BAYESIAN MIXED FACTOR ANOVA

Mixed factor ANOVA (another two-way ANOVA) is a combination of both independent and repeated measures ANOVA involving more than 1 independent variable (known as factors). Below is a design with time as the within and group as the between factor:

Independent variable (Factor 2)	*Independent variable (Factor 1) = time or condition*		
	Time/condition 1	**Time/condition 2**	**Time/condition 3**
Group 1	Dependent variable	Dependent variable	Dependent variable
Group 2	Dependent variable	Dependent variable	Dependent variable

The factors are split into levels, therefore, in this case, Factor 1 has 3 levels and Factor 2 has 2 levels. This results in 6 possible combinations.

A "main effect" is the effect of one of the independent variables on the dependent variable, ignoring the effects of any other independent variables. There are 2 main effects tested: in this case comparing data across factor 1 (i.e., time) is known as the "**repeated measures**" factor while comparing differences between factor 2 (i.e., groups) is known as the "**between-subjects**" factor. I**nteraction** is where one factor influences the other factor.

The standard frequentist approach to ANOVA is to compare the variances between levels of a defined factor where the H_0 is that these variances are equal.

The Bayesian ANOVA compares the predictive performances of the possible competing models, i.e., how likely a set of data is under one model compared to another. In most cases, one model is the null model (H_0) suggesting that the data is purely random and the alternative model (H_1) that one or more of the factors have an effect. In the mixed factor analysis, multiple models are tested.

ASSUMPTIONS

Like all other analyses, mixed factor ANOVA makes a series of assumptions which should either be addressed in the research design or can the tested for.

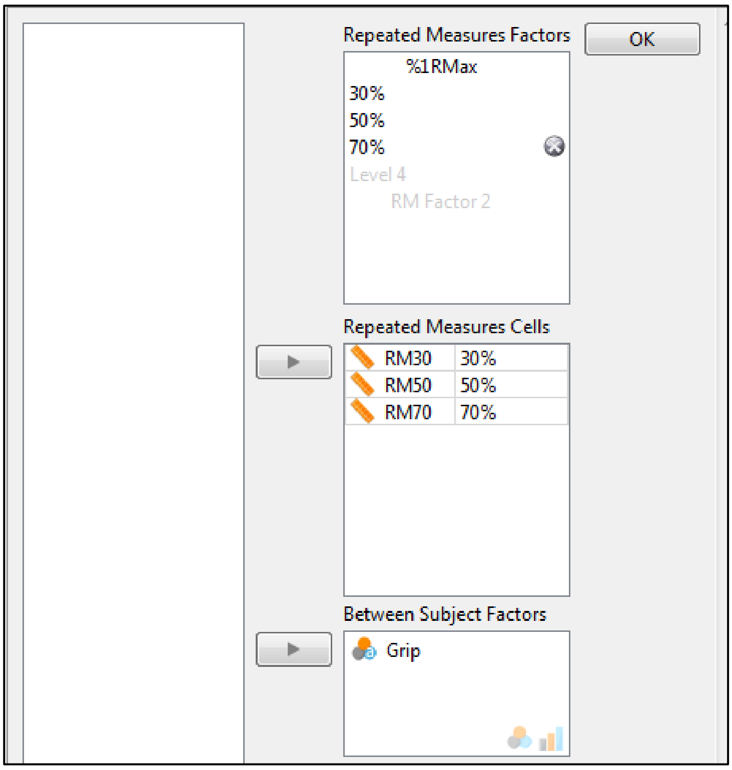
1. The "**Repeated measures**" factor should contain at least two related (repeated measures) categorical groups (levels).
2. The "**Between-subjects**" factor should have at least two categorical independent groups (levels).
3. The dependent variable should be continuous and approximately normally distributed for all combinations of factors.
4. There should be homogeneity of variance between the groups.
5. There should be no outliers.

RUNNING THE MIXED FACTOR BAYESIAN ANOVA

Open **Bayesian Mixed ANOVA.csv** in JASP. This contains 4 columns of data relating to the type of weightlifting grip and speed of the lift at 3 different loads (%1RM) for deadlifting. Column 1 contains the grip type, columns 2-4 contain the 3 repeated measures (30, 50 and70%). Check for outliers using boxplots.

