and effect sizes are reported in Table DS.

The main effects of the number of participants ($\omega_p^2 = 0.68$), number of timepoints per participant ($\omega_p^2 = 0.68$) and compliance ($\omega_p^2 = 0.66$) crossed the cut-off for effect size.

Additionally, the interaction between the number of timepoints per participant and compliance ($\omega_p^2 = 0.28$), number of participants and timepoints per participants ($\omega_p^2 = 0.07$), and between the number of participants and compliance ($\omega_p^2 = 0.22$) was found.

Figure 4 depicts the interaction between the number of timepoints per participant and compliance, while Figure 5 shows the interaction between the number of participants and compliance.

Bias in the estimation of the fixed slope and bias in the estimation of the person-specific means. TODO.

Outcome: Statistical power

Descriptive statistics. The statistical power for each combination of the manupulated parameters is reported in Table (BIG TABLE TODO). As an illustration, the effect of compliance, missingness type, the

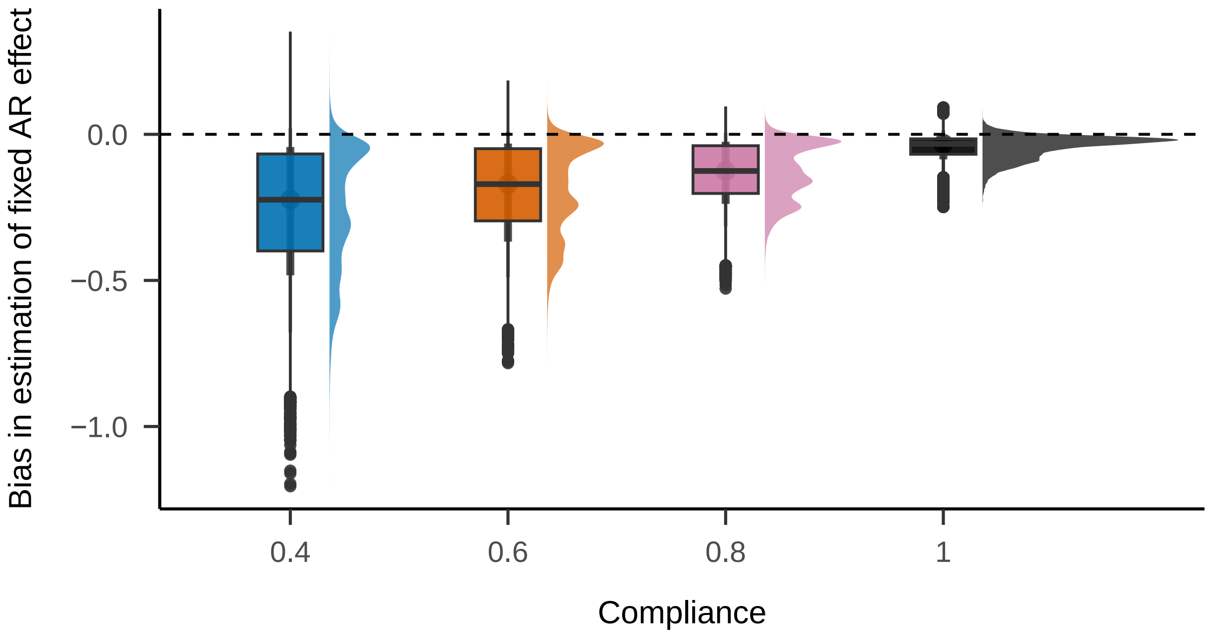

Figure 2: The effect of missingness type on the bias in estimation of the fixed slopes.

