Routinely Randomize Potential Sources of Measurement Reactivity to Estimate and Adjust for Biases in Subjective Reports

Contents

1 Additional descriptive statistics

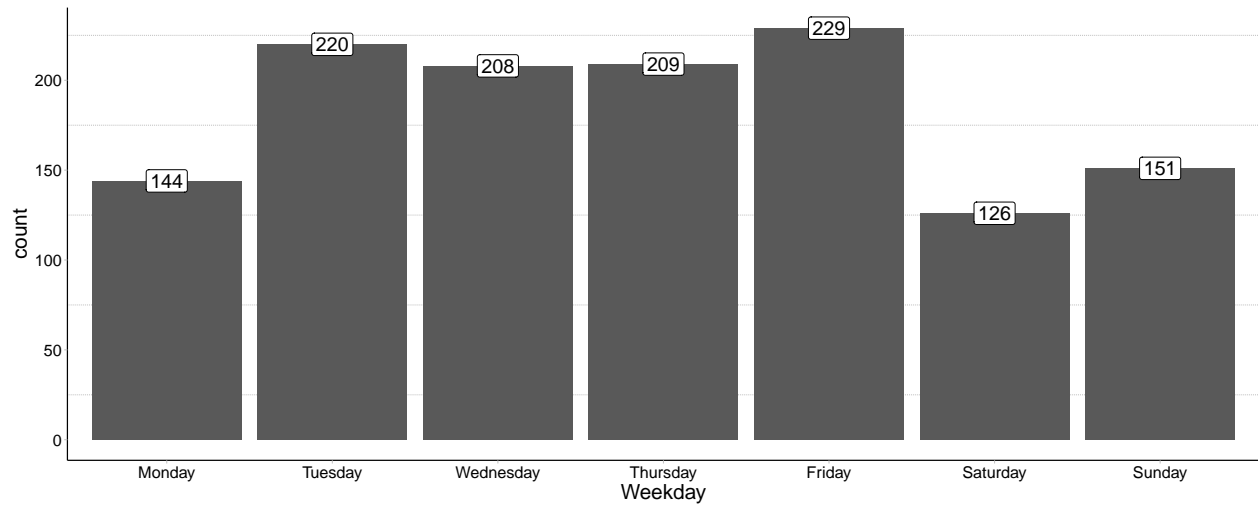Figure 1: Weekday distribution on diary starting day. We see that women were more likely to start the diary on Tuesday-Friday.
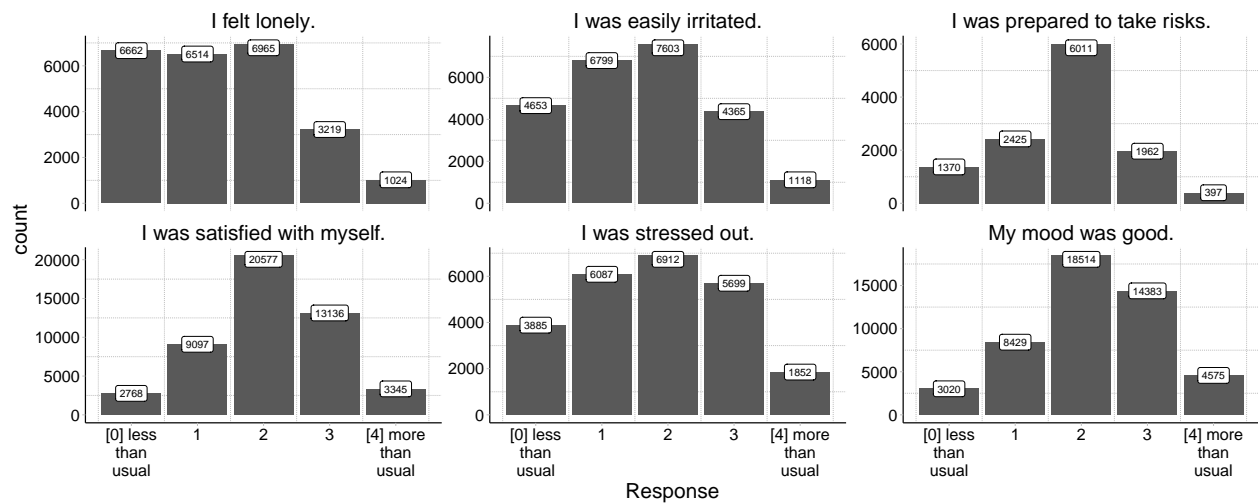
Figure 2: Distributions of items.
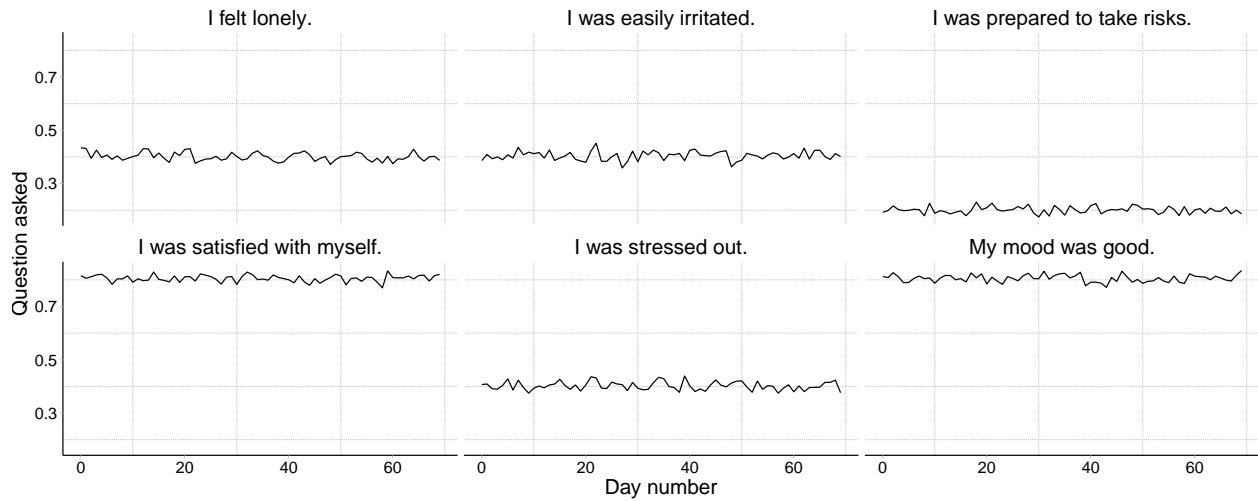
Figure 3: A quick verification that fluctuation in whether a question is asked is as random as intended over time.
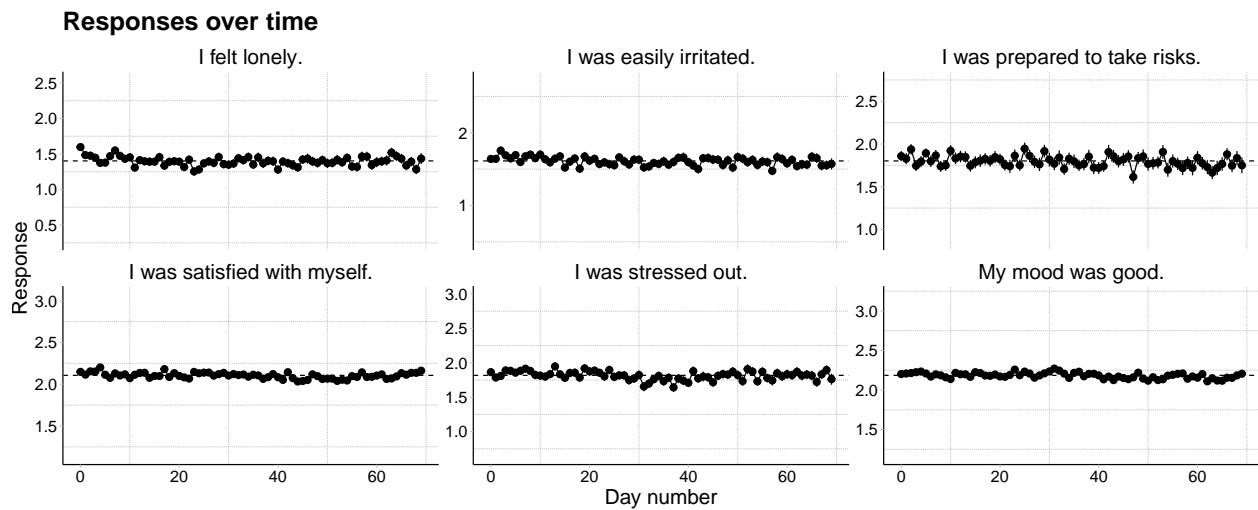
Figure 4: No strong time trends in the responses. The dashed line shows the overall mean. Y axis range is overall mean ±1SD.

2 Potential sources of measurement reactivity

2.1 Initial elevation bias
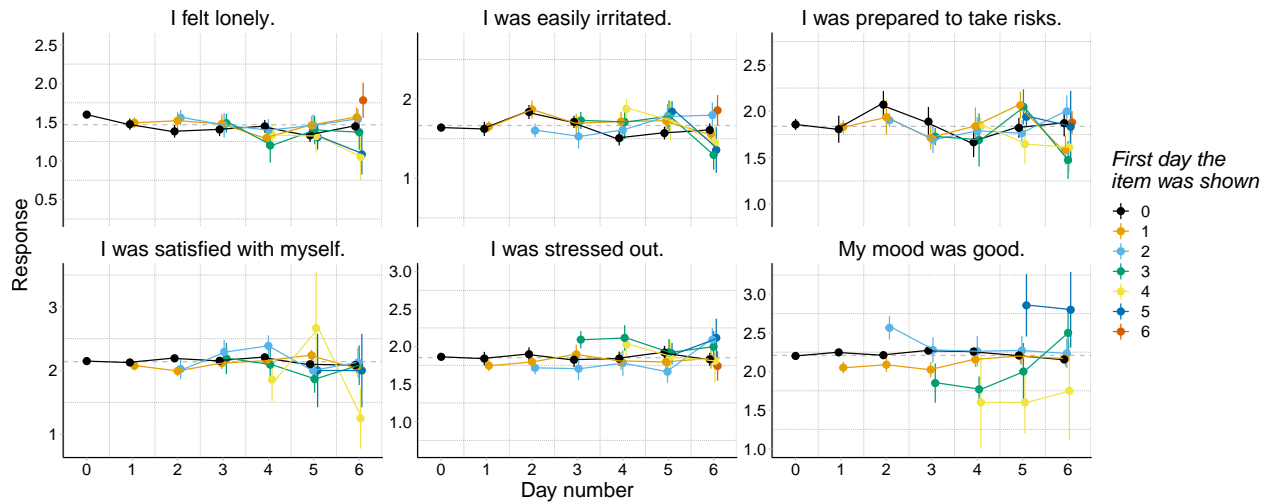
Figure 5: Figure 2, including combinations with fewer than 20 observations.
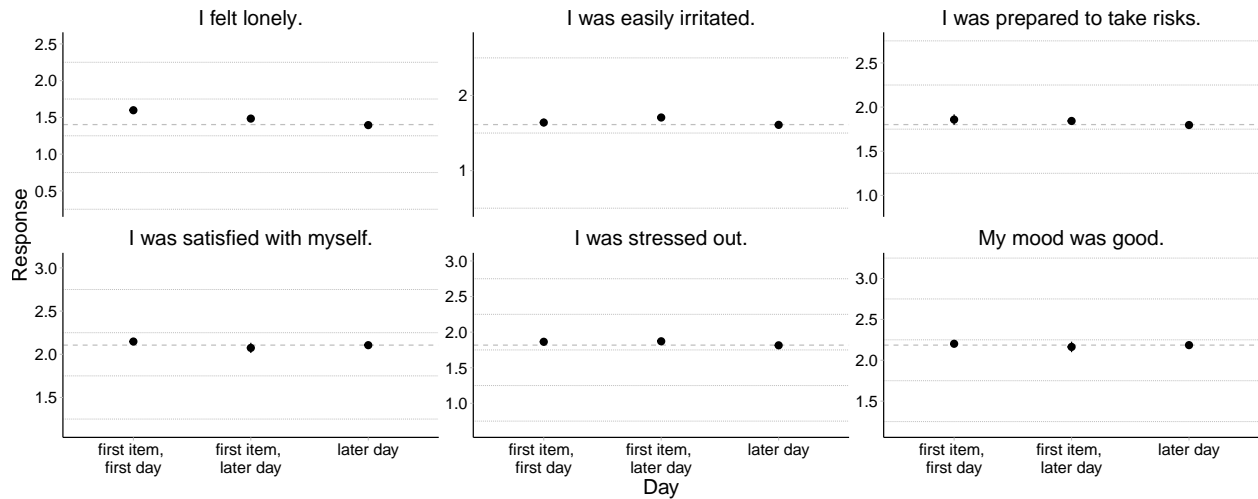
Figure 6: An alternative, more focused visualisation of initial elevation bias. Are responses elevated when shown for the first time, even if it is not the first day of the diary? The dashed line shows the overall mean for the item. Standard errors cannot be seen, because they are so narrow (and ignore the multilevel structure of the data)