NOTE: When running the analysis using the included dataset the results are always likely to be very slightly different to the ones in this chapter. This is because the analyses are based on numerical algorithms like Markov chain Monte Carlo (MCMC) which reports an error percentage. The higher the error percentage the higher the fluctuation of the results.

Go to ANOVA > Bayesian Repeated measures ANOVA. Define the Repeated Measures Factor, %1RMax, and add 3 levels (30, 50 and 70%). Add the appropriate variable to the Repeated measures Cells and add Grip to the Between-Subjects Factors:



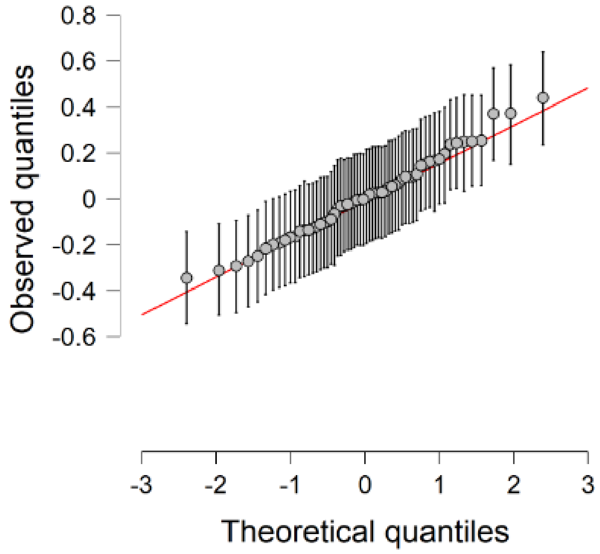
Then select the following options:

- ✓ Bayes Factor – BF_{10}
- ✓ Order – compare to best model
- ✓ Tables – Effects – Across all models
- ✓ Q-Q plots of residuals.
- ✓ Descriptives

In Descriptive plots move %1Rmax to the horizontal axis and Grip to 'Separate lines'
The output should initially comprise of 4 tables and 3 graphs.

UNDERSTANDING THE OUTPUT

Firstly, it is important to test the assumption of normality, in this case, that the residuals are normally distributed. This can easily be done by looking at the Q-Q plot.



If the residuals are normally distributed, they should lie consistently along the diagonal line. Any obvious deviations along the line would suggest that the assumption of normality has been violated. The assumption of homogeneity of variance can be assessed using Levene's test which is calculated as part of the classical ANOVA analysis.

Comparison of the competing models – Best model

The first column lists all models determined: four alternative models and one null model. The models

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
%1RMax + Grip + %1RMax ⁎ Grip	0.200	0.997	1265.774	1.000	
%1RMax + Grip	0.200	0.003	0.012	0.003	2.680
%1RMax	0.200	6.193e −5	2.477e −4	6.212e −5	1.838
Grip	0.200	4.660e −17	1.864e −16	4.675e −17	1.893
Null model (incl. subject)	0.200	2.057e −17	8.228e −17	2.064e −17	1.700

Note. All models include subject

are ordered by their predictive performance relative to the best model in this case.

In the other columns, results are presented for:

P(M): for the ANOVA, the analysis sets the prior probabilities of each of the five models to be equal (i.e., 0.2).

P(M │data): shows the updated probabilities having now seen the data.

BFM: shows how much the data have changed the prior model odds

BF_{10}: shows the Bayes comparison with the best model; for the first row, it is always 1 since it is being compared to itself.

H(1):%1RM + Grip + %1RM*Grip

A model based on the alternative hypothesis that lift speed depends on %1RM, grip type and the interaction between these two factors. This is the best model and has a BF_{10}=1 since it is being compared to itself.

H(1):%1RM + Grip

A model based on the alternative hypothesis that lift speed depends on %1RM and grip type. This a BF_{10} of 0.003 or a BF_{01} of 322, suggesting that the data are 322 times more likely under the best model than under the model with main effects only.

H(1): %1RM,   H(1): grip

Models based on the alternative hypothesis that lift speed depends on either %1RM or grip alone have extremely small BF_{10} values, as does the null model.