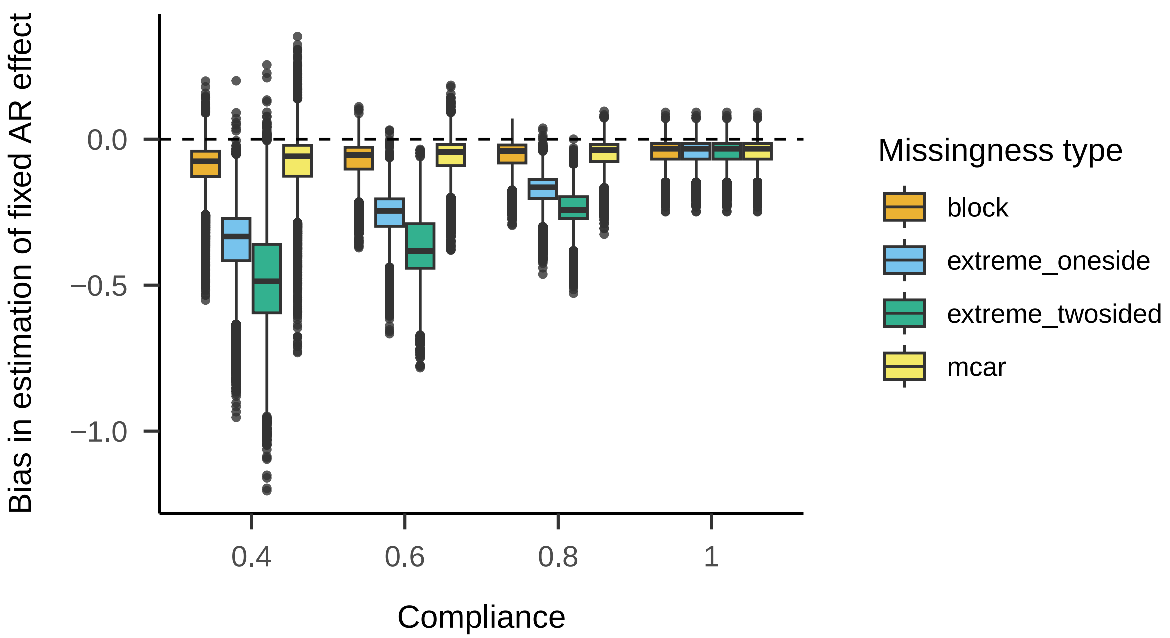

Figure 3: The effect of the interaction between missingness type and compliance on the bias in estimation of the fixed slopes.

Table 1: ANOVA results, simulation A. Outcome: Estimation bias

	Df	Sum Sq	Mean Sq	F value	p-value	Partial omega-squared
N	1	0.11	0.11	32.53	<0.001	0.00
T.obs	1	354.93	354.93	101753.95	<0.001	0.26
miss_type	3	2657.10	885.70	253921.29	<0.001	0.73
compliance	1	1706.99	1706.99	489377.69	<0.001	0.63
B1_sim	1	169.00	169.00	48449.29	<0.001	0.14
N:T.obs	1	0.02	0.02	4.40	0.0360	0.00
N:miss_type	3	0.00	0.00	0.31	0.8216	0.00
T.obs:miss_type	3	14.11	4.70	1348.32	<0.001	0.01
N:compliance	1	0.03	0.03	8.09	0.0044	0.00
T.obs:compliance	1	22.78	22.78	6529.83	<0.001	0.02
miss_type:compliance	3	1157.71	385.90	110634.77	<0.001	0.54
N:B1_sim	1	0.03	0.03	8.41	0.0037	0.00
T.obs:B1_sim	1	1.75	1.75	502.84	<0.001	0.00
miss_type:B1_sim	3	148.38	49.46	14179.70	<0.001	0.13
compliance:B1_sim	1	59.28	59.28	16994.38	<0.001	0.06
Residuals	287974	1004.48	0.00		NA	

Table 2: Simulation A. Average bias in estimation of the fixed slope for each combination of missingness type and compliance.

	Missingness type			
compliance	block	extreme_oneside	extreme_twosided	mcar
0.4	-0.09	-0.36	-0.48	-0.08
0.6	-0.07	-0.26	-0.37	-0.06
0.8	-0.05	-0.17	-0.24	-0.05
1.0	-0.04	-0.04	-0.04	-0.04

Table 3: Simulation A. Average standard error in the estimation of the fixed slope for each combination of number of participants, number of timepoints/participant, and compliance.