4
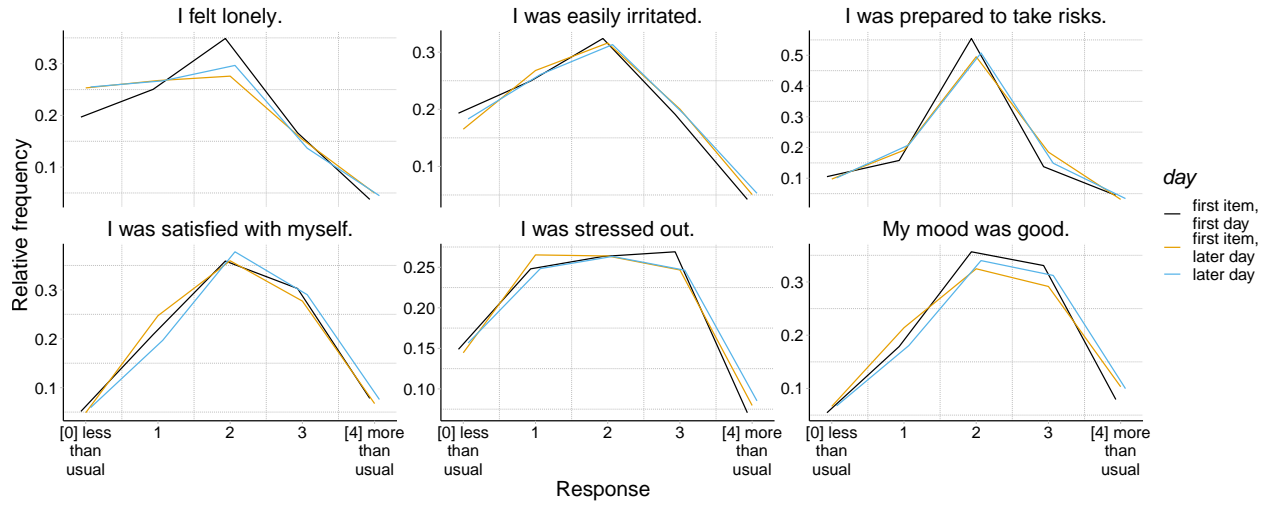
Figure 7: A response bias to a Likert scale may not only affect the mean response, but also the dispersion or the propensity to choose the middle or extreme categories. Such biases would balance out and not show up in the mean response. We therefore compute the relative frequency of certain responses for first days and later days. Days on which the item is first shown have very similar response distributions as later days.

2.1.1 Response time
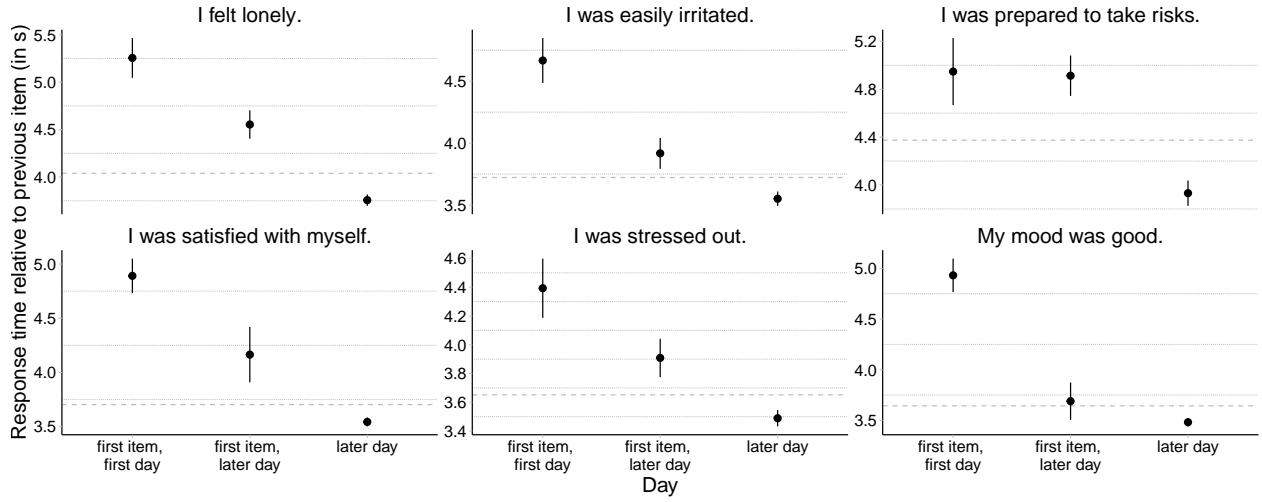
Figure 8: We can additionally examine whether responses slow down when items are first shown. We only examine the response time relative to the answer to the previous item here. This means the first item on each page is excluded from consideration. We do this, because responses relative to the time the page loaded are strongly biased upwards through participants who clicked the link and did something else until the page loaded, or participants who first familiarise themselves with all items. We also excluded responses that were made out of order (negative response times relative to the previous item), and responses that took longer than 30 seconds. Responses to the very first item take almost 8000ms, much longer than responses to later items. The dashed line shows the overall mean for the item. The points show means plus standard errors.

Figure 9: We can also examine response times as in Figure 2, but switching to the time series view somewhat clutters the display. Limited to combinations with at least 20 responses. Lines show 10% trimmed means.

2.1.2 Only participants who participated every day in the first week
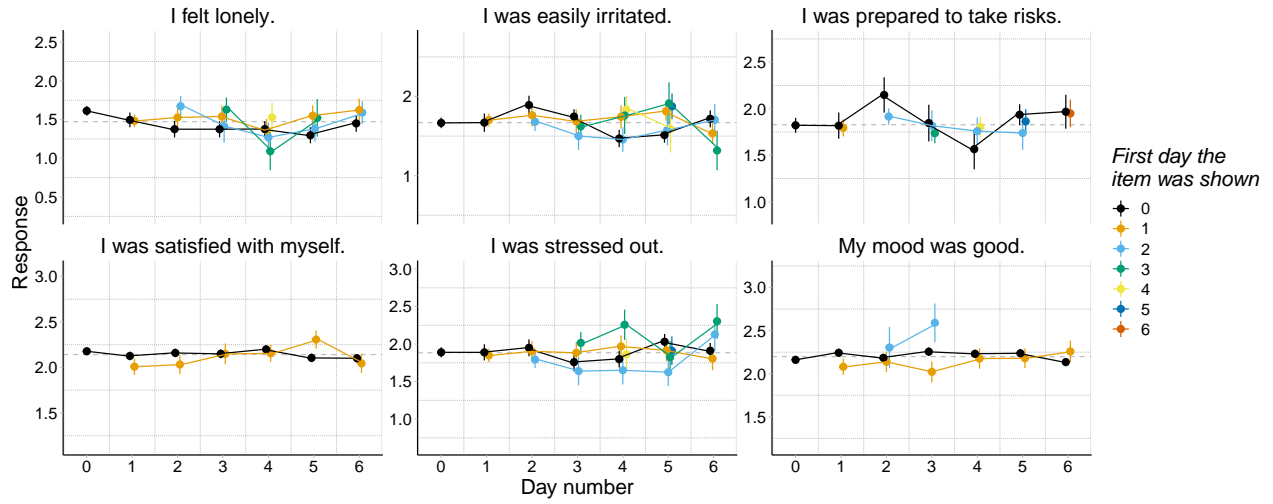
Figure 10: Figure 2, but only with those who didn't miss a day in the first week (ruling out selective dropout as an explanation). Patterns seem unchanged.

2.2 Item order
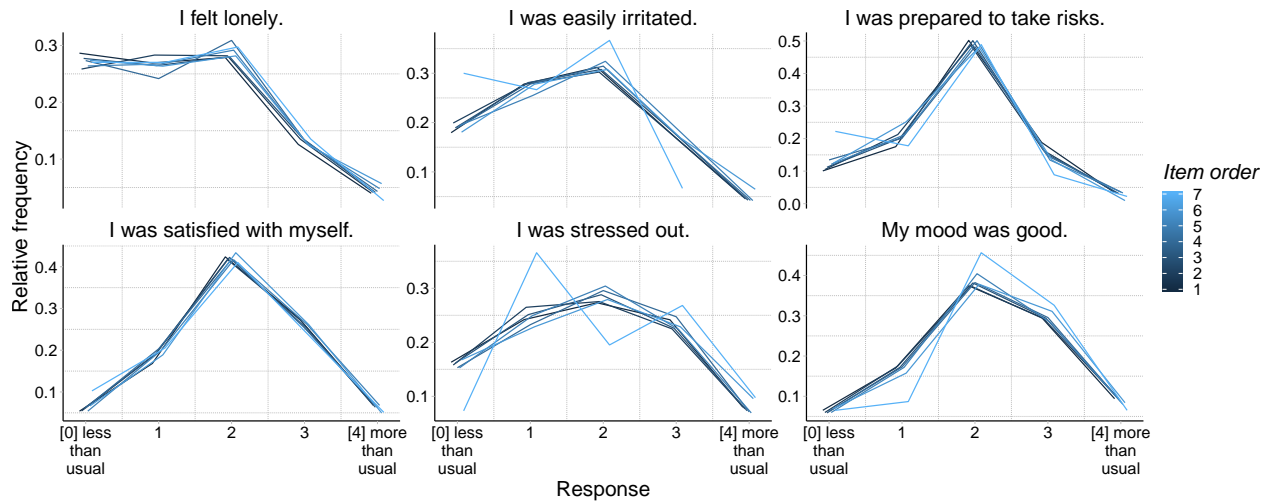
Figure 11: As above, we can not only examine the mean but also the response distribution according to item order.
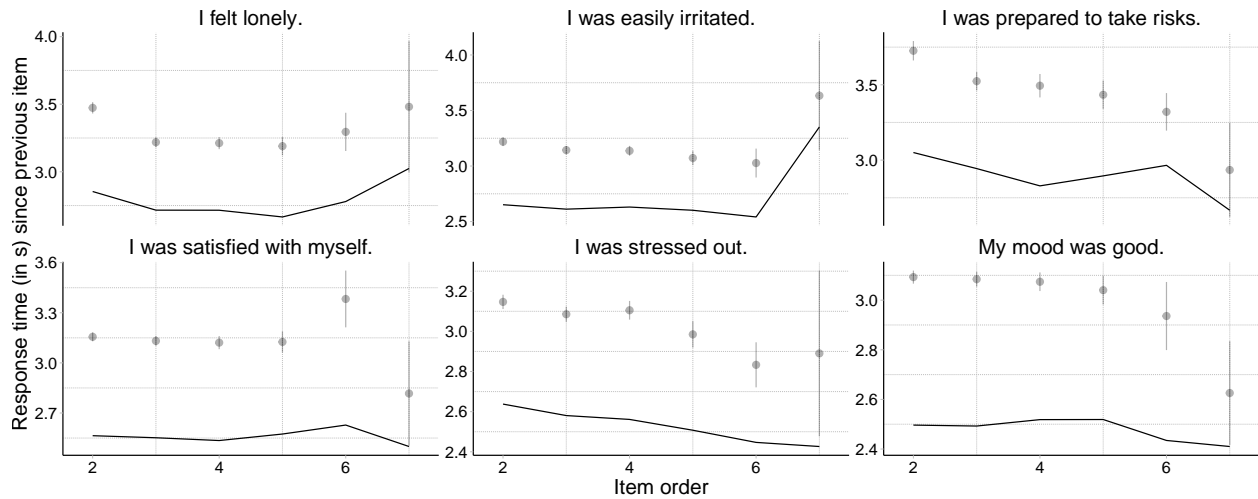
Figure 12: Again, we can also examine the response time to each item according to item order. The line shows the 10% trimmed means, the points show means plus standard errors. We excluded responses that were made out of order (negative response times relative to the previous item), and responses that took longer than 30 seconds.

