WLR intervention analyses

Matteo Lisi

Contents

1 Participants

Participants were tested at three time points: T1 (in February before the WLR reading program), T2 (in April, after the program), and T3 (later follow up in June).

Table 1: Composition of the dataset (number of subjects per timepoint)

Group	Nationality	refugee status	T1	T2	T3
control	Jordanian	0	20	20	9
WLR	Jordanian	0	25	22	16
control	Syrian	1	22	12	9
WLR	Syrian	1	26	15	12

Table 2: Mean age per group and timepoint

Group	Nationality	Timepoint	Mean age	Std.
control	Jordanian	T1	8.6	0.9
WLR	Jordanian	T1	8.4	0.9
control	Syrian	T1	9.9	1.5
WLR	Syrian	T1	8.7	1.4
control	Jordanian	T2	8.6	0.9
WLR	Jordanian	T2	8.6	0.5
control	Syrian	T2	10.1	1.2
WLR	Syrian	T2	8.5	1.2
control	Jordanian	T3	9.0	1.0
WLR	Jordanian	T3	8.6	0.5
control	Syrian	T3	10.2	1.2
WLR	Syrian	T3	8.5	1.2

2 Modelling

To analyse the data, we used a multilevel Bayesian generalized-linear model (GLM). The model is fully specified by the following expressions (the first 2 define the likelihood, the rest are the priors and hyperpriors). Note that differnetly from a classical GLM it accounts for the possibility of lapses, or random responses (parameter λ).

$$p(\text{responding 'sad'}) = \frac{\lambda}{2} + (1-\lambda)\,\Phi(\varphi)$$

$$\begin{aligned}
\varphi = {} & \beta_0 + u_0 + \alpha_0\,\text{gender} + (\beta_1 + u_1 + \alpha_1\,\text{gender})\ \text{img} + \beta_2\,\text{rfg} + \beta_3\,\text{group} \\
& + \text{T2}\ (\beta_4 + \beta_5\,\text{rfg} + \beta_6\,\text{group} + \beta_7\,\text{group} \times \text{rfg}) \\
& + \text{T3}\ (\beta_8 + \beta_9\,\text{rfg} + \beta_{10}\,\text{group} + \beta_{11}\,\text{group} \times \text{rfg}) \\
& + \text{img T2}\ (\beta_{12} + \beta_{13}\,\text{rfg} + \beta_{14}\,\text{group} + \beta_{15}\,\text{group} \times \text{rfg}) \\
& + \text{img T3}\ (\beta_{16} + \beta_{17}\,\text{rfg} + \beta_{18}\,\text{group} + \beta_{19}\,\text{group} \times \text{rfg})
\end{aligned}$$

$$\begin{aligned}
\beta_{0,2,\ldots,19} &\sim \mathcal{N}(0,1) \\
\beta_1 &\sim \mathcal{N}(2,5) \\
\lambda &\sim \text{Beta}(0.5,15) \\
\mathbf{u} &\sim \mathcal{N}(0,\Sigma) \\
\Sigma &= \left(\sigma_u^2 I\right)\ \mathbf{R}\ \left(\sigma_u^2 I\right) \\
\sigma_u &\sim \text{HalfCauchy}(0,1) \\
\mathbf{R} &\sim \text{LKJcorr}(2)
\end{aligned}$$

where 'img' indicate the level of morphing (centered and rescaled to have standard deviation of 1); 'rfg', 'group' are dummy variables that indicate whether the subject was a refugee (Syrian or Jordanian) or a member of the experimental group (WLR), respectively; and 'T1' and 'T2' are also dummy variables that indicate whether the measurements was taken at T2 or T3 (the first timepoint, T1, is the baseline). 'gender' is the gender of the actor whose face was morphed to generate the stimuli, coded using deviation (or sum) contrasts (gender = 1 for female faces, and gender = −1 for male faces). λ is a lapse rate parameter, which is assumed to be constant for all subjects (since it was the experimenters, not the children, who pushed buttons for all group and at all time points). Note that the model allow for possible differences between groups and nationality (or refugee status), and also for differences in changes across timepoints.

Finally, note also that although the model is formulated as generalized linear multilevel model (GLMM), the coefficients can be directly mapped to those of the psychometric function. To see this, consider the following simplified example:

$$\begin{aligned}
p(y_i) &= \Phi\left(\frac{x_i - \mu}{\sigma}\right) \\
&= \Phi(-\frac{\mu}{\sigma} + \frac{1}{\sigma}x_i) \\
&= \Phi(\beta_0 + \beta_1 x_i)
\end{aligned}$$

$$\mu = -\frac{\beta_0}{\beta_1},\ \ \sigma = \frac{1}{\beta_1}$$

3 Happy-sad psychophysical task

The following results concern the happy-sad bias task.

The PSE, estimates via the model, are plotted in Figure 3 as a function of the time point and group. Error bands are 95% Bayesian credible intervals (specifically HPDI, highest posterior density intervals).

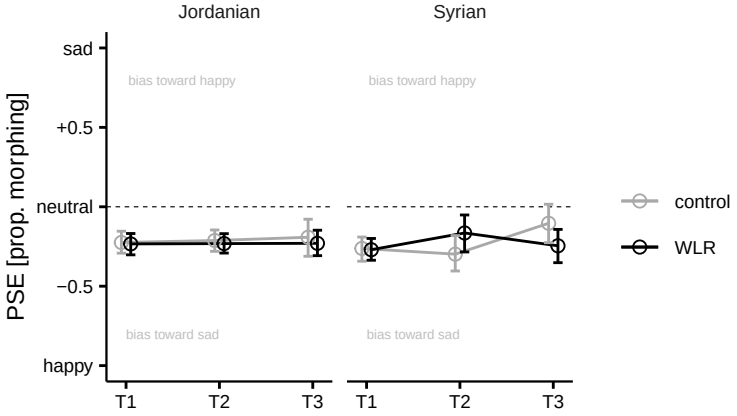

Figure 1: Group level estimate of PSE, plot by group and timepoint.

3.1 Adjusting for differences in actor identity (Syrian control, T3)

Due to a mix-up occurred while collecting data on the field, Syrian control group at T3 was not presented with the same actors as the others. This create a possible confound for assessing the effect of the intervention at T3 in the Syrian group.