		Compliance			
N	T.obs	0.4	0.6	0.8	1
20	20	0.14	0.08	0.06	0.05
	50	0.07	0.05	0.04	0.03
	100	0.05	0.03	0.02	0.02
100	20	0.06	0.04	0.03	0.02
	50	0.03	0.02	0.02	0.01
	100	0.02	0.01	0.01	0.01

Table 4: ANOVA results, simulation A. Outcome: Standard error

	Df	Sum Sq	Mean Sq	F value	p-value	Partial omega-squared
N	1	0.06	0.06	625.92	<0.001	0.68
T.obs	1	0.06	0.06	621.16	<0.001	0.68
miss_type	3	0.00	0.00	14.11	<0.001	0.12
compliance	1	0.05	0.05	556.59	<0.001	0.66
B1_sim	1	0.00	0.00	21.75	<0.001	0.07
N:T.obs	1	0.01	0.01	91.92	<0.001	0.24
N:miss_type	3	0.00	0.00	2.13	0.096	0.01
T.obs:miss_type	3	0.00	0.00	1.48	0.220	0.00
N:compliance	1	0.01	0.01	82.89	<0.001	0.22
T.obs:compliance	1	0.01	0.01	114.06	<0.001	0.28
miss_type:compliance	3	0.00	0.00	13.31	<0.001	0.11
N:B1_sim	1	0.00	0.00	3.14	0.078	0.01
T.obs:B1_sim	1	0.00	0.00	1.37	0.243	0.00
miss_type:B1_sim	3	0.00	0.00	0.20	0.895	0.00
compliance:B1_sim	1	0.00	0.00	1.55	0.214	0.00
Residuals	262	0.03	0.00		NA	

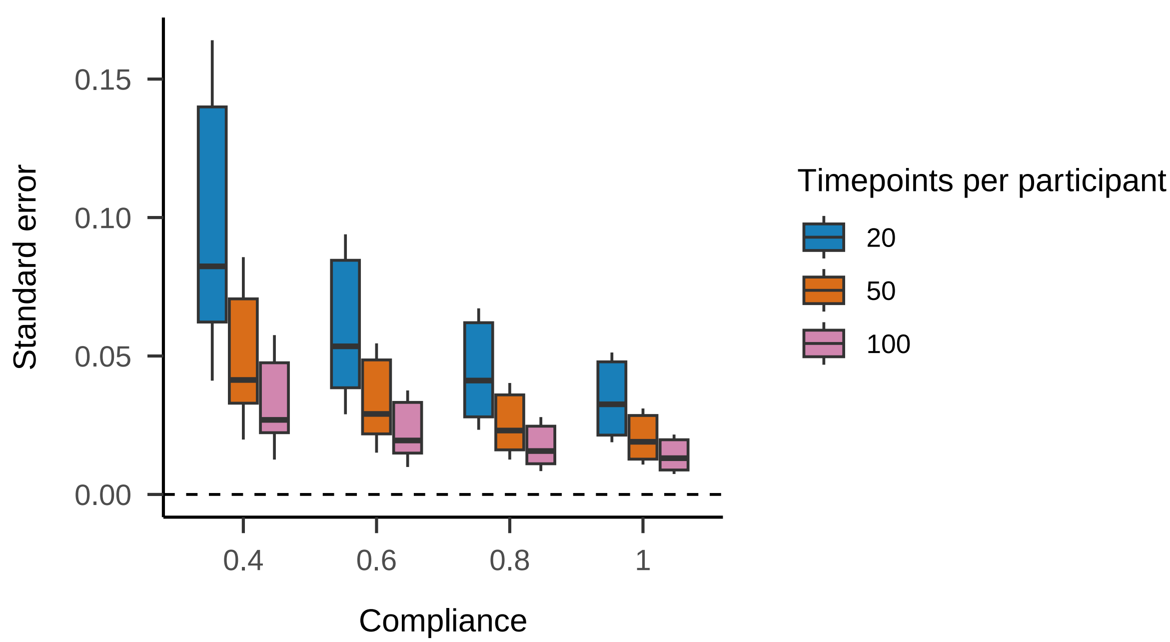

Figure 4: The effect of the interaction between number of timepoints and compliance on standard error of estimation of the fixed slopes. Simulation A.