2.3 Number of items shown

As mentioned above, the number of items shown is also a randomised variable. It is in turn confounded with item order, because items shown on a page with more items are likely to have a later item order. Most importantly, when only one item is shown, item order is necessarily 1.
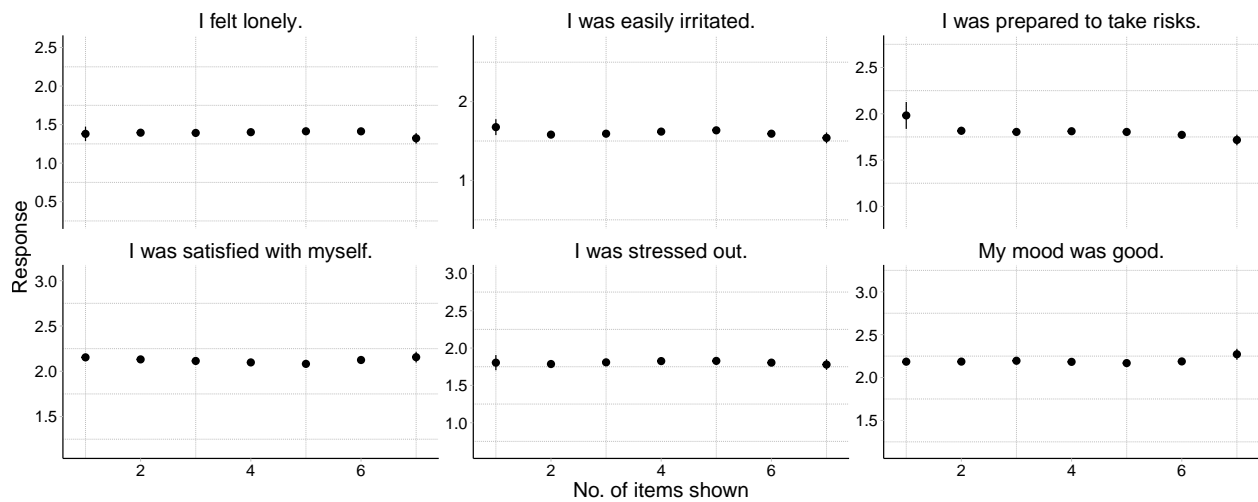
Figure 13: Response means according to the number of items shown. The Y axis scale is displayed from each item's mean $\pm$ 1 SD; values ranged from 0 to 4. The standard errors for the means do not account for the person-level structure of the data.
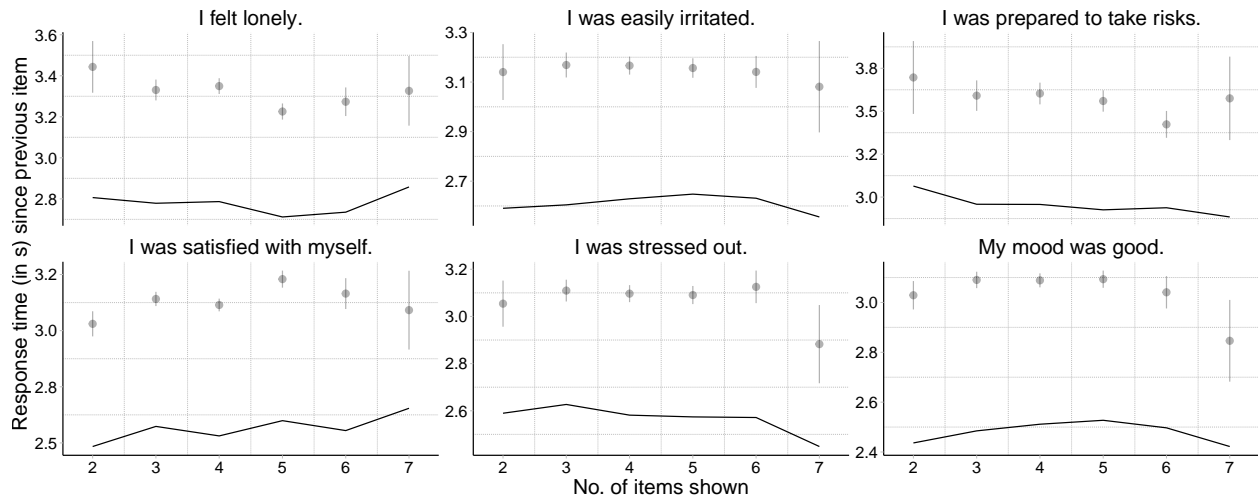
Figure 14: Again, we can also examine the response time to each item according to number of items shown. The line shows the 10% trimmed means, the points show means plus standard errors. We excluded responses that were made out of order (negative response times relative to the previous item), and responses that took longer than 30 seconds.

2.4 Last item

Figure 15: A different way to think about item order is to consider the identity of the immediately preceding item. The Y axis here shows which item (if any) preceded the given item. The X axes include the global mean ± 1 global standard deviation for each item. Standard errors (not visible) do not account for the multilevel structure of the data.
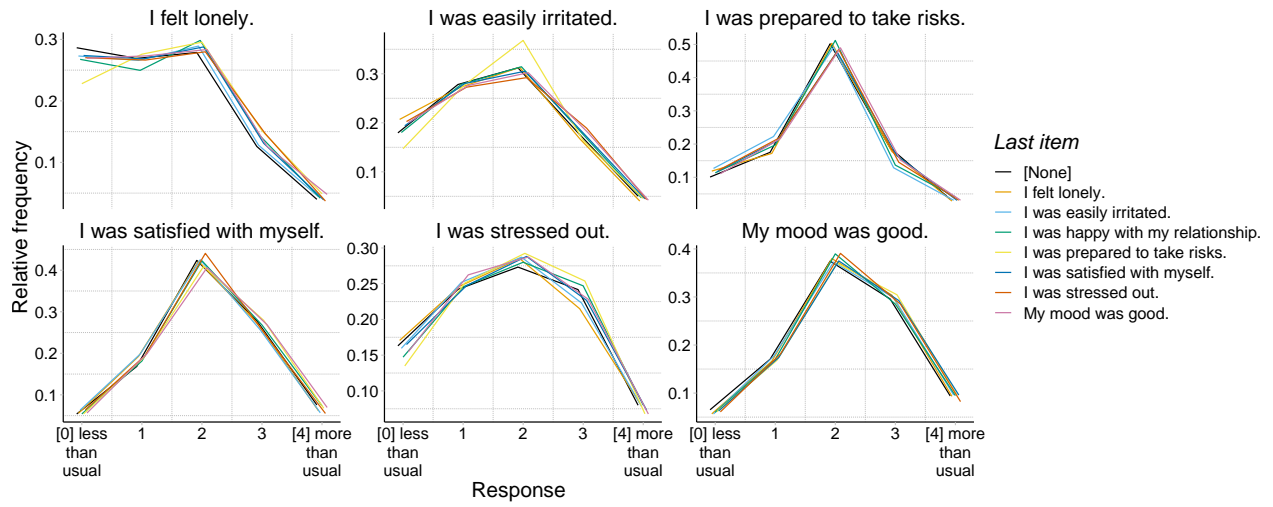
Figure 16: As above, we can not only examine the mean but also the distribution of each response.

3 Multilevel analysis
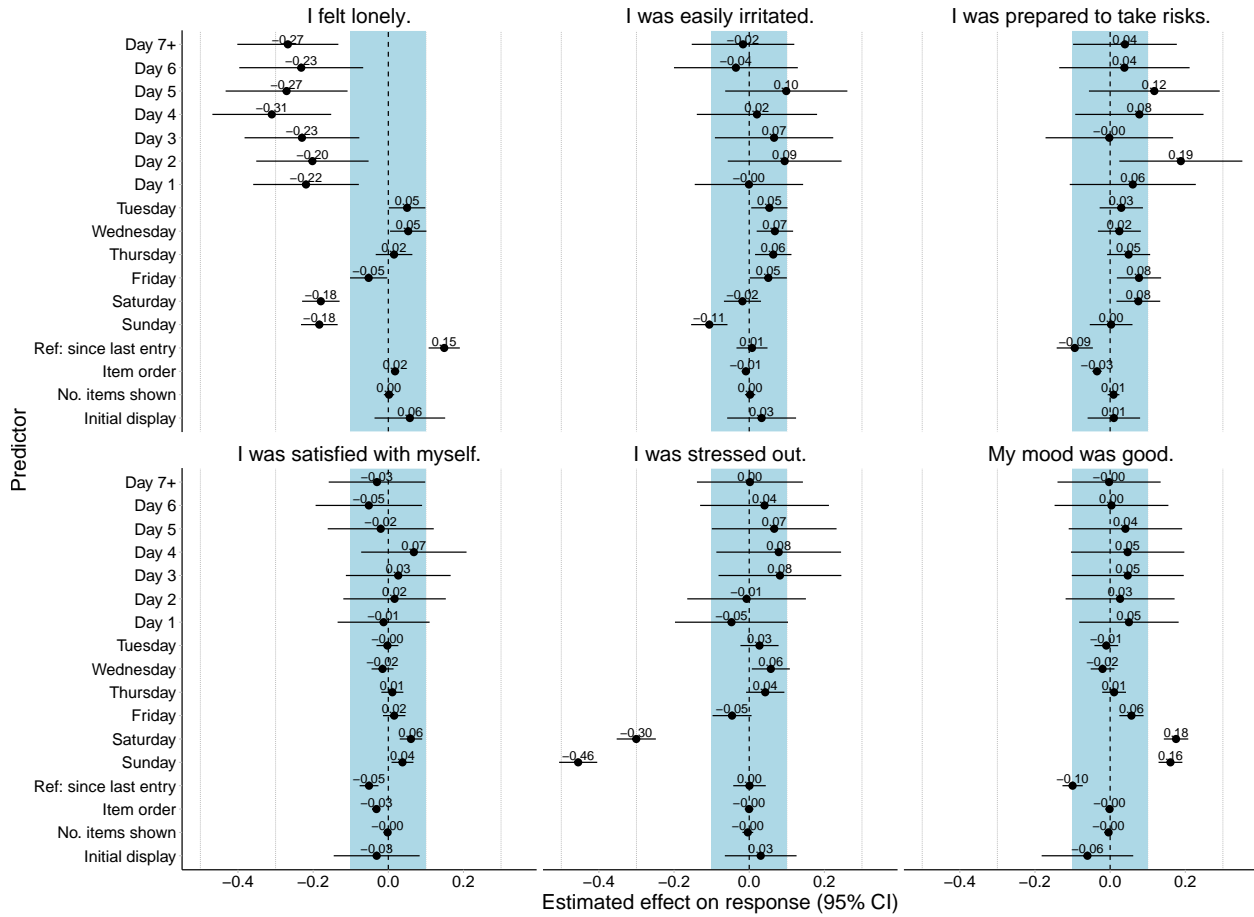
Figure 17: Figure 6 only showed the randomised variables. Here, we additionally show the estimates for the covariates.