However, it is possible to adjust for this confound, by including in the analysees some additiona data that we collected. Specifically, we have a fairly large sample of Syrian kids that have been tested with both set of faces. The approach I have adopted to adjust the analysis is the following:

- Estimate the difference in the parameters of the psychometric function using these additional dataset, which consist of the children tested at T1, and the new children tested at T3 (see Table 4). The output is a posterior distribution that gives the probability of over possible changes in the parameters due to the different actor identity.
- Next, the posterior distribution is used as a prior in the full model (using a Gaussian approximation, see below) to constrain additional parameters, added to account for the effect of actor identity.

Overall, the outcome will be identical to the previous model, except for the estiamted parameters of the Syrian control at T3, which will be adjusted to account for the effect of actor. Importantly, using this approach uncertainty is propagated exactly, such that the confidence interval around the estimated parameters for the Syrian T3 group will be larger (reflecing the added uncertainty about the true effect of actor identity).

Table 3: Composition of the supplementary dataset used to estimate the effect of actor identity.

Actor identity	Actor gender	Nationality	n. subjects
1	F	Syrian	26
2	F	Syrian	11
1	M	Syrian	23

Actor identity	Actor gender	Nationality	n. subjects
2	M	Syrian	11

The model is a simplified version of the full model described above, with the difference that in this case there are no parameter to code for group effects, but only for the effect of actor gender and identity. The simplified model is fully specified by the following expressions:

$$p(\text{responding 'sad'}) = \frac{\lambda}{2} + (1 - \lambda)\,\Phi(\varphi)$$

$$\begin{aligned}\varphi = {} & \beta_0 + u_0 + \beta_2\,\text{gender} + \beta_3\,\text{actor} + \beta_4\,\text{gender} \times \text{actor} \\ & + \text{img}\,(\beta_1 + u_1 + \beta_5\,\text{gender} + \beta_6\,\text{actor} + \beta_7\,\text{gender} \times \text{actor})\end{aligned}$$

$$\begin{aligned}\beta_{0,2,\ldots,7} &\sim \mathcal{N}(0, 1) \\ \beta_1 &\sim \mathcal{N}(2, 5) \\ \lambda &\sim \text{Beta}(0.5, 15) \\ \mathbf{u} &\sim \mathcal{N}(0, \Sigma) \\ \Sigma &= \left(\sigma_u^2 I\right)\ \mathbf{R}\ \left(\sigma_u^2 I\right) \\ \sigma_u &\sim \text{HalfCauchy}(0, 1) \\ \mathbf{R} &\sim \text{LKJcorr}(2)\end{aligned}$$

The posterior distribution obtained for the parameters of interest (i.e., $\beta_3, \beta_4, \beta_6, \beta_7$) are shown in the next figure. The histogram represents the samples of the posterior distribution obtained via Markov-Chain Montecarlo (MCMC) sampling. The red curves are Gaussian functions, fit on the posterior samples, which will be used to

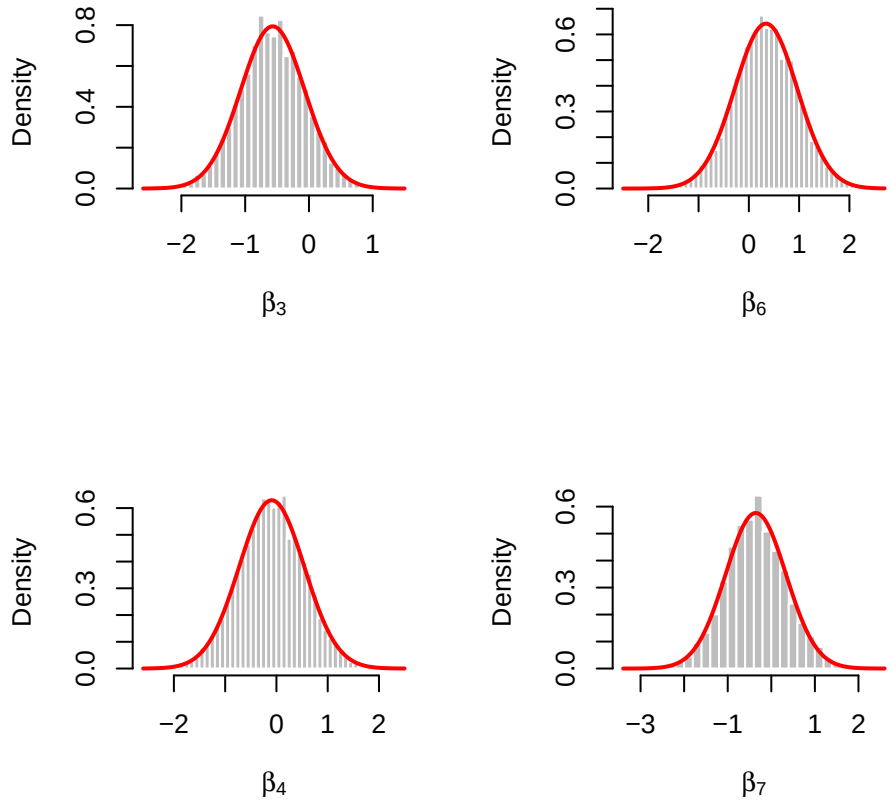

Figure 2: The grey histogram shows the posterior distributions over the effect of actor identity on bias and psychometric slopes. Red lines are the Gaussian approximation that was used the as informative prior in the main analysis.

Table 4: Summary statistics for posterior distributions of the effect of actor identity.

Parameter	name.in.updated.model	Mean	Std
beta_3	beta_20	-0.569	0.502
beta_4	beta_21	-0.092	0.634
beta_6	beta_22	0.341	0.621
beta_7	beta_23	-0.356	0.692