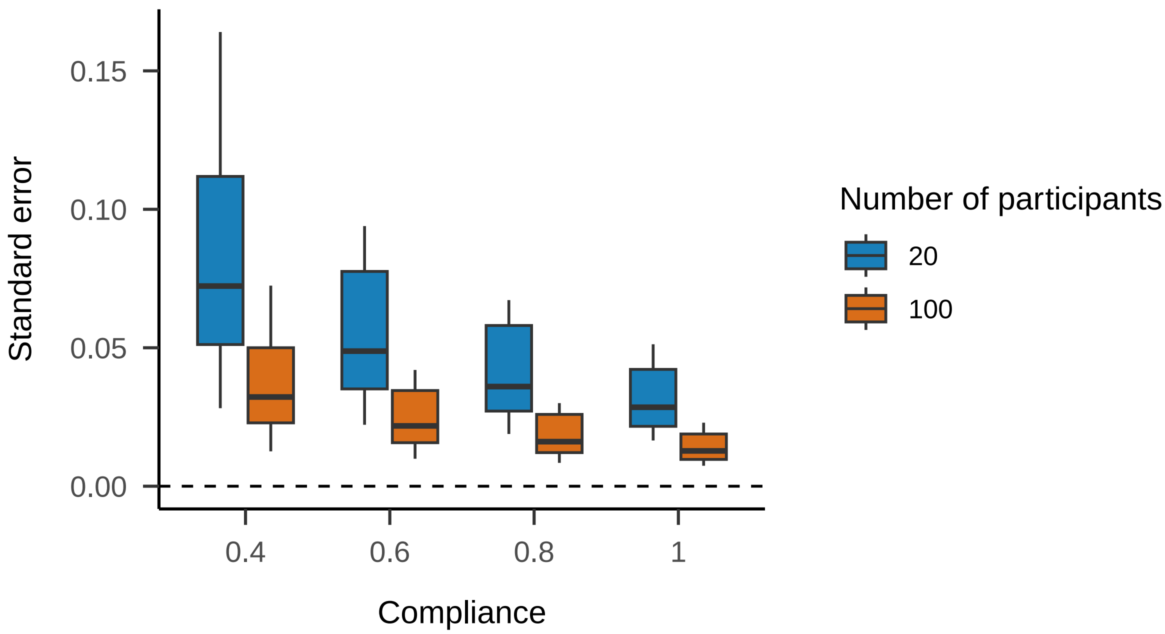

Figure 5: The effect of the interaction between number of participants and compliance on standard error of estimation of the fixed slopes. Simulation A.

number of participants and the number of timepoints per participant when the simulated fixed slope is 0.3 are visualised in Figure 6. Consistent with the results about estimation bias, statistical power is the lowest in the two conditions with the most extreme datapoints missing. For the conditions with data missing completely at random and data missing in consecutive blocks, power is very high even when the compliance is low for most conditions (except for the two conditions with T = 20).

A peculiar pattern is worth pointing out in the plot: in the two conditions with T = 20 and the most extreme data missing at both sides (green dashed line), the statistical power is higher when compliance is 0.4 compared to when compliance is 0.6. This counterintuitive result is likely due to the fact that the underestimation is the most severe when the most extreme values at both sides. As such, some of the estimates of the fixed slope will be negative, and their magnitude will be large enough for them to reach statistical significance.

ANOVA. Aa $4 \times 2 \times 3 \times 4 \times 3$ factorial Type I ANOVA was used to analyse the effect of the manipulated parameters (288 conditions in total) on statistical power. The results are reported in Table 5.

Four main effect above the cut-off for the effect size were found: the effect of compliance ($\omega_p^2 = 0.43$), of missingness type ($\omega_p^2 = 0.36$), simulated fixed slope ($\omega_p^2 = 0.18$), and the effect of the number of timepoints per participant ($\omega_p^2 = 0.17$).

Simulation B

Field, Andy, Jeremy Miles, and Zoë Field. 2012. "Discovering Statistics Using r (2012)." *Great Britain: Sage Publications, Ltd* 958.

Hamaker, Ellen L., and Raoul P. P. P. Grasman. 2015. "To Center or Not to Center? Investigating Inertia with a Multilevel Autoregressive Model." *Frontiers in Psychology* 5 (January). https://doi.org/10.3389/fpsyg.2014.01492.

Okada, Kensuke. 2013. "Is Omega Squared Less Biased? A Comparison of Three Major Effect Size Indices in One-Way ANOVA." *Behaviormetrika* 40 (2): 129147.