Table 1: Residual correlation

Response variable	Correlation
I felt lonely.	1.00
I was easily irritated.	1.00
I was prepared to take risks.	1.00
I was satisfied with myself.	1.00
I was stressed out.	1.00
My mood was good.	1.00

Note:
Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

Table 2: Multilevel regression coefficients

term	good_mood	irritable	loneliness	risk_taking	self_esteem	stressed
Intercept	2.24 [2.09;2.38]	1.62 [1.47;1.77]	1.52 [1.37;1.67]	1.86 [1.71;2.02]	2.24 [2.11;2.37]	1.93 [1.78;2.08]
Initial display	-0.06 [-0.18;0.06]	0.03 [-0.06;0.12]	0.06 [-0.04;0.15]	0.01 [-0.06;0.08]	-0.03 [-0.15;0.08]	0.03 [-0.06;0.13]
No. items shown	-0.00 [-0.01;0.00]	0.00 [-0.01;0.02]	0.00 [-0.01;0.02]	0.01 [-0.01;0.02]	-0.00 [-0.01;0.01]	-0.00 [-0.02;0.01]
Item order	-0.00 [-0.01;0.01]	-0.01 [-0.02;0.00]	0.02 [0.01;0.03]	-0.03 [-0.05;-0.02]	-0.03 [-0.04;-0.02]	-0.00 [-0.01;0.01]
Ref: since last entry	-0.10 [-0.13;-0.07]	0.01 [-0.03;0.05]	0.15 [0.11;0.19]	-0.09 [-0.14;-0.05]	-0.05 [-0.08;-0.03]	0.00 [-0.04;0.04]
Sunday	0.16 [0.13;0.19]	-0.11 [-0.15;-0.06]	-0.18 [-0.23;-0.13]	0.00 [-0.05;0.06]	0.04 [0.01;0.07]	-0.46 [-0.51;-0.40]
Saturday	0.18 [0.14;0.21]	-0.02 [-0.07;0.03]	-0.18 [-0.23;-0.13]	0.08 [0.02;0.13]	0.06 [0.03;0.09]	-0.30 [-0.35;-0.25]
Friday	0.06 [0.02;0.09]	0.05 [0.00;0.10]	-0.05 [-0.10;-0.00]	0.08 [0.02;0.14]	0.02 [-0.01;0.05]	-0.05 [-0.10;0.01]
Thursday	0.01 [-0.02;0.04]	0.06 [0.02;0.11]	0.02 [-0.03;0.06]	0.05 [-0.01;0.11]	0.01 [-0.02;0.04]	0.04 [-0.01;0.09]
Wednesday	-0.02 [-0.05;0.01]	0.07 [0.02;0.12]	0.05 [0.00;0.10]	0.02 [-0.03;0.08]	-0.02 [-0.04;0.01]	0.06 [0.01;0.11]
Tuesday	-0.01 [-0.04;0.02]	0.05 [0.01;0.10]	0.05 [0.00;0.10]	0.03 [-0.03;0.09]	-0.00 [-0.03;0.03]	0.03 [-0.02;0.08]
Day 1	0.05 [-0.08;0.18]	-0.00 [-0.14;0.14]	-0.22 [-0.36;-0.08]	0.06 [-0.11;0.23]	-0.01 [-0.13;0.11]	-0.05 [-0.20;0.10]
Day 2	0.03 [-0.12;0.17]	0.09 [-0.06;0.25]	-0.20 [-0.35;-0.05]	0.19 [0.02;0.35]	0.02 [-0.12;0.15]	-0.01 [-0.16;0.15]
Day 3	0.05 [-0.10;0.20]	0.07 [-0.09;0.22]	-0.23 [-0.38;-0.08]	-0.00 [-0.17;0.17]	0.03 [-0.11;0.17]	0.08 [-0.08;0.25]
Day 4	0.05 [-0.10;0.20]	0.02 [-0.14;0.18]	-0.31 [-0.47;-0.15]	0.08 [-0.09;0.25]	0.07 [-0.07;0.21]	0.08 [-0.09;0.24]
Day 5	0.04 [-0.11;0.19]	0.10 [-0.06;0.26]	-0.27 [-0.43;-0.11]	0.12 [-0.06;0.29]	-0.02 [-0.16;0.12]	0.07 [-0.10;0.23]
Day 6	0.00 [-0.15;0.16]	-0.04 [-0.20;0.13]	-0.23 [-0.40;-0.07]	0.04 [-0.14;0.21]	-0.05 [-0.19;0.09]	0.04 [-0.13;0.21]
Day 7+	-0.00 [-0.14;0.13]	-0.02 [-0.15;0.12]	-0.27 [-0.40;-0.13]	0.04 [-0.10;0.18]	-0.03 [-0.16;0.10]	0.00 [-0.14;0.14]
sd (Intercept)	0.35	0.41	0.47	0.42	0.39	0.42
sd Observation	0.96	1.04	1.03	0.85	0.89	1.09

Note:
Estimates from the multilevel regression model displayed in Figure 6, with 95% confidence intervals.

3.1 Robustness checks

Here, we examine several alternative model specifications to see whether the results are robust to additional constraints or relaxed assumptions.

3.1.1 Complete first week
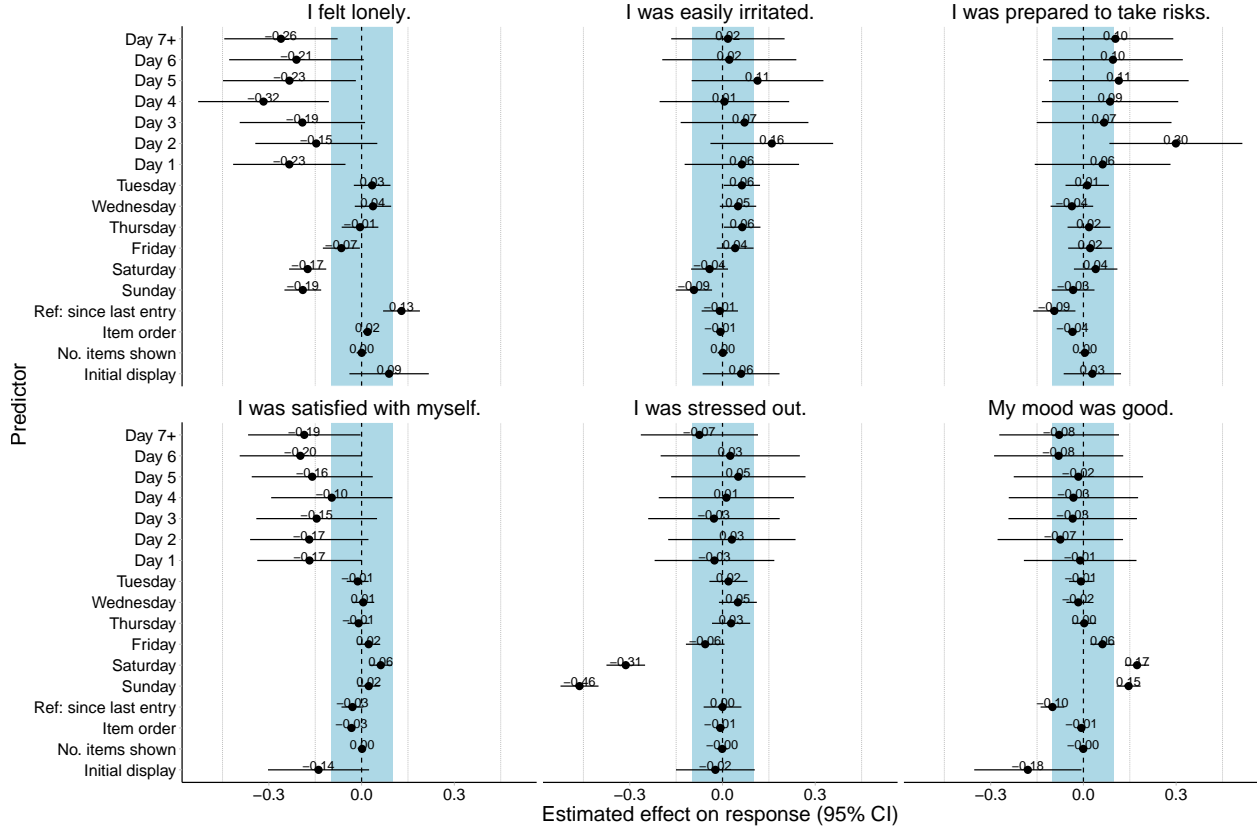
Figure 18: Here, we fit the same model as before, but include only participants who participated the entire first week.

Table 3: Residual correlation

Response variable	Correlation
I felt lonely.	1.00
I was easily irritated.	1.00
I was prepared to take risks.	1.00
I was satisfied with myself.	1.00
I was stressed out.	1.00
My mood was good.	1.00

Note:
Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

3.1.2 Last item identity
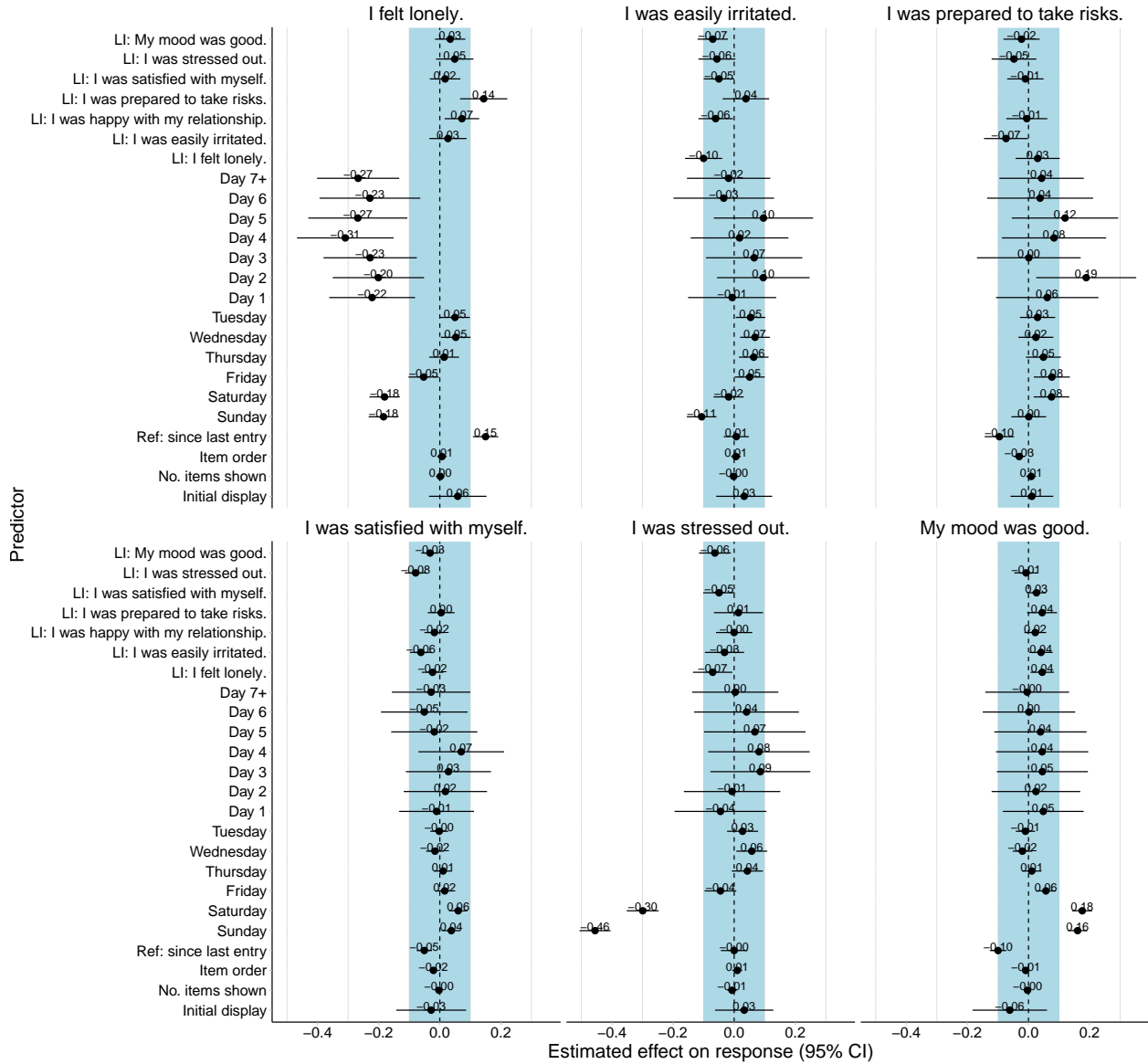
Figure 19: Here, we additionally test for an effect of last item identity (which is heavily collinear with item order).

Table 5: Model comparison

response	df	AIC	logLik	statistic	Chi.Df	p.value
I felt lonely.	8	72991	-36488	NA	NA	NA
I felt lonely.	14	72986	-36479	17.6	6	0.01
I was easily irritated.	8	73125	-36554	NA	NA	NA
I was easily irritated.	14	73117	-36544	20.4	6	0.00
I was prepared to take risks.	8	32023	-16003	NA	NA	NA
I was prepared to take risks.	14	32025	-15999	9.2	6	0.16

14

I was satisfied with myself.	8	129959	-64972	NA	NA	NA
I was satisfied with myself.	14	129941	-64956	30.6	6	0.00
I was stressed out.	8	75664	-37824	NA	NA	NA
I was stressed out.	14	75663	-37817	13.9	6	0.03
My mood was good.	8	137637	-68810	NA	NA	NA
My mood was good.	14	137636	-68804	12.4	6	0.05

Note:

Do models additionally accounting for last item identity fit better? We test this separately because item order 1 and last item "[None]" are identical, and estimates become less precise because of similar, less severe multicollinearity.