The posterior distribution, summarized in the table, are then used as prior in the updated model used to analyze the effect of the intervention, which becomes

$$p(\text{responding 'sad'}) = \frac{\lambda}{2} + (1-\lambda)\,\Phi(\varphi)$$

$$\begin{aligned}
\varphi = {} & \beta_0 + u_0 + \alpha_0\,\text{gender} + (\beta_1 + u_1 + \alpha_1\,\text{gender})\ \text{img} + \beta_2\,\text{rfg} + \beta_3\,\text{group} \\
& + \text{T2}\ (\beta_4 + \beta_5\,\text{rfg} + \beta_6\,\text{group} + \beta_7\,\text{group} \times \text{rfg}) \\
& + \text{T3}\ (\beta_8 + \beta_9\,\text{rfg} + \beta_{10}\,\text{group} + \beta_{11}\,\text{group} \times \text{rfg}) \\
& + \text{img T2}\ (\beta_{12} + \beta_{13}\,\text{rfg} + \beta_{14}\,\text{group} + \beta_{15}\,\text{group} \times \text{rfg}) \\
& + \text{img T3}\ (\beta_{16} + \beta_{17}\,\text{rfg} + \beta_{18}\,\text{group} + \beta_{19}\,\text{group} \times \text{rfg}) \\
& + \beta_{20}\,\text{actor} + \beta_{21}\,\text{gender} \times \text{actor} \\
& + \text{img}\,(\beta_{22}\,\text{actor} + \beta_{23}\,\text{gender} \times \text{actor})
\end{aligned}$$

where I have added to the bottom the new part (last two lines), and omitted the prior description for brevity. The prior description is the same as before, with the exception that the new parameters $\beta_{20}, \ldots, \beta_{23}$ have priors that are normal distribution with mean and standard deviation as indicated in the table (i.e. they correspond to the red curves). The results of this new analysis are represented in Figure 6. Notice that the main differences are that while previously the Syrian control seemed to exhibit an decrease in bias at T3, this effect can be entirely explained by the change in stimulus. The confidence bands is increased to reflect the increased uncertainty, however it seems safe to conclude that these results indicate that the none of the control groups, Syrian or Jordanian, exhibited significant changes in bias across timepoints.

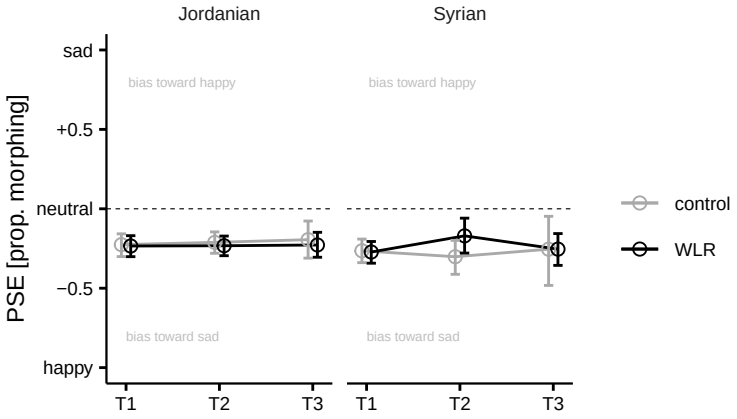

Figure 3: Group level estimate of PSE, plot by group and timepoint. Estimated using all data, and with adjustment for the bias of Syrian control at T3, due to the different stimulus (see text).

This can be seen even more clearly in Figure 7, where the same data is plotted as differences from T1

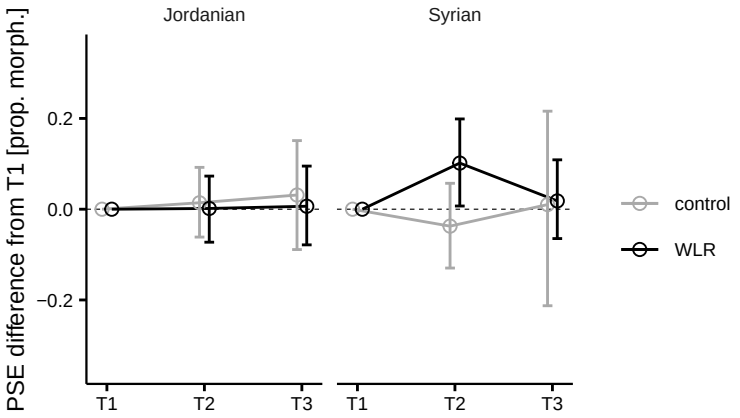

Figure 4: Group level estimate of PSE, plotted as a difference from T1.

Table 5: Mean and 95% credible intervals of differences from T1 (happy-sad task).

time	ethni	group	mu	mu_se	mu_lb	mu_ub
T2	Syrian	control	-0.037	0.048	-0.130	0.057
T2	Syrian	WLR	0.102	0.050	0.007	0.199
T2	Jordanian	control	0.014	0.039	-0.061	0.092
T2	Jordanian	WLR	0.002	0.037	-0.073	0.073
T3	Syrian	control	0.010	0.111	-0.213	0.216
T3	Syrian	WLR	0.018	0.044	-0.065	0.109
T3	Jordanian	control	0.031	0.061	-0.089	0.151
T3	Jordanian	WLR	0.006	0.044	-0.079	0.095

4 Fear-anger psychophysical task

Analyses of the fear-anger task. The analysis procedure is exactly the same.

The only difference is that in this case I have two Jordanian control that were tested using actor 2 at T3, therefore in the additional model used to estimate the effect of the actor identity I have included both Syrian and Jordanian (see table). The results indicates no effect of the treatment, since the confidence intervals are always largely overlapping. There is a significant change in bias at T2 but only for the Jordanians, and for both the control and experimental group.

Table 6: Composition of the supplementary dataset used to estimate the effect of actor identity.

Actor identity	Actor gender	Nationality	n. subjects
1	F	Jordanian	23
2	F	Jordanian	8
1	M	Jordanian	23
2	M	Jordanian	8
1	F	Syrian	23
2	F	Syrian	11
1	M	Syrian	26
2	M	Syrian	11

Figure 10 and 11 represent the PSE in the FA task, using the same coventions adopted in the previous figures.

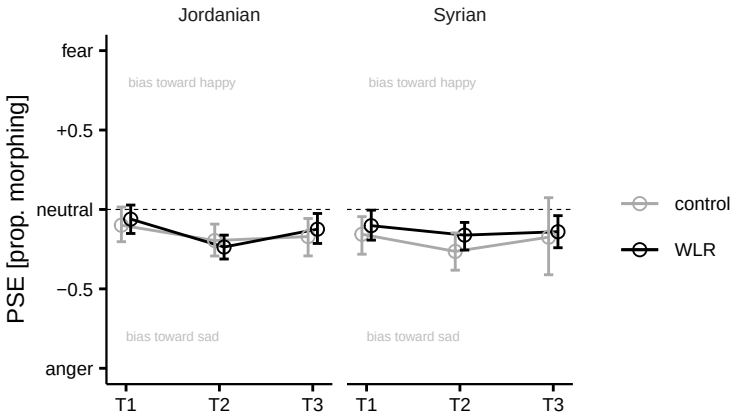

Figure 5: Group level estimate of PSE in the FA task, plot by group and timepoint. Estimated using all data, and with adjustment for the bias of Syrian control at T3, due to the different stimulus (see text).