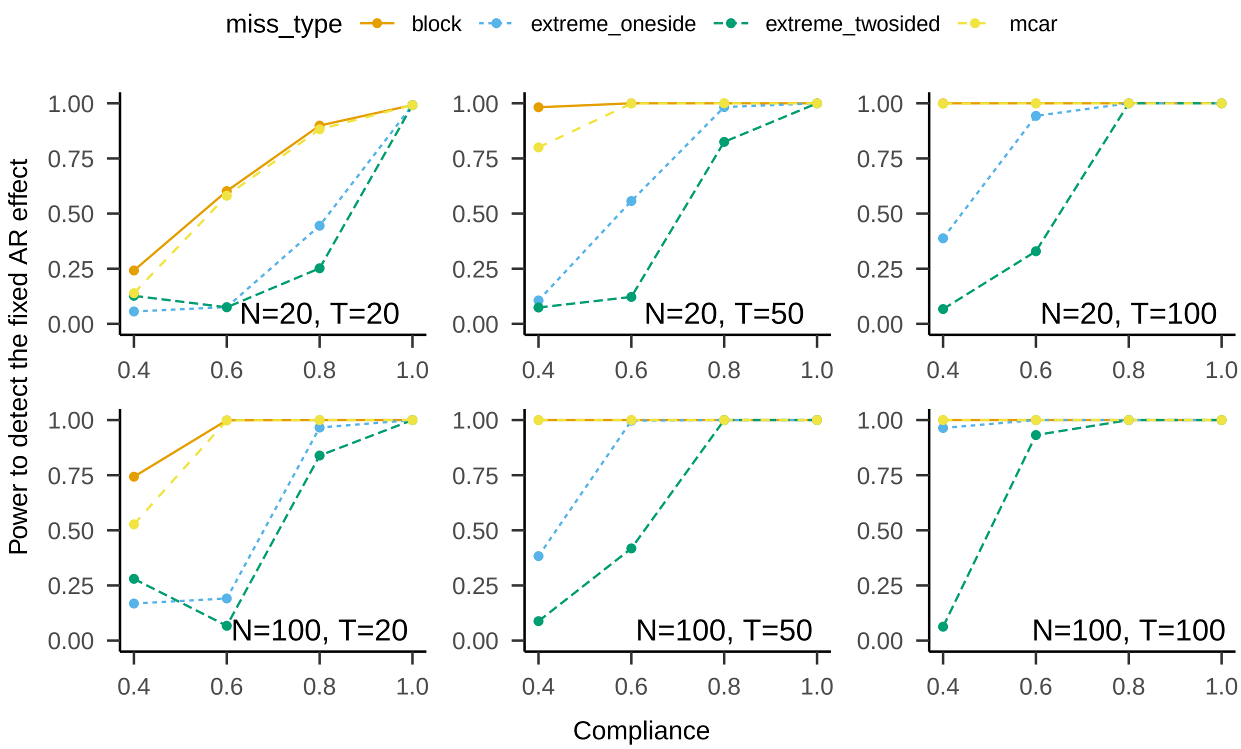

Figure 6: Simulation A. Statistical power to detect the fixed slope for all combinations of compliance, missingness type, number of participants and timepoints per participant when the simulated fixed slope is 0.3.

Table 5: ANOVA results, simulation A. Outcome: Power to detect the fixed slope

	Df	Sum Sq	Mean Sq	F value	p-value	Partial omega-squared
N	1	0.42	0.42	20.32	<0.001	0.06
T.obs	1	1.21	1.21	58.92	<0.001	0.17
miss_type	3	3.36	1.12	54.56	<0.001	0.36
compliance	1	4.47	4.47	217.19	<0.001	0.43
B1_sim	1	1.35	1.35	65.78	<0.001	0.18
N:T.obs	1	0.08	0.08	3.80	0.0524	0.01
N:miss_type	3	0.10	0.03	1.66	0.1753	0.01
T.obs:miss_type	3	0.34	0.11	5.49	0.0011	0.04
N:compliance	1	0.22	0.22	10.93	0.0011	0.03
T.obs:compliance	1	0.79	0.79	38.66	<0.001	0.12
miss_type:compliance	3	3.12	1.04	50.65	<0.001	0.34
N:B1_sim	1	0.09	0.09	4.25	0.0403	0.01
T.obs:B1_sim	1	0.24	0.24	11.50	<0.001	0.04
miss_type:B1_sim	3	0.44	0.15	7.19	<0.001	0.06
compliance:B1_sim	1	0.76	0.76	37.10	<0.001	0.11
Residuals	262	5.39	0.02		NA	