3.1.3 Varying effects
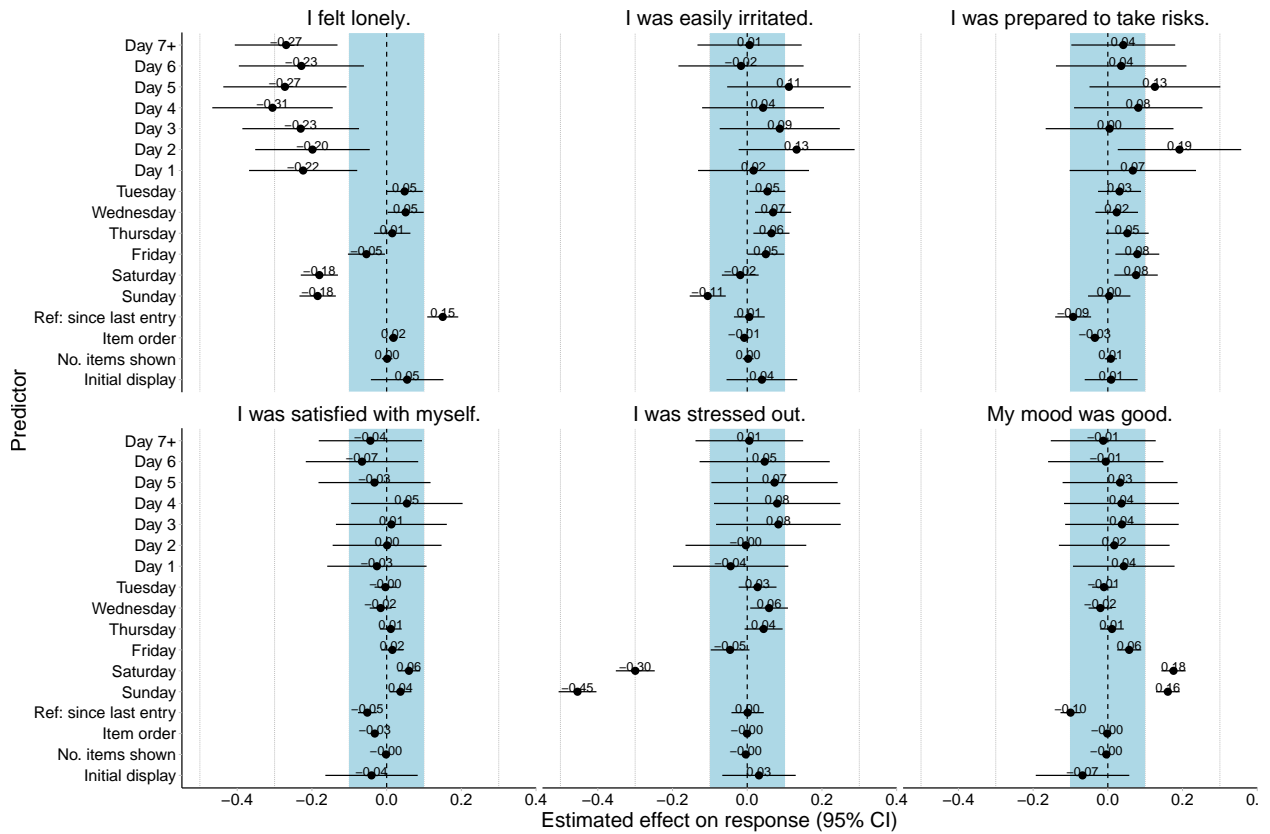
Figure 20: Here, we allow the effect of the randomised variables to vary by woman.

Table 4: Residual correlation

Response variable	Correlation
I felt lonely.	1.00
I was easily irritated.	1.00
I was prepared to take risks.	1.00
I was satisfied with myself.	1.00
I was stressed out.	1.00
My mood was good.	1.00

Note:
Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

Table 6: Residual correlation

Response variable	Correlation
I was prepared to take risks.	0.99
I felt lonely.	1.00
I was easily irritated.	1.00
I was satisfied with myself.	1.00
I was stressed out.	1.00
My mood was good.	1.00

Note:
Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

Table 7: Multilevel regression coefficients

term	good_mood	irritable	loneliness	risk_taking	self_esteem	stressed
Intercept	2.24 [2.10;2.39]	1.59 [1.44;1.75]	1.53 [1.38;1.68]	1.86 [1.70;2.02]	2.25 [2.11;2.40]	1.93 [1.77;2.08]
Initial display	-0.07 [-0.19;0.06]	0.04 [-0.06;0.13]	0.05 [-0.04;0.15]	0.01 [-0.06;0.08]	-0.04 [-0.16;0.08]	0.03 [-0.07;0.13]
No. items shown	-0.00 [-0.01;0.01]	0.00 [-0.01;0.02]	0.00 [-0.01;0.02]	0.01 [-0.01;0.02]	-0.00 [-0.01;0.01]	-0.00 [-0.02;0.01]
Item order	-0.00 [-0.01;0.01]	-0.01 [-0.02;0.00]	0.02 [0.01;0.03]	-0.03 [-0.05;-0.02]	-0.03 [-0.04;-0.02]	-0.00 [-0.01;0.01]
Ref: since last entry	-0.10 [-0.13;-0.07]	0.01 [-0.04;0.05]	0.15 [0.11;0.19]	-0.09 [-0.14;-0.04]	-0.05 [-0.08;-0.03]	0.00 [-0.04;0.04]
Sunday	0.16 [0.13;0.19]	-0.11 [-0.15;-0.06]	-0.18 [-0.23;-0.14]	0.00 [-0.05;0.06]	0.04 [0.01;0.07]	-0.45 [-0.50;-0.40]
Saturday	0.18 [0.14;0.21]	-0.02 [-0.07;0.03]	-0.18 [-0.23;-0.13]	0.08 [0.02;0.13]	0.06 [0.03;0.09]	-0.30 [-0.35;-0.25]
Friday	0.06 [0.03;0.09]	0.05 [0.00;0.10]	-0.05 [-0.10;-0.00]	0.08 [0.02;0.14]	0.02 [-0.01;0.05]	-0.05 [-0.10;0.01]
Thursday	0.01 [-0.02;0.04]	0.06 [0.02;0.11]	0.01 [-0.03;0.06]	0.05 [-0.00;0.11]	0.01 [-0.02;0.04]	0.04 [-0.01;0.09]

Wednesday	-0.02 [-0.05;0.01]	0.07 [0.02;0.12]	0.05 [0.00;0.10]	0.02 [-0.03;0.08]	-0.02 [-0.05;0.01]	0.06 [0.01;0.11]
Tuesday	-0.01 [-0.04;0.02]	0.05 [0.01;0.10]	0.05 [-0.00;0.10]	0.03 [-0.03;0.09]	-0.00 [-0.03;0.03]	0.03 [-0.02;0.08]
cor (Intercept) Initial display	-0.48	-0.31	-0.46	-0.07	-0.45	-0.63
cor (Intercept) Item order	0.07	-0.39	0.13	0.14	-0.20	-0.01
cor (Intercept) No. items shown	-0.48	-0.64	-0.27	-0.52	-0.59	-0.57
cor Item order Initial display	-0.17	-0.58	-0.01	-0.41	0.06	0.78
cor No. items shown Initial display	0.35	0.14	0.08	-0.09	0.41	0.99
cor No. items shown Item order	-0.91	0.38	-0.99	-0.68	0.81	0.82
Day 1	0.04 [-0.09;0.18]	0.02 [-0.13;0.17]	-0.22 [-0.37;-0.08]	0.07 [-0.10;0.24]	-0.03 [-0.16;0.11]	-0.04 [-0.20;0.11]
Day 2	0.02 [-0.13;0.17]	0.13 [-0.02;0.29]	-0.20 [-0.35;-0.05]	0.19 [0.03;0.36]	0.00 [-0.14;0.15]	-0.00 [-0.17;0.16]
Day 3	0.04 [-0.11;0.19]	0.09 [-0.07;0.25]	-0.23 [-0.39;-0.07]	0.00 [-0.17;0.18]	0.01 [-0.14;0.16]	0.08 [-0.08;0.25]
Day 4	0.04 [-0.12;0.19]	0.04 [-0.12;0.21]	-0.31 [-0.47;-0.14]	0.08 [-0.09;0.25]	0.05 [-0.09;0.20]	0.08 [-0.09;0.25]
Day 5	0.03 [-0.12;0.19]	0.11 [-0.05;0.28]	-0.27 [-0.44;-0.11]	0.13 [-0.05;0.30]	-0.03 [-0.18;0.12]	0.07 [-0.10;0.24]
Day 6	-0.01 [-0.16;0.15]	-0.02 [-0.18;0.15]	-0.23 [-0.40;-0.06]	0.04 [-0.14;0.21]	-0.07 [-0.22;0.08]	0.05 [-0.13;0.22]
Day 7+	-0.01 [-0.15;0.13]	0.01 [-0.13;0.15]	-0.27 [-0.41;-0.13]	0.04 [-0.10;0.18]	-0.04 [-0.18;0.09]	0.01 [-0.14;0.15]
sd (Intercept)	0.38	0.56	0.51	0.52	0.43	0.45
sd Initial display	0.32	0.40	0.40	0.22	0.47	0.39
sd Item order	0.05	0.04	0.07	0.07	0.03	0.02
sd No. items shown	0.02	0.07	0.05	0.07	0.02	0.02
sd Observation	0.96	1.03	1.03	0.84	0.88	1.08

Note:

Estimates from the multilevel regression model with varying slopes. With 95% confidence intervals.
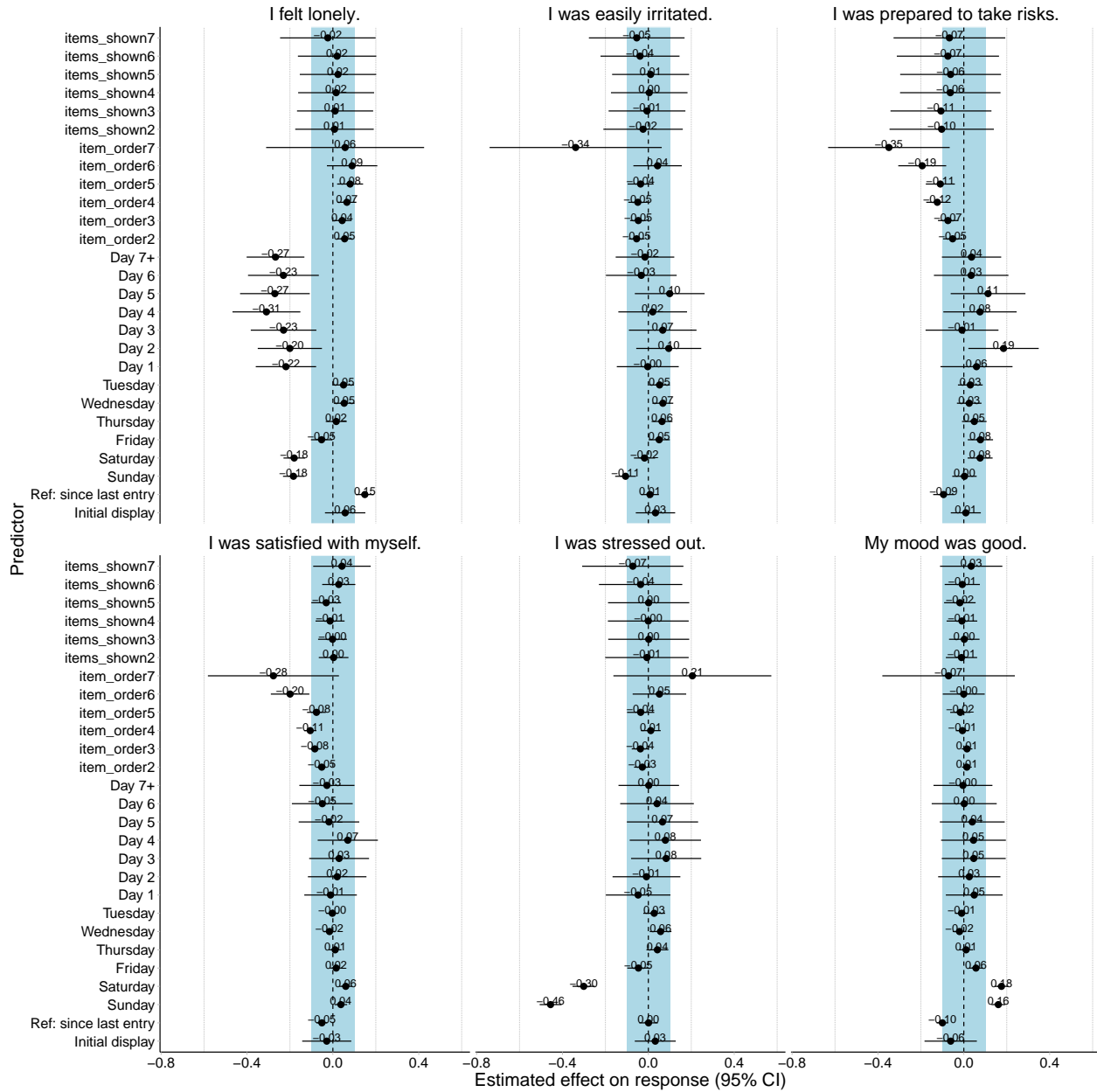
3.1.4 Nonlinear effects

Figure 21: Here, we allow the effects of the randomised variables to be nonlinear.