Comparison of the competing models – Null model

Alternatively, the data can be compared to the null model rather than the best model. In the options change the order to 'compare to the null model'. The model comparison has tested 5 models and compares the alternative models to the null model (H_0) which that states lift speed is not dependent on any other factors.

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Null model (incl. subject)	0.200	2.057e −17	8.228e −17	1.000	
%1RMax + Grip + %1RMax ✻ Grip	0.200	0.997	1265.774	4.846e +16	1.700
%1RMax + Grip	0.200	0.003	0.012	1.501e +14	2.071
%1RMax	0.200	6.193e −5	2.477e −4	3.011e +12	0.699
Grip	0.200	4.660e −17	1.864e −16	2.266	0.833

Note. All models include subject

H(1): grip

A model based on the alternative hypothesis that lift speed depends on grip type alone. This has a very small Bayes factor of 2.26 suggesting that there is very little evidence for this model, compared to the null model.

H(1): %1RM

A model based on the alternative hypothesis that lift speed depends on %1RM alone. This has an extremely large BF_{10} (i.e., $3.01*10^{12}$), decisively supporting this model over the null model.

H(1):%1RM + Grip

A model based on the alternative hypothesis that lift speed depends on %1RM and grip type. This also has an extremely large BF_{10} (i.e., $1.5*10^{14}$), decisively supporting this model over the null model.

H(1):%1RM + Grip + %1RM*Grip

A model based on the alternative hypothesis that lift speed depends on %1RM, grip type and the interaction between these two factors. This is the best model and has the largest BF_{10} (i.e., $4.86*10^{16}$), against the null model.

In order to compare the %1RM + Grip model against the %1RM + Grip + %1RM*Grip model, one can divide out the null hypothesis by computing $4.86*10^{16}$)/ $1.5*10^{14}$) = 324, which should give (approximately, due to rounding) the same result as the earlier 'compare to best model' analysis (i.e., BF = 322).

Whether one wants to compare to either the best or the null models is a matter of personal choice, the result is effectively the same.

Analysis of effects

This table shows the prior and posterior inclusion probability and the inclusion Bayes factor for each of the model's predictors. These data are based on all the models simultaneously.

%1Rmax and grip are considered as the main effects and the %1Rmax*Grip the interaction.

Analysis of Effects

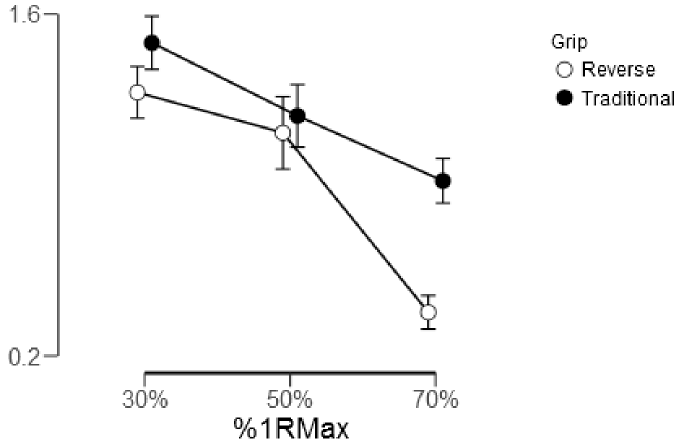
Effects	P(incl)	P(incl|data)	BF_{incl}
%1RMax	0.600	1.000	∞
Grip	0.600	1.000	10764.591
%1RMax * Grip	0.200	0.997	1265.774

The data suggests that there is infinite evidence for the inclusion of %1Rmax than a model without this predictor. (it is 'infinite' because of the computer's limited ability to present very small or very large numbers,). There is also decisive evidence for the inclusion of Grip and the interaction as predictors.

Descriptive data and plots are shown below.

Descriptives

%1RMax	Grip	Mean	SD	N
30%	Reverse	1.279	0.178	10.000
	Traditional	1.482	0.217	10.000
50%	Reverse	1.114	0.198	10.000
	Traditional	1.183	0.256	10.000
70%	Reverse	0.379	0.105	10.000
	Traditional	0.917	0.086	10.000



If the evidence suggested that the data is best predicted by the null model or that the evidence for the alternative was inconsequential. Although evidence for a lack of an effect is still information – there is no point in following up with further analyses.