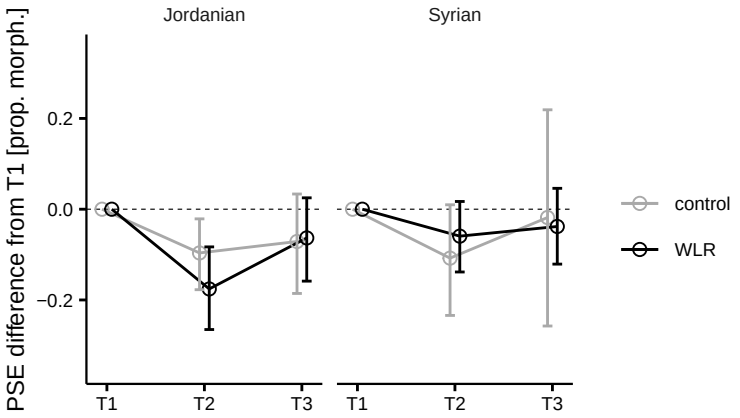

Figure 6: Group level estimate of PSE, plotted as a difference from T1.

Table 7: Mean and 95% credible intervals of differences from T1 (fear-anger task).

time	ethni	group	mu	mu_se	mu_lb	mu_ub
T2	Syrian	control	-0.108	0.062	-0.234	0.010
T2	Syrian	WLR	-0.059	0.040	-0.138	0.017
T2	Jordanian	control	-0.096	0.040	-0.177	-0.021
T2	Jordanian	WLR	-0.176	0.047	-0.265	-0.083
T3	Syrian	control	-0.018	0.123	-0.257	0.219
T3	Syrian	WLR	-0.038	0.042	-0.121	0.046
T3	Jordanian	control	-0.071	0.056	-0.186	0.033
T3	Jordanian	WLR	-0.063	0.047	-0.158	0.025

5 Questionnaires analysis

5.1 Whole-sample comparison at T1

The different questionnaire measures obtained at T1 where compared across groups (Syrian vs. Jordanian) using the Mann–Whitney U test, with a normal approximation to compute the p-values. See table for the results

Table 8: Comparison of questionnaire responses at T1 (Syrian vs Jordanian). TEC = Traumatic Events

questionnaire	W	p
OPTIMISM_1	1314.0	0.0904962
PPL_IN_HOUSEHOLD	996.5	0.4157287
CRIES_8	811.0	0.1240409
CRIES_INTRUSION	831.5	0.0893334
CRIES_AVOIDANCE	909.5	0.2126538
AYMHS_1	1000.5	0.9258415
HIDS_INSECURITY_1	1153.5	0.2537263
HIDS_DISTRESS_1	1036.0	0.9904620
TEC	82.0	0.0000000
CRIES_ADJUSTED	615.0	0.0000187
CRIES_SCORE	710.0	0.0002571

Longitudinal tests for changes in the questionnaire scores at the different time points were done using the non-parametric marginal model (Brunner and Puri 2001,@Brunner1999), a generalization of classical non-parametric tests to repeated-measures, factorial designs. For each test the statistical significance is assessed using the ANOVA-type statistic (ATS). The ATS is typically preferred to more classical χ^2 commonly used by rank-score tests as it requires less assumptions on the covariance matrix and it has superior small sample performances (Brunner and Puri 2001). Moreover, for the main effects and interactions involving only the between subjects effects, I report the modified ANOVA-type statistic with Box approximation, which is preferred to ATS (Brunner and Puri 2001).

Notice that the longitudinal tests are restricted to the subset of participants for which a measure is available at all timepoints.

5.2 Longitudinal analysis (3 time points)

The following plots contain data from all participants (regardless of whether they showed up at T2 or T3).

5.2.1 Optimism

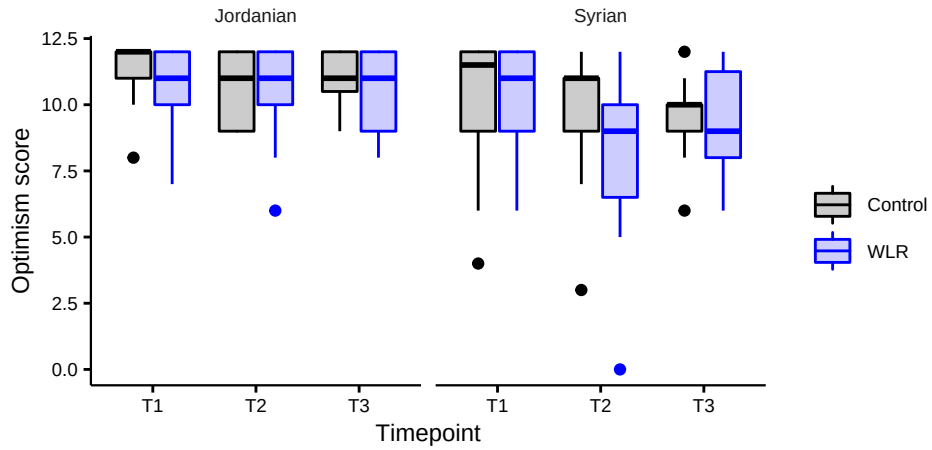

Figure 7: Optimism scores across timepoints

Table 9: Optimism: ANOVA-Type Statistc (ATS).

	Statistic	df	p-value
wlr_group	3.5923634	1.000000	0.0580456
nationality	4.9052342	1.000000	0.0267754
time	3.4456288	1.837244	0.0357247
wlr_group:nationality	1.0709478	1.000000	0.3007315
wlr_group:time	1.3613788	1.837244	0.2560062
nationality:time	0.8612858	1.837244	0.4144877
wlr_group:nationality:time	1.4252773	1.837244	0.2408272

Table 10: Optimism: Modified ANOVA-Type Statistic for the between-subjects factors.