Table 8: Residual correlation

Response variable	Correlation
I felt lonely.	1.00
I was easily irritated.	1.00
I was prepared to take risks.	1.00
I was satisfied with myself.	1.00
I was stressed out.	1.00
My mood was good.	1.00

Note:
Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

4   Other items

We focused on the most general items on the first page of our study. However, the diary also contained randomised (in order and odds of appearing) items about sexual desire, time use, and partner jealousy, among others. We show that the overall results (that residualising for estimated biases has negligible effects) holds here too, even though the partner jealousy items were asked on a response scale from "not at all" to "very much", the desire items on a scale of "very inaccurate" to "very accurate", and the time use items on the same "less than usual" to "more than usual" scale as the items on the first page.

4.1   Time items

On a "less than usual" to "more than usual" response scale.

Table 9: How often did how many women see each item?

item_name	label_english	women	days
time_family	I looked for contact with my family.	1326	42701
time_friends	I spent time with friends.	1331	42856
time_people	I looked for contact with other people.	1330	42796
time_sports	I worked out.	1335	42591
time_work_uni	I spent time working/studying.	1328	42527
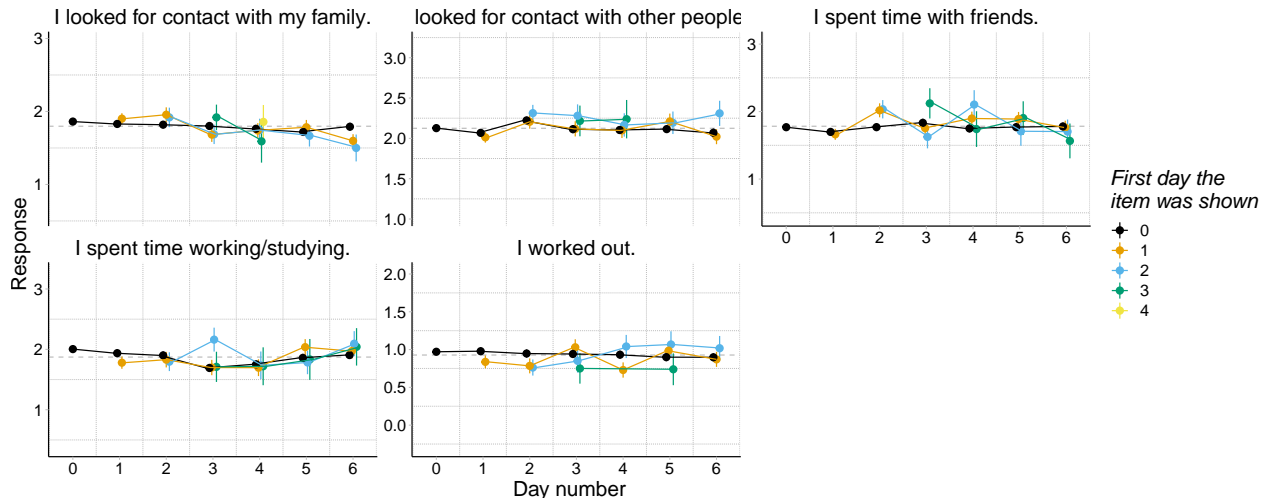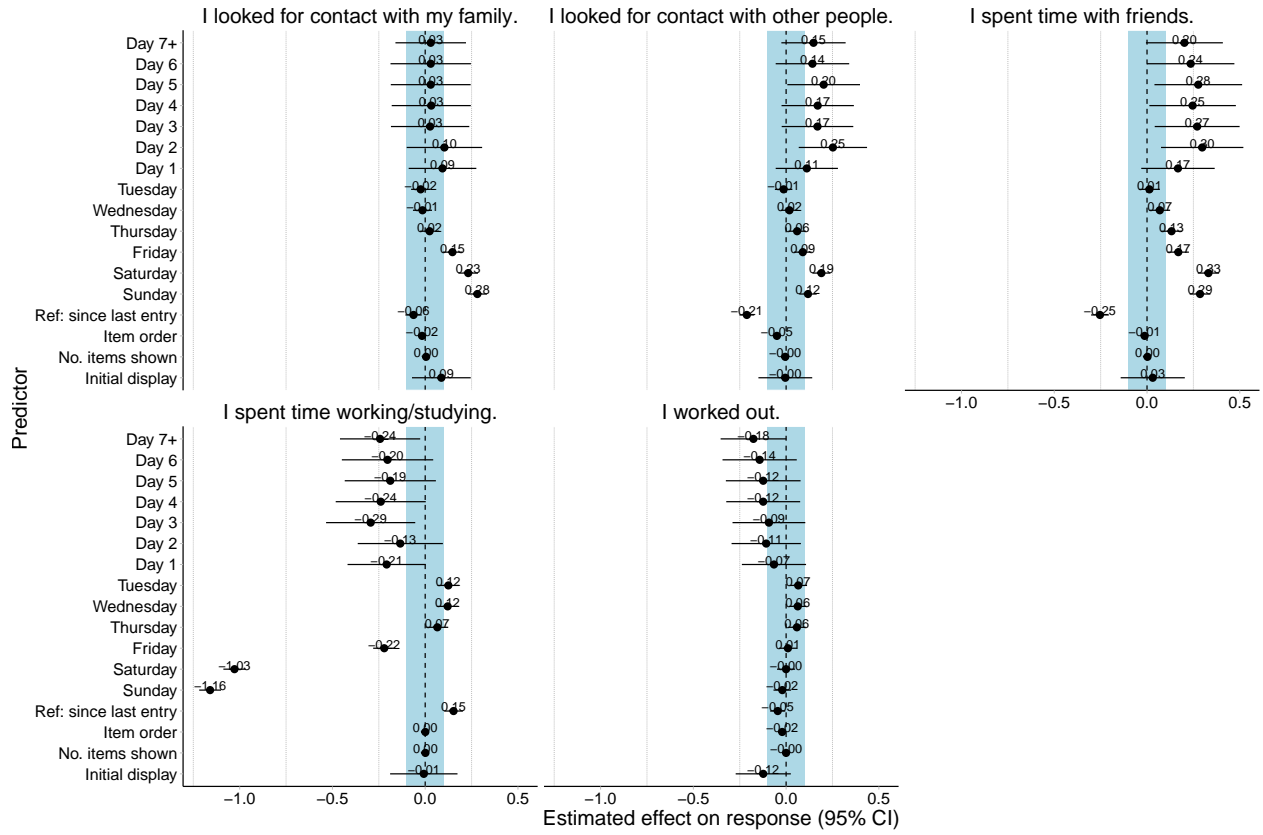
Figure 22: Estimates for biases and covariates for the time items.

Table 10: Residual correlation

Response variable	Correlation
I looked for contact with my family.	1.00
I looked for contact with other people.	1.00
I spent time with friends.	1.00
I spent time working/studying.	1.00
I worked out.	1.00

Note:
Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

Table 11: Multilevel regression coefficients

term	time_family	time_friends	time_people	time_sports	time_work_uni
Intercept	1.72	1.64	2.21	1.11	2.23
	[1.51;1.92]	[1.42;1.86]	[2.03;2.40]	[0.92;1.30]	[2.00;2.46]
Initial display	0.09	0.03	-0.00	-0.12	-0.01
	[-0.07;0.25]	[-0.14;0.20]	[-0.15;0.14]	[-0.27;0.03]	[-0.19;0.17]
No. items	0.00	0.00	-0.00	-0.00	0.00
shown	[-0.01;0.02]	[-0.02;0.02]	[-0.02;0.01]	[-0.02;0.02]	[-0.02;0.02]
Item order	-0.02	-0.01	-0.05	-0.02	0.00
	[-0.03;-0.00]	[-0.03;0.00]	[-0.06;-0.04]	[-0.03;-0.01]	[-0.01;0.02]
Ref: since last	-0.06	-0.25	-0.21	-0.05	0.15
entry	[-0.11;-0.02]	[-0.30;-0.21]	[-0.25;-0.17]	[-0.09;-0.00]	[0.10;0.20]
Sunday	0.28	0.29	0.12	-0.02	-1.16
	[0.23;0.33]	[0.23;0.34]	[0.07;0.16]	[-0.07;0.03]	[-1.22;-1.10]
Saturday	0.23	0.33	0.19	-0.00	-1.03
	[0.18;0.28]	[0.27;0.39]	[0.14;0.24]	[-0.05;0.05]	[-1.09;-0.97]
Friday	0.15	0.17	0.09	0.01	-0.22
	[0.09;0.20]	[0.11;0.23]	[0.04;0.14]	[-0.04;0.06]	[-0.28;-0.16]
Thursday	0.02	0.13	0.06	0.06	0.07
	[-0.03;0.08]	[0.08;0.19]	[0.01;0.11]	[0.01;0.11]	[0.01;0.12]
Wednesday	-0.01	0.07	0.02	0.06	0.12
	[-0.07;0.04]	[0.01;0.12]	[-0.03;0.06]	[0.02;0.11]	[0.06;0.18]
Tuesday	-0.02	0.01	-0.01	0.07	0.12
	[-0.08;0.03]	[-0.04;0.07]	[-0.06;0.03]	[0.02;0.11]	[0.07;0.18]
Day 1	0.09	0.17	0.11	-0.07	-0.21
	[-0.09;0.28]	[-0.03;0.36]	[-0.06;0.28]	[-0.24;0.11]	[-0.42;0.00]
Day 2	0.10	0.30	0.25	-0.11	-0.13
	[-0.10;0.31]	[0.08;0.52]	[0.07;0.44]	[-0.29;0.08]	[-0.36;0.10]
Day 3	0.03	0.27	0.17	-0.09	-0.29
	[-0.18;0.24]	[0.04;0.50]	[-0.02;0.36]	[-0.29;0.10]	[-0.53;-0.05]
Day 4	0.03	0.25	0.17	-0.12	-0.24
	[-0.18;0.25]	[0.01;0.48]	[-0.02;0.37]	[-0.32;0.08]	[-0.48;0.00]
Day 5	0.03	0.28	0.20	-0.12	-0.19
	[-0.19;0.25]	[0.04;0.51]	[0.01;0.40]	[-0.32;0.08]	[-0.43;0.06]
Day 6	0.03	0.24	0.14	-0.14	-0.20
	[-0.19;0.25]	[-0.00;0.47]	[-0.06;0.34]	[-0.34;0.06]	[-0.45;0.04]
Day 7+	0.03	0.20	0.15	-0.18	-0.24
	[-0.16;0.22]	[-0.01;0.41]	[-0.03;0.32]	[-0.35;0.00]	[-0.46;-0.03]
sd (Intercept)	0.45	0.47	0.39	0.57	0.54
sd Observation	1.11	1.21	1.02	1.03	1.26

Note:
Estimates from the multilevel regression model, with 95% confidence intervals.

4.2   Desire items

On a "very inaccurate" to "very accurate" response scale.