POST HOC TESTING

If the ANOVA yields meaningful predictors (i.e., models outperforming the null model), post hoc testing can now be carried out. In Post Hoc Tests add %1RM to the analysis box on the right. Bayesian post hoc testing is based on pairwise comparisons using Bayesian t-tests. As in frequentist analyses, multiple t-tests will increase familywise error. In JASP, methods are used to correct for multiplicity based on adjusting the prior odds.

In the analysis options, now:

- ✓ Plots – Model averaged posteriors – Group levels in a single plot

Add %1Rmax and Grip to the right in 'Post Hoc tests'. Select Null control**.**

Post Hoc Comparisons - %1RMax

		Prior Odds	Posterior Odds	$BF_{10, U}$	error %
30%	50%	0.587	20.362	34.664	3.050e −4
	70%	0.587	5.135e +8	8.742e +8	5.573e −14
50%	70%	0.587	6918.088	11777.453	2.806e −8

Note. The posterior odds have been corrected for multiple testing by fixing to 0.5 the prior probability that the null hypothesis holdsacross all comparisons (Westfall, Johnson, & Utts, 1997). Individualcomparisons are based on the default t-test with a Cauchy (0, r =1/sqrt(2)) prior. The "U" in the Bayes factor denotes that it is uncorrected.

The adjusted posterior odds show that there is strong evidence for a difference between 30% and 50% %1Rmax whereas there is decisive evidence for differences between 30 and 70% as well as 50 and 70%1Rmax.

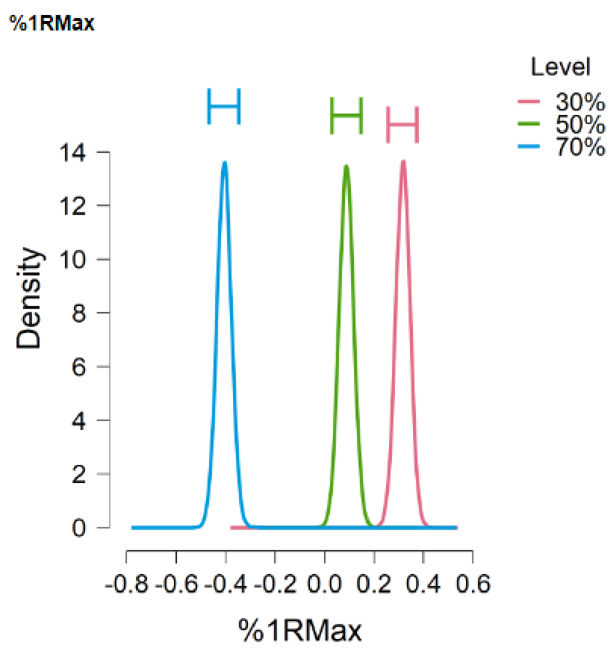
Post Hoc Comparisons - Grip

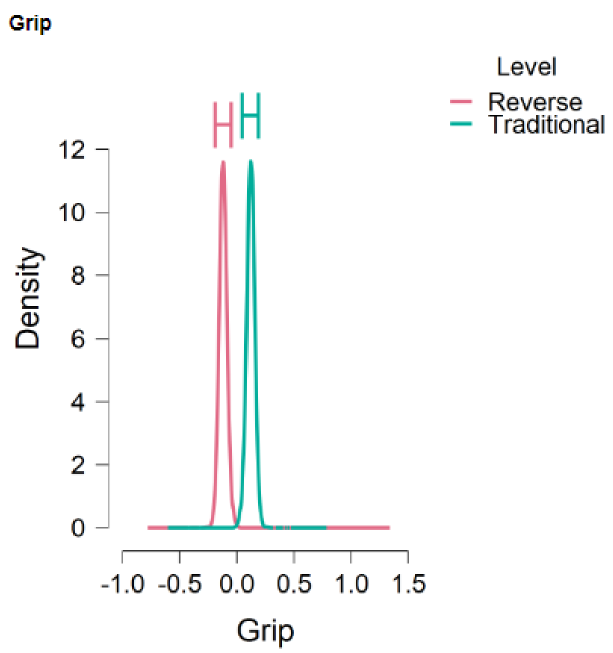
		Prior Odds	Posterior Odds	$BF_{10,U}$	error %
Reverse	Traditional	1.000	6.541	6.541	0.002