	Statistic	df1	df2	p-value
wlr_group	3.592363	1	25.94952	0.0692370
nationality	4.905234	1	25.94952	0.0357630
wlr_group:nationality	1.070948	1	25.94952	0.3102799

5.2.2 Human Insecurity Scale (HIS)

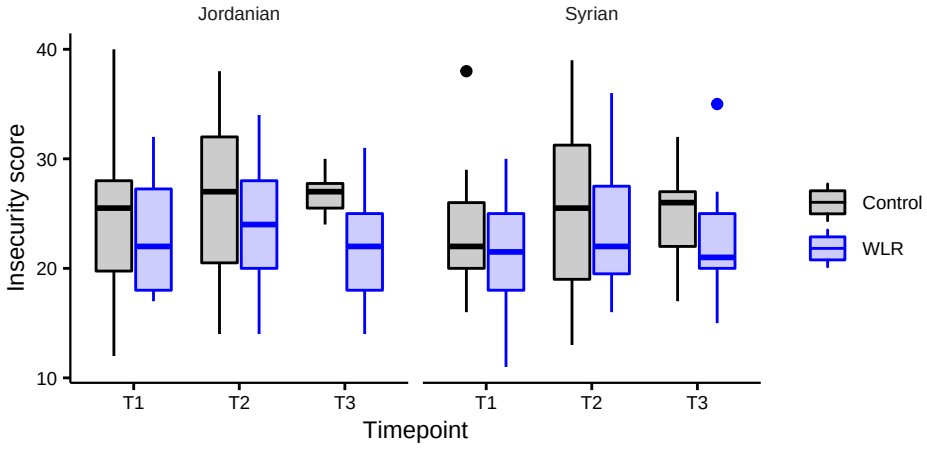

Figure 8: Insecurity scores across timepoints

Table 11: Insecurity: ANOVA-Type Statistc (ATS).

	Statistic	df	p-value
wlr_group	2.2265276	1.000000	0.1356584
nationality	2.9460533	1.000000	0.0860876
time	1.8402814	1.589945	0.1671087
wlr_group:nationality	0.0001575	1.000000	0.9899877
wlr_group:time	1.3898812	1.589945	0.2482679
nationality:time	0.3437137	1.589945	0.6592568
wlr_group:nationality:time	1.0947785	1.589945	0.3234342

Table 12: Insecurity: Modified ANOVA-Type Statistic for the between-subjects factors.

	Statistic	df1	df2	p-value
wlr_group	2.2265276	1	11.15105	0.1633990
nationality	2.9460533	1	11.15105	0.1137017
wlr_group:nationality	0.0001575	1	11.15105	0.9902094

5.2.3 Human Distress Scale (HDS)

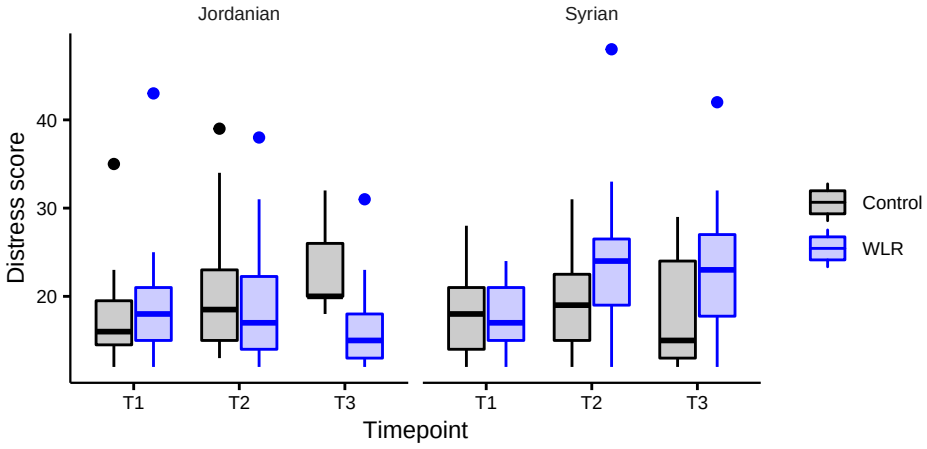

Figure 9: Distress scores across timepoints

Table 13: Distress: ANOVA-Type Statistc (ATS).

	Statistic	df	p-value
wlr_group	0.1665382	1.000000	0.6832069
nationality	0.0241826	1.000000	0.8764211
time	4.1466123	1.593206	0.0235174
wlr_group:nationality	3.9598536	1.000000	0.0465978
wlr_group:time	0.9117982	1.593206	0.3823730
nationality:time	0.2685269	1.593206	0.7132980
wlr_group:nationality:time	6.1912807	1.593206	0.0043112

Table 14: Distress: Modified ANOVA-Type Statistic for the between-subjects factors.

	Statistic	df1	df2	p-value
wlr_group	0.1665382	1	29.29467	0.6861744
nationality	0.0241826	1	29.29467	0.8774877
wlr_group:nationality	3.9598536	1	29.29467	0.0560033

5.2.4 Arab Youth Mental Health Scale (AYMHS)

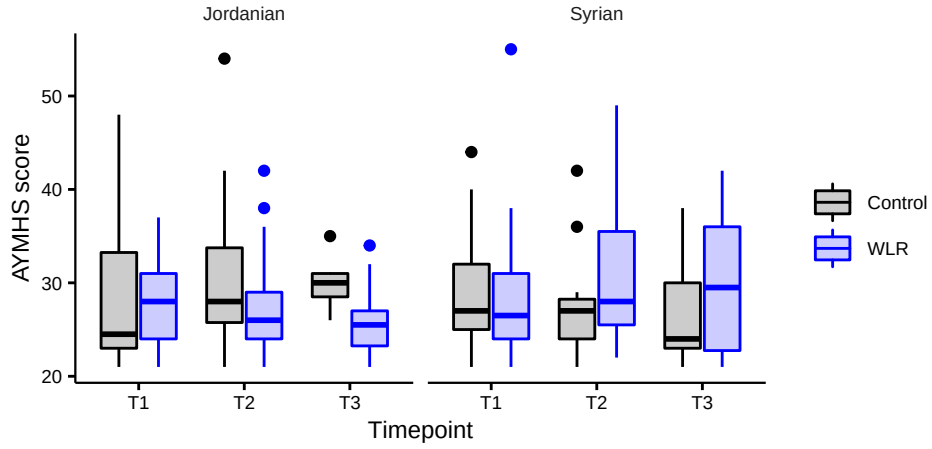

Figure 10: Mental health (AYMHS) scores across timepoints

Table 15: Mental health (AYMHS): ANOVA-Type Statistc (ATS).

	Statistic	df	p-value
wlr_group	0.0019947	1.00000	0.9643769
nationality	0.6386774	1.00000	0.4241901
time	6.2171464	1.58446	0.0042890
wlr_group:nationality	1.7724457	1.00000	0.1830797
wlr_group:time	0.7615615	1.58446	0.4389209
nationality:time	1.4786086	1.58446	0.2294960
wlr_group:nationality:time	4.0916958	1.58446	0.0248288

Table 16: Mental health (AYMHS): Modified ANOVA-Type Statistic for the between-subjects factors.