Table 12: How often did how many women see each item?

label_english	women	days
I fantasized about having sex with my partner.	862	12003
I fantasized about tenderness with my partner.	859	12064
I felt attracted towards my partner.	860	12108
I felt sexually attracted to my partner.	859	12126
I thought my partner was attractive.	865	12229
I wanted to be sexually active with my partner.	859	12053
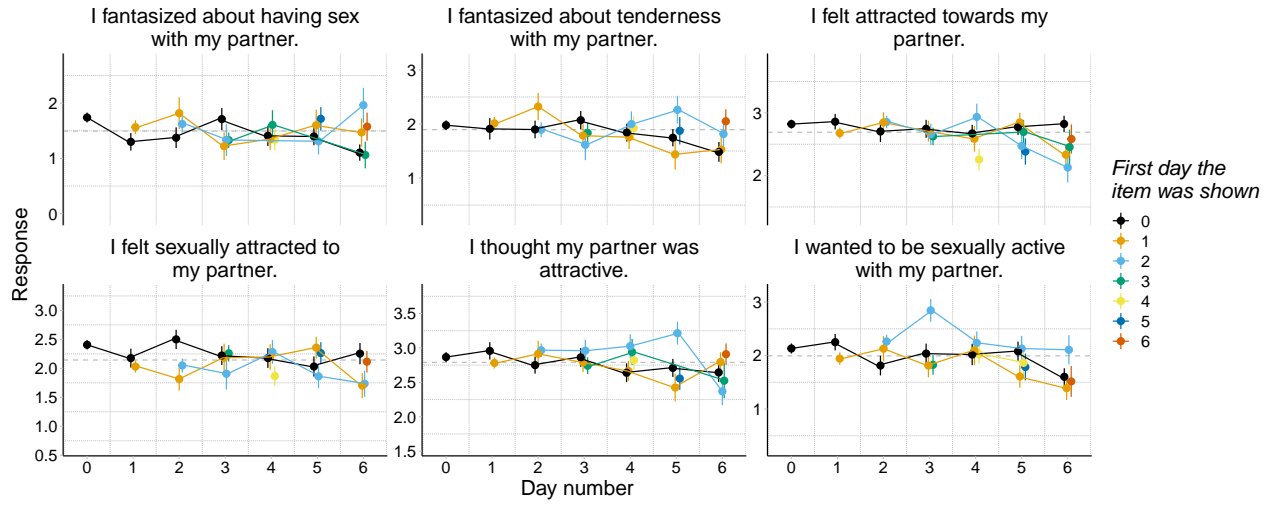
Figure 23: Estimates for biases and covariates for the desire items.

Table 13: Residual correlation

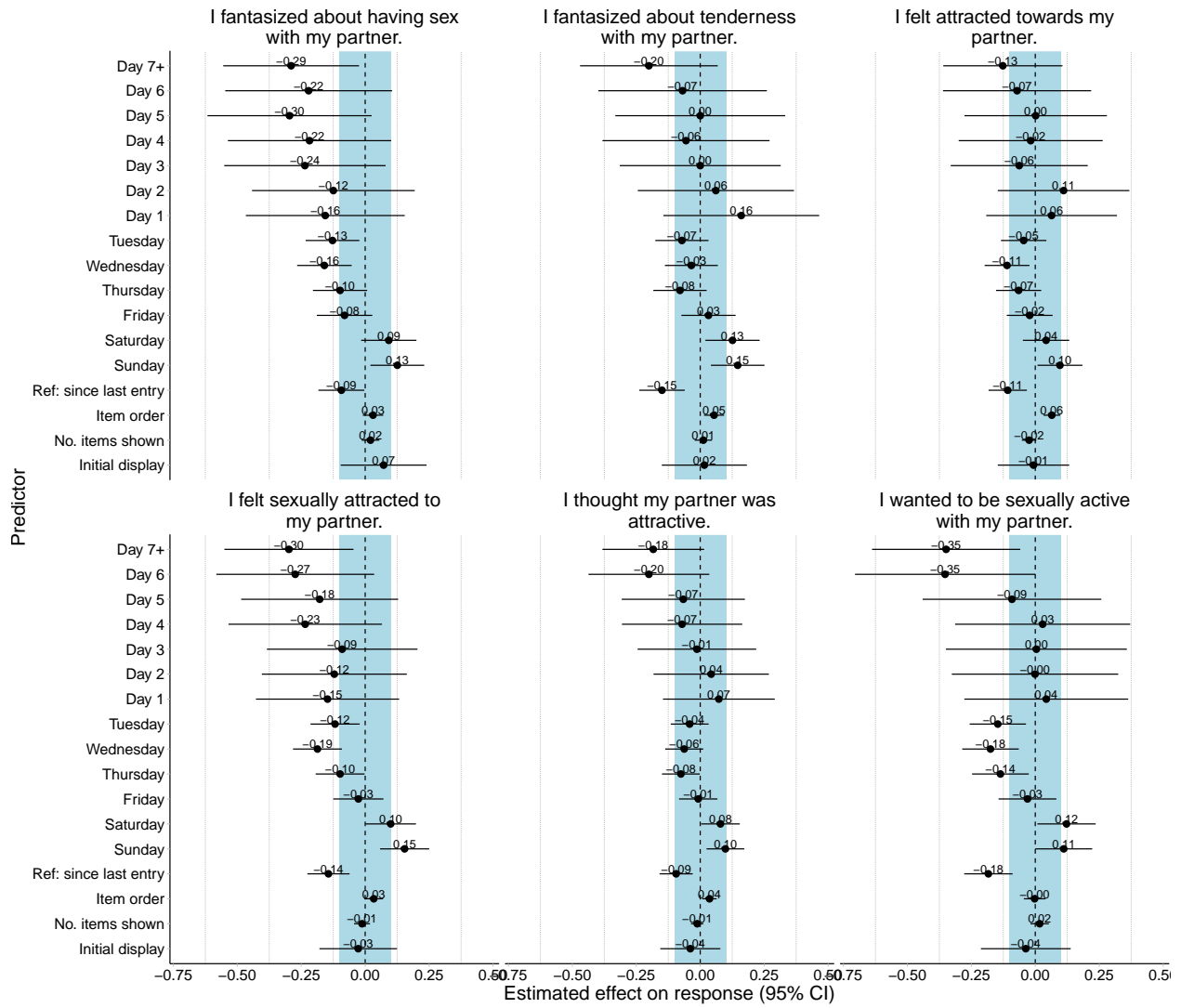| Response variable	Correlation
I fantasized about having sex with my partner.	1.00
I fantasized about tenderness with my partner.	1.00
I felt attracted towards my partner.	1.00
I felt sexually attracted to my partner.	1.00
I thought my partner was attractive.	1.00
I wanted to be sexually active with my partner.	1.00

Note:

Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

Table 14: Multilevel regression coefficients

term	desire_10	desire_11	desire_13	desire_14	desire_7	desire_8
Intercept	2.37	2.21	1.87	1.64	2.74	2.91
	[2.10;2.64]	[1.90;2.51]	[1.58;2.15]	[1.35;1.92]	[2.49;2.99]	[2.70;3.13]
Initial display	-0.03	-0.04	0.02	0.07	-0.01	-0.04
	[-0.18;0.12]	[-0.21;0.14]	[-0.15;0.18]	[-0.10;0.24]	[-0.15;0.13]	[-0.16;0.08]
No. items shown	-0.01	0.02	0.01	0.02	-0.02	-0.01
	[-0.04;0.02]	[-0.02;0.05]	[-0.02;0.05]	[-0.01;0.06]	[-0.05;0.00]	[-0.04;0.01]
Item order	0.03	-0.00	0.05	0.03	0.06	0.04
	[-0.00;0.07]	[-0.04;0.04]	[0.01;0.09]	[-0.01;0.07]	[0.03;0.10]	[0.01;0.06]
Ref: since last entry	-0.14	-0.18	-0.15	-0.09	-0.11	-0.09
	[-0.23;-0.06]	[-0.28;-0.09]	[-0.24;-0.06]	[-0.18;-0.00]	[-0.18;-0.03]	[-0.16;-0.03]
Sunday	0.15	0.11	0.15	0.13	0.10	0.10
	[0.06;0.25]	[0.00;0.22]	[0.04;0.25]	[0.02;0.23]	[0.01;0.18]	[0.02;0.17]
Saturday	0.10	0.12	0.13	0.09	0.04	0.08
	[0.00;0.20]	[0.01;0.24]	[0.02;0.23]	[-0.02;0.20]	[-0.05;0.13]	[0.00;0.15]
Friday	-0.03	-0.03	0.03	-0.08	-0.02	-0.01
	[-0.12;0.07]	[-0.14;0.08]	[-0.07;0.14]	[-0.19;0.03]	[-0.11;0.07]	[-0.08;0.07]
Thursday	-0.10	-0.14	-0.08	-0.10	-0.07	-0.08
	[-0.19;-0.00]	[-0.25;-0.03]	[-0.18;0.03]	[-0.20;0.01]	[-0.15;0.02]	[-0.15;-0.00]
Wednesday	-0.19	-0.18	-0.03	-0.16	-0.11	-0.06
	[-0.28;-0.09]	[-0.29;-0.06]	[-0.14;0.07]	[-0.27;-0.05]	[-0.20;-0.02]	[-0.14;0.01]
Tuesday	-0.12	-0.15	-0.07	-0.13	-0.05	-0.04
	[-0.21;-0.02]	[-0.26;-0.04]	[-0.18;0.03]	[-0.23;-0.02]	[-0.13;0.04]	[-0.12;0.03]
Day 1	-0.15	0.04	0.16	-0.16	0.06	0.07
	[-0.43;0.13]	[-0.28;0.36]	[-0.14;0.47]	[-0.47;0.15]	[-0.19;0.32]	[-0.15;0.29]
Day 2	-0.12	-0.00	0.06	-0.12	0.11	0.04
	[-0.40;0.16]	[-0.33;0.32]	[-0.24;0.37]	[-0.44;0.19]	[-0.15;0.37]	[-0.18;0.27]
Day 3	-0.09	0.00	0.00	-0.24	-0.06	-0.01
	[-0.38;0.20]	[-0.35;0.36]	[-0.31;0.31]	[-0.55;0.08]	[-0.33;0.21]	[-0.24;0.22]

|						
Day 4	-0.23	0.03	-0.06	-0.22	-0.02	-0.07
	[-0.53;0.07]	[-0.31;0.37]	[-0.38;0.27]	[-0.54;0.10]	[-0.30;0.26]	[-0.31;0.16]
Day 5	-0.18	-0.09	0.00	-0.30	0.00	-0.07
	[-0.48;0.13]	[-0.44;0.26]	[-0.33;0.33]	[-0.62;0.02]	[-0.28;0.28]	[-0.31;0.17]
Day 6	-0.27	-0.35	-0.07	-0.22	-0.07	-0.20
	[-0.58;0.04]	[-0.70;-0.00]	[-0.40;0.26]	[-0.55;0.11]	[-0.36;0.22]	[-0.44;0.04]
Day 7+	-0.30	-0.35	-0.20	-0.29	-0.13	-0.18
	[-0.55;-0.05]	[-0.64;-0.06]	[-0.47;0.07]	[-0.55;-0.02]	[-0.36;0.11]	[-0.38;0.02]
sd (Intercept)	0.79	0.74	0.83	0.74	0.74	0.74
sd Observation	1.08	1.25	1.18	1.20	0.99	0.84

Note:

Estimates from the multilevel regression model, with 95% confidence intervals.

4.3 Retention items

On a "not at all" to "very much" response scale.

Table 15: How often did how many women see each item?