Note. The posterior odds have been corrected for multiple testing by fixing to 0.5 the prior probability that the null hypothesis holdsacross all comparisons (Westfall, Johnson, & Utts, 1997). Individualcomparisons are based on the default t-test with a Cauchy (0, r =1/sqrt(2)) prior. The "U" in the Bayes factor denotes that it is uncorrected.

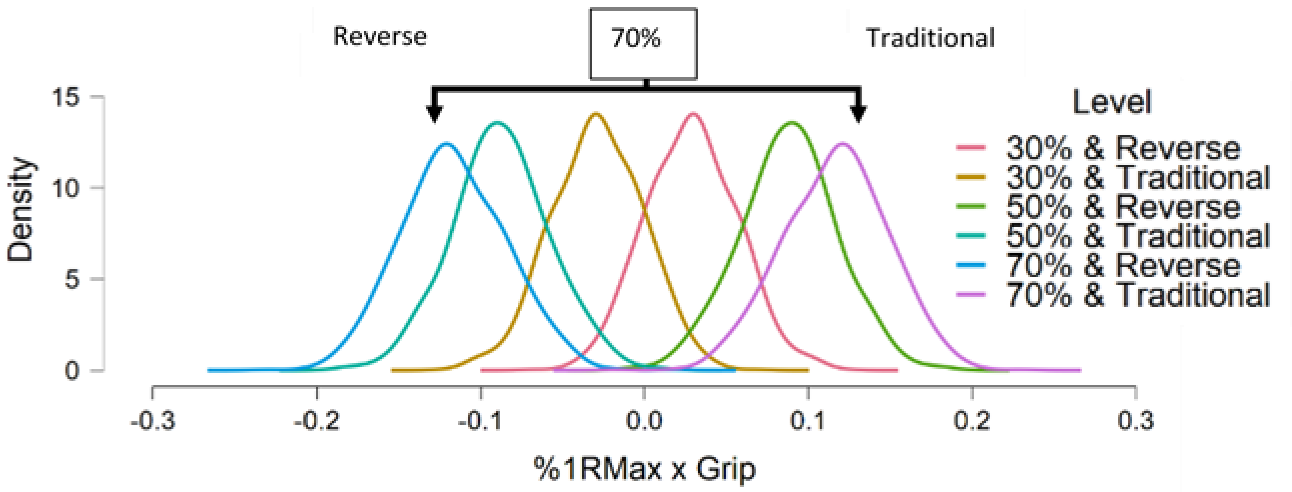
There is also moderate evidence for a difference between reverse and traditional grips $BF_{10} = 6.54$.

The model average posterior distributions for the main effects are shown below. There is a clear separation between the %1Rmax levels with 30% having the highest lift velocity and 70% the lowest. For grip, the two distributions are closer but still separate without overlapping credible intervals, with the traditional grip exhibiting higher lift velocities than the reverse grip.





The model-averaged posterior distributions for the interactions are shown below. As can be seen, the largest separation is between 70% traditional and reverse lifts.



REPORTING THE RESULTS

This study determined the velocity of deadlifts using two different grips and 3 loads based on %1Rmax. Examination of the Q-Q plots suggested that the assumption of normality was not violated. A Bayesian mixed factor ANOVA determined that the data were best represented by a model that included both main factors, grip and load, and the grip*load interaction. The Bayes factor (BF_{10}) was $4.86*10^{16}$, indicating decisive evidence in favour of this model when compared to the null model. The BF_{10} in favour of indicating the interaction effect (on top of the two main effects) equalled 322.

Post hoc comparisons (Bayesian t-tests controlled for multiplicity) were subsequently performed. For the load, the adjusted posteriors show that there is strong evidence for a difference between 30% and 50% %1Rmax (20.6) whereas there is decisive evidence for differences between 30 and 70% as well as 50 and 70%1Rmax ($5.1*10^{8}$ and 6918 respectively).