	Statistic	df1	df2	p-value
wlr_group	0.0019947	1	28.12391	0.9646925
nationality	0.6386774	1	28.12391	0.4308882
wlr_group:nationality	1.7724457	1	28.12391	0.1937713

5.3 Longitudinal analysis (2 time points)

The following plots contain data from only participants that were re-tested at T2.

5.3.1 Optimism

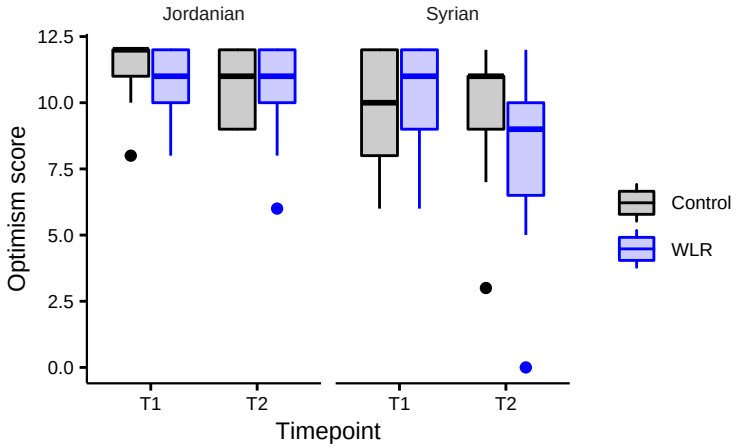

Figure 11: Optimism scores across 2 timepoints

Table 17: Optimism: ANOVA-Type Statistc (ATS). 2 time points only.

	Statistic	df	p-value
wlr_group	2.0382184	1	0.1533892
nationality	4.7440722	1	0.0293994
time	9.1556824	1	0.0024795
wlr_group:nationality	0.2938879	1	0.5877398
wlr_group:time	0.1886453	1	0.6640465
nationality:time	0.0394533	1	0.8425534
wlr_group:nationality:time	4.0472755	1	0.0442427

Table 18: Optimism: Modified ANOVA-Type Statistic for the between-subjects factors. 2 time points only.

	Statistic	df1	df2	p-value
wlr_group	2.0382184	1	42.05632	0.1607730
nationality	4.7440722	1	42.05632	0.0350566
wlr_group:nationality	0.2938879	1	42.05632	0.5906008

5.3.2 Human Insecurity Scale (HIS)

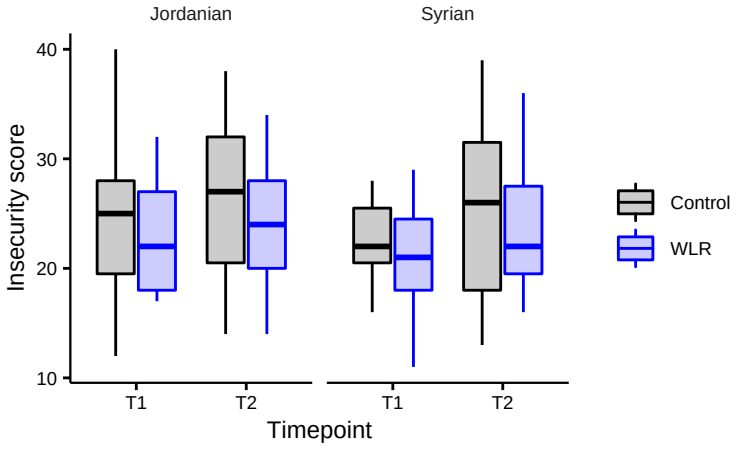

Figure 12: Insecurity scores across 2 timepoints

Table 19: Insecurity: ANOVA-Type Statistc (ATS). 2 time points only.

	Statistic	df	p-value
wlr_group	1.2846914	1	0.2570286
nationality	0.6141009	1	0.4332479
time	3.8104314	1	0.0509343
wlr_group:nationality	0.0012357	1	0.9719579
wlr_group:time	0.0484404	1	0.8257997
nationality:time	0.4908202	1	0.4835617
wlr_group:nationality:time	0.3493180	1	0.5544994

Table 20: Insecurity: Modified ANOVA-Type Statistic for the between-subjects factors. 2 time points only.

	Statistic	df1	df2	p-value
wlr_group	1.2846914	1	45.37821	0.2629815
nationality	0.6141009	1	45.37821	0.4373192
wlr_group:nationality	0.0012357	1	45.37821	0.9721121

5.3.3 Human Distress Scale (HDS)

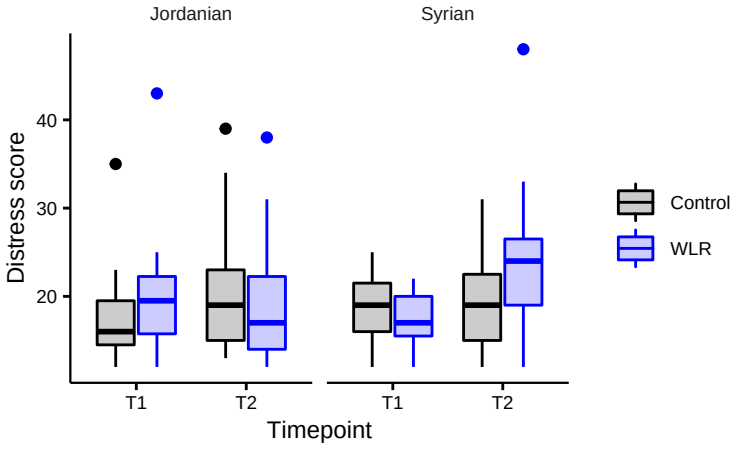

Figure 13: Distress scores across 2 timepoints

Table 21: Distress: ANOVA-Type Statistc (ATS). 2 time points only.

	Statistic	df	p-value
wlr_group	0.2732606	1	0.6011528
nationality	0.2901686	1	0.5901125
time	3.4912129	1	0.0616954
wlr_group:nationality	0.0003559	1	0.9849496
wlr_group:time	0.0892181	1	0.7651736
nationality:time	0.7210084	1	0.3958133
wlr_group:nationality:time	6.6378831	1	0.0099832

Table 22: Distress: Modified ANOVA-Type Statistic for the between-subjects factors. 2 time points only.