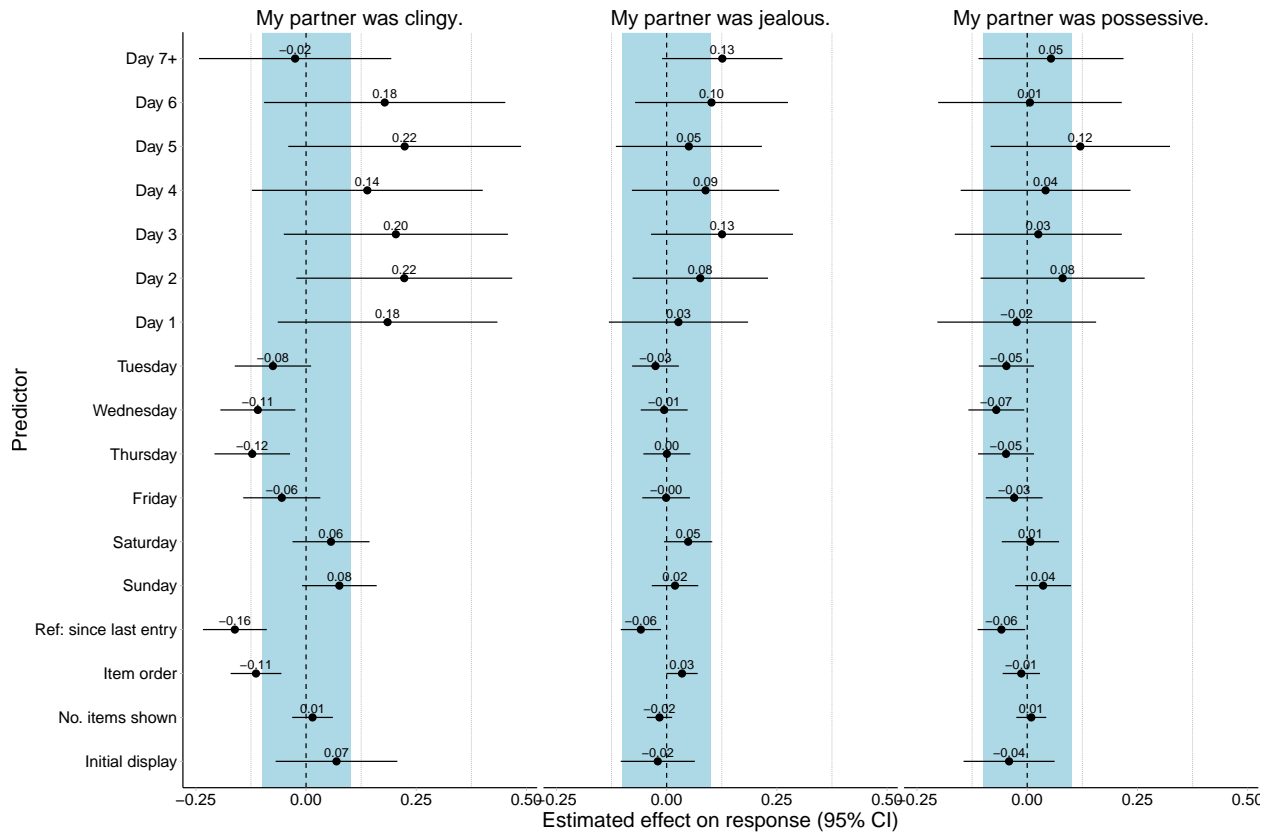| item_name	label_english	women	days
mate_retention1	My partner was clingy.	864	12300
mate_retention2	My partner was jealous.	860	12125
mate_retention3	My partner was possessive.	858	12109
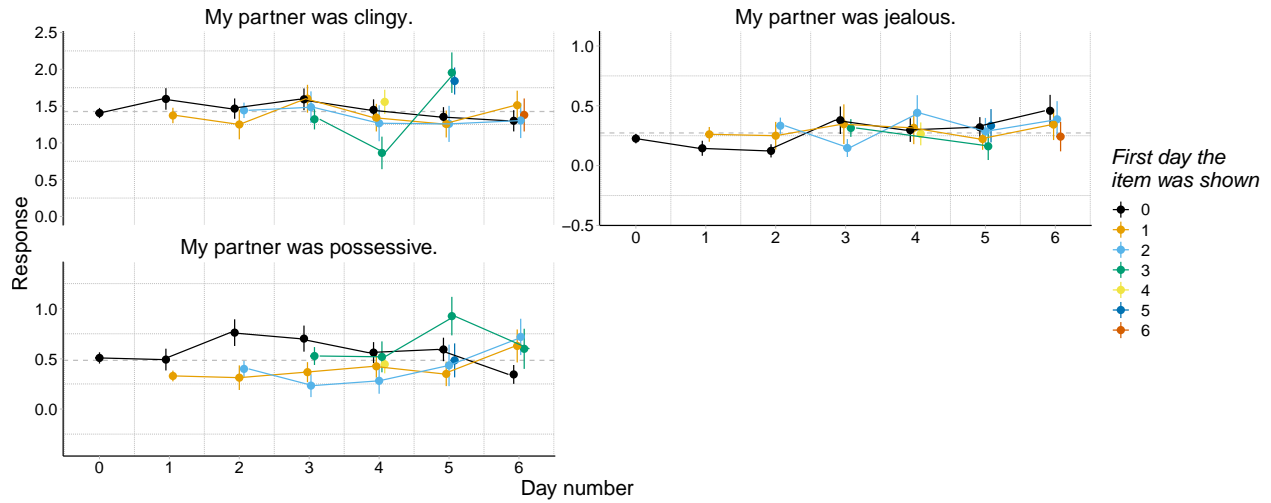
Figure 24: Estimates for biases and covariates for the mate retention items.

Table 16: Residual correlation

Response variable	Correlation
My partner was clingy.	1.00
My partner was jealous.	1.00
My partner was possessive.	1.00

Note:
Correlation between the residuals of a covariate-only model and a model additionally including the randomised variables.

Table 17: Multilevel regression coefficients

term	mate_retention1	mate_retention2	mate_retention3
Intercept	1.49 [1.26;1.73]	0.23 [0.08;0.38]	0.54 [0.36;0.72]
Initial display	0.07 [-0.07;0.21]	-0.02 [-0.10;0.06]	-0.04 [-0.14;0.06]
No. items shown	0.01 [-0.03;0.06]	-0.02 [-0.04;0.01]	0.01 [-0.02;0.04]
Item order	-0.11 [-0.17;-0.06]	0.03 [-0.00;0.07]	-0.01 [-0.06;0.03]
Ref: since last entry	-0.16 [-0.23;-0.09]	-0.06 [-0.10;-0.01]	-0.06 [-0.11;-0.00]
Sunday	0.08 [-0.01;0.16]	0.02 [-0.03;0.07]	0.04 [-0.03;0.10]
Saturday	0.06 [-0.03;0.14]	0.05 [-0.01;0.10]	0.01 [-0.06;0.07]
Friday	-0.06 [-0.14;0.03]	-0.00 [-0.06;0.05]	-0.03 [-0.09;0.03]
Thursday	-0.12 [-0.21;-0.04]	0.00 [-0.05;0.05]	-0.05 [-0.11;0.02]
Wednesday	-0.11 [-0.19;-0.02]	-0.01 [-0.06;0.05]	-0.07 [-0.13;-0.01]
Tuesday	-0.08 [-0.16;0.01]	-0.03 [-0.08;0.03]	-0.05 [-0.11;0.02]
Day 1	0.18 [-0.06;0.43]	0.03 [-0.13;0.18]	-0.02 [-0.20;0.16]
Day 2	0.22 [-0.02;0.47]	0.08 [-0.08;0.23]	0.08 [-0.11;0.27]
Day 3	0.20 [-0.05;0.46]	0.13 [-0.04;0.29]	0.03 [-0.16;0.21]
Day 4	0.14 [-0.12;0.40]	0.09 [-0.08;0.26]	0.04 [-0.15;0.23]
Day 5	0.22 [-0.04;0.49]	0.05 [-0.12;0.22]	0.12 [-0.08;0.32]
Day 6	0.18 [-0.10;0.45]	0.10 [-0.07;0.28]	0.01 [-0.20;0.21]
Day 7+	-0.02 [-0.24;0.19]	0.13 [-0.01;0.26]	0.05 [-0.11;0.22]
sd (Intercept)	0.70	0.43	0.52
sd Observation	0.97	0.60	0.71

Note:
Estimates from the multilevel regression model, with 95% confidence intervals.

5 Dropout

The additional workload on each day (total number of randomised items that were shown) is a random variable, conditional on relationship status (women in relationships had to answer slightly more questions, some of which were also randomised in appearance). We can use this variable to predict whether a woman who had to answer more questions is less likely to answer the diary on the next day.
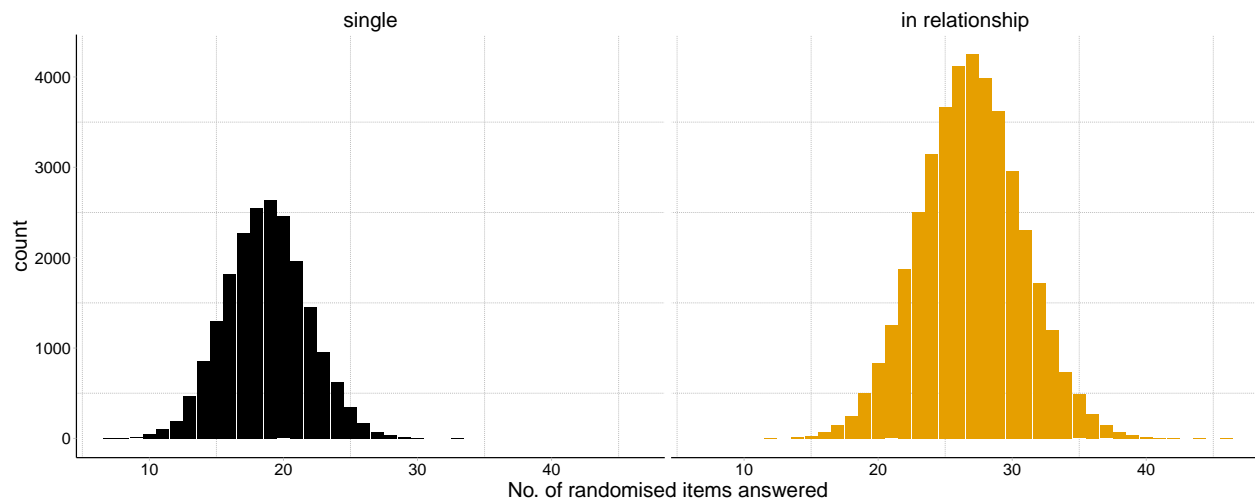

Figure 25: Distribution of the number of randomised items assigned on each day, split by relationship status.
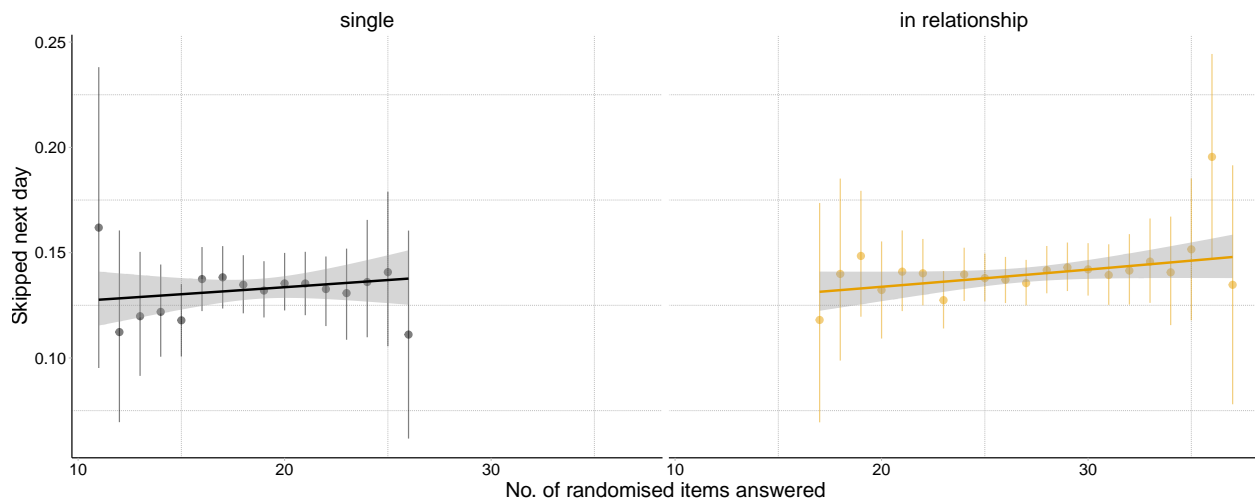

Figure 26: Is there any relationship between the number of items today and whether the participant takes part again on the next day? Only randomised variable counts with at least 100 instances are shown. The line and shading reflects a fit from a logistic regression that does not account for the multilevel structure of the data.

Table 18: Model comparison

term	df	AIC	logLik	statistic	Chi.Df	p.value
m0	3	43043	-21518	NA	NA	NA
m1	7	43046	-21516	5	4	0.29

Note:

Can we predict whether participants will skip the next day any better through the randomised number of items they answered today? Compares a multilevel model adjusted only for being in a relationship with one that allows for an effect of the no. of randomised items, which is allowed to vary between women and moderated by relationship status.