	Statistic	df1	df2	p-value
wlr_group	0.2732606	1	43.12443	0.6038283
nationality	0.2901686	1	43.12443	0.5928825
wlr_group:nationality	0.0003559	1	43.12443	0.9850366

5.3.4 Arab Youth Mental Health Scale (AYMHS)

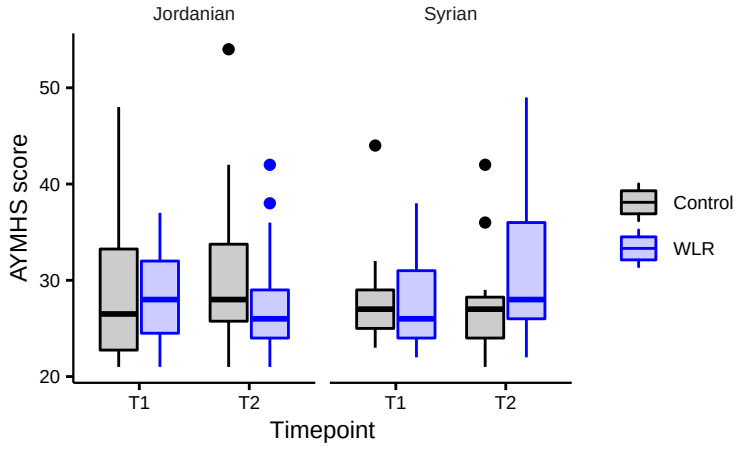

Figure 14: Mental health (AYMHS) scores across timepoints

Table 23: Mental health (AYMHS): ANOVA-Type Statistc (ATS). 2 time points only.

	Statistic	df	p-value
wlr_group	0.0008576	1	0.9766380
nationality	0.0024131	1	0.9608208
time	2.2759074	1	0.1313982
wlr_group:nationality	0.6510359	1	0.4197426
wlr_group:time	0.1778623	1	0.6732168
nationality:time	0.0320515	1	0.8579147
wlr_group:nationality:time	7.4279115	1	0.0064220

Table 24: Mental health (AYMHS): Modified ANOVA-Type Statistic for the between-subjects factors. 2 time points only.

	Statistic	df1	df2	p-value
wlr_group	0.0008576	1	49.63691	0.9767554
nationality	0.0024131	1	49.63691	0.9610179
wlr_group:nationality	0.6510359	1	49.63691	0.4235920

5.3.5 Post-hoc tests

Some of the analyses revealed significant changes in the scores that could be attributed to the WLR reading program. Here I use the same non parametric approach to do post-hoc comparisons. There are three tables, one for optimism, one for distress and one for mental health. Control of family-wise error rate is done by applying the Bonferroni correction: the adjusted p-value is the p-value multiplied by the number of comparisons, 4, correspondign to alpha level of 0.0125.[1]

Table 25: Post-hoc comparisons for Optimism scale. The table also show the adjusted p value (Bonferroni corrected).

Statistic	df	p-value	p_adjusted	Comparison
5.9273992	1	0.0149072	0.0596289	Syrian, WLR: T2 vs T1
0.1564633	1	0.6924338	1.0000000	Syrian, control: T2 vs T1
0.3577598	1	0.5497531	1.0000000	Jordanian, WLR: T2 vs T1
7.8400000	1	0.0051103	0.0204410	Jordanian, control: T2 vs T1

Table 26: Post-hoc comparisons for distress scale (HDS). The table also show the adjusted p value (Bonferroni corrected).

Statistic	df	p-value	p_adjusted	Comparison
7.1439114	1	0.0075219	0.0300876	Syrian, WLR: T2 vs T1
0.0000000	1	NA	NA	Syrian, control: T2 vs T1
0.5597584	1	0.4543576	1.0000000	Jordanian, WLR: T2 vs T1
2.6041406	1	0.1065849	0.4263397	Jordanian, control: T2 vs T1

Table 27: Post-hoc comparisons for meantal health scale (AYMHS). The table also show the adjusted p value (Bonferroni corrected).

Statistic	df	p-value	p_adjusted	Comparison
10.3541699	1	0.0012918	0.0051673	Syrian, WLR: T2 vs T1
0.5678776	1	0.4511035	1.0000000	Syrian, control: T2 vs T1
0.2471749	1	0.6190714	1.0000000	Jordanian, WLR: T2 vs T1
2.1470646	1	0.1428428	0.5713710	Jordanian, control: T2 vs T1

[1] A more restrictive control perhaps would be to consider this a family of 3 times 4 = 12 comparisons, corresponding to an alpha level of 0.00416, in which case the only remaining significant comparison would be Syrian, WLR: T2 vs T1, with a p-value of =0.0155.

Same tests as in previous page, but using Wilcoxon signed rank test.

Table 28: Post-hoc comparisons for Optimism scale (Wilcoxon signed rank tests). The table also show the adjusted p value (Bonferroni corrected).

statistic	p.value	p_adjusted	Comparison	p.adjusted.BH
66.0	0.0360975	0.1443899	Syrian, WLR: T2 vs T1	0.0721949
23.5	0.9512007	1.0000000	Syrian, control: T2 vs T1	0.9512007
92.0	0.4701008	1.0000000	Jordanian, WLR: T2 vs T1	0.6268010
50.0	0.0232606	0.0930425	Jordanian, control: T2 vs T1	0.0721949

Table 29: Post-hoc comparisons for distress scale (HDS), Wilcoxon signed rank tests. The table also show the adjusted p value (Bonferroni corrected).

statistic	p.value	p_adjusted	Comparison	p.adjusted.BH
19.5	0.0229575	0.0918300	Syrian, WLR: T2 vs T1	0.0918300
24.0	0.7593110	1.0000000	Syrian, control: T2 vs T1	0.7593110
87.0	0.6346577	1.0000000	Jordanian, WLR: T2 vs T1	0.7593110
46.5	0.0929422	0.3717688	Jordanian, control: T2 vs T1	0.1858844

Table 30: Post-hoc comparisons for meantal health scale (AYMHS); Wilcoxon signed rank tests. The table also show the adjusted p value (Bonferroni corrected).

statistic	p.value	p_adjusted	Comparison	p.adjusted.BH
6.5	0.0206066	0.0824263	Syrian, WLR: T2 vs T1	0.0824263
41.5	0.4749361	1.0000000	Syrian, control: T2 vs T1	0.6332481
89.0	0.8957066	1.0000000	Jordanian, WLR: T2 vs T1	0.8957066
18.5	0.0634621	0.2538486	Jordanian, control: T2 vs T1	0.1269243

5.3.6 Post-hoc tests: T3 vs T1

Table 31: Post-hoc comparisons for Optimism scale. The table also show the adjusted p value (Bonferroni corrected).

Statistic	df	p-value	p_adjusted	Comparison
1.5519268	1	0.2128512	0.8514048	Syrian, WLR: T3 vs T1
0.0917431	1	0.7619727	1.0000000	Syrian, control: T3 vs T1
0.0054735	1	0.9410241	1.0000000	Jordanian, WLR: T3 vs T1
1.7553047	1	0.1852112	0.7408450	Jordanian, control: T3 vs T1

Table 32: Post-hoc comparisons for distress scale (HDS). The table also show the adjusted p value (Bonferroni corrected).

Statistic	df	p-value	p_adjusted	Comparison
7.9950150	1	0.0046906	0.0187625	Syrian, WLR: T3 vs T1
0.0232875	1	0.8787117	1.0000000	Syrian, control: T3 vs T1
2.9827353	1	0.0841569	0.3366278	Jordanian, WLR: T3 vs T1
5.5867902	1	0.0180964	0.0723857	Jordanian, control: T3 vs T1

Table 33: Post-hoc comparisons for meantal health scale (AYMHS). The table also show the adjusted p value (Bonferroni corrected).

Statistic	df	p-value	p_adjusted	Comparison
0.0034843	1	0.9529297	1.0000000	Syrian, WLR: T3 vs T1
1.3281453	1	0.2491354	0.9965418	Syrian, control: T3 vs T1
1.2463214	1	0.2642562	1.0000000	Jordanian, WLR: T3 vs T1
0.2463343	1	0.6196681	1.0000000	Jordanian, control: T3 vs T1

Same tests as in previous page, but using Wilcoxon signed rank test.

Table 34: Post-hoc comparisons for Optimism scale (Wilcoxon signed rank tests). The table also show the adjusted p value (Bonferroni corrected).

statistic	p.value	p_adjusted	Comparison	p.adjusted.BH
21.0	0.2675839	1.0000000	Syrian, WLR: T3 vs T1	0.5351678
23.0	0.5176341	1.0000000	Syrian, control: T3 vs T1	0.6901788
48.5	0.8594663	1.0000000	Jordanian, WLR: T3 vs T1	0.8594663
12.5	0.2030918	0.8123672	Jordanian, control: T3 vs T1	0.5351678

Table 35: Post-hoc comparisons for distress scale (HDS), Wilcoxon signed rank tests. The table also show the adjusted p value (Bonferroni corrected).

statistic	p.value	p_adjusted	Comparison	p.adjusted.BH
4.5	0.0217387	0.0869547	Syrian, WLR: T3 vs T1	0.0869547
22.5	1.0000000	1.0000000	Syrian, control: T3 vs T1	1.0000000
82.5	0.2093514	0.8374058	Jordanian, WLR: T3 vs T1	0.2791353
3.0	0.0738343	0.2953373	Jordanian, control: T3 vs T1	0.1476686

Table 36: Post-hoc comparisons for meantal health scale (AYMHS); Wilcoxon signed rank tests. The table also show the adjusted p value (Bonferroni corrected).

statistic	p.value	p_adjusted	Comparison	p.adjusted.BH
17.5	0.5933057	1.0000000	Syrian, WLR: T3 vs T1	0.7344021
26.5	0.2608580	1.0000000	Syrian, control: T3 vs T1	0.5217161
63.0	0.2322284	0.9289136	Jordanian, WLR: T3 vs T1	0.5217161
11.5	0.7344021	1.0000000	Jordanian, control: T3 vs T1	0.7344021

5.4 Correlations between questionnaire scores and bias in emotion task

The table reports both Kendall and Spearman rank-correlation coefficients between bias in happy-sad task and questionnaire score at T1, each with both the original p-value and the p-value adjusted after controlling for false discovery rate (*Benjamini–Hochberg procedure*).

Table 37: Non-parametric correlation coefficients between questionnaire scores and bias in HS task.

quest	kendall	kendall_p	kendall_p_adj	spearman	spearman_p	spearman_p_adj
OPTIMISM_1	0.2076	0.0075	0.0597	0.2853	0.0056	0.0446
TEC	-0.0971	0.2016	0.4033	-0.1248	0.2333	0.3921
AYMHS_1	0.0468	0.5245	0.5245	0.0710	0.5061	0.5061
CRIES_ADJUSTED	-0.1617	0.0461	0.1228	-0.2115	0.0441	0.1177
CRIES_INTRUSION	-0.0739	0.3471	0.4628	-0.0985	0.3558	0.4745
CRIES_AVOIDANCE	-0.0894	0.2544	0.4071	-0.1231	0.2451	0.3921
HIDS_INSECURITY_1	-0.0566	0.4446	0.5081	-0.0819	0.4454	0.5061
HIDS_DISTRESS_1	0.1779	0.0161	0.0643	0.2423	0.0214	0.0855

Table 38: Non-parametric correlation coefficients between questionnaire scores and bias in FA task.

quest	kendall	kendall_p	kendall_p_adj	spearman	spearman_p	spearman_p_adj
OPTIMISM_1	-0.0783	0.3129	0.8345	-0.1018	0.3313	0.8754
TEC	-0.1108	0.1449	0.5796	-0.1579	0.1306	0.5401
AYMHS_1	-0.0129	0.8612	0.8834	-0.0222	0.8352	0.8754
CRIES_ADJUSTED	-0.0388	0.6325	0.8834	-0.0496	0.6403	0.8754
CRIES_INTRUSION	0.0197	0.8024	0.8834	0.0318	0.7659	0.8754
CRIES_AVOIDANCE	0.0115	0.8834	0.8834	0.0167	0.8754	0.8754
HIDS_INSECURITY_1	-0.1113	0.1326	0.5796	-0.1597	0.1350	0.5401
HIDS_DISTRESS_1	0.0124	0.8666	0.8834	0.0216	0.8397	0.8754

References

Brunner, Edgar, Ulrich Munzel, and Madan L Puri. 1999. "Rank-Score Tests in Factorial Designs with Repeated Measures." *Journal of Multivariate Analysis* 70 (2): 286–317. https://doi.org/10.1006/jmva.1999.1821.

Brunner, Edgar, and Madan L. Puri. 2001. "Nonparametric methods in factorial designs." *Statistical Papers* 42 (1): 1–52. https://doi.org/10.1007/s003620000039.