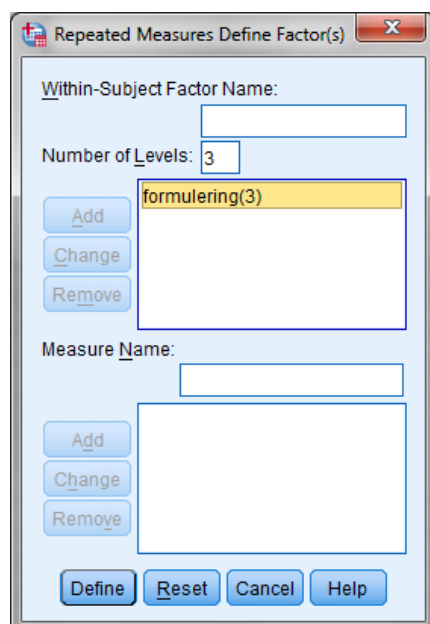


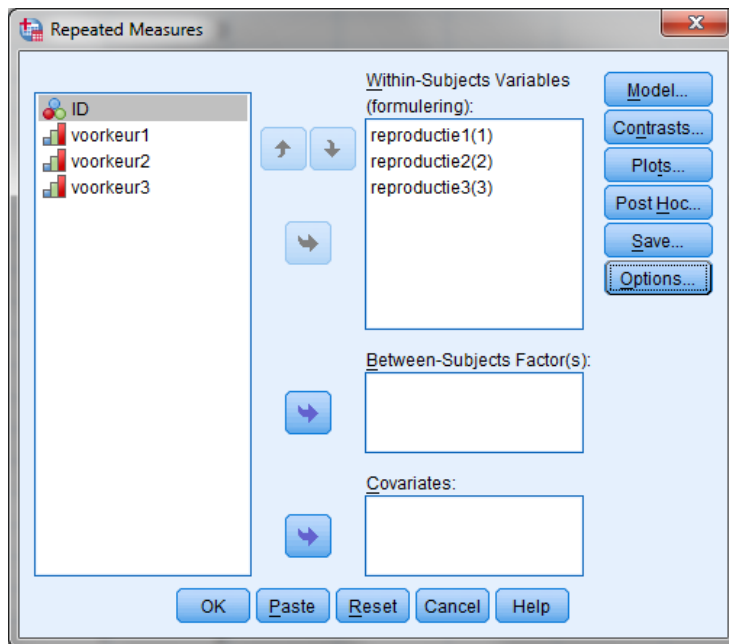
OPLOSSINGEN BIJ HOOFDSTUK 8

1. Eén van de nadelige gevolgen van de moderne welvaart is een monstrueus mobiliteitsprobleem. Om één of andere bizarre reden willen we graag allemaal om acht uur 's ochtends in de wagen springen (liefst ook alleen) om met onze nieuwe auto een mooie plaats in de file te bemachtigen. Allerlei maatregelen ten spijt lijkt het probleem enkel nog groter te worden en duwen we meer en meer wagens tegelijkertijd de weg op. Eén van de maatregelen die worden getroffen in de hoop om mensen te behoeden voor de file is het melden van verkeersinformatie op de radio. Zo'n verkeersbericht bevat, gezien de steeds uitdijende proporties, een overdosis aan informatie die soms moeilijk te onthouden is. Daarom probeert men de berichten te presenteren op een manier die de informatie beter verwerkbaar maakt. Om de optimale volgorde en formulering te vinden test een verkeersredactie een aantal berichten uit door een groep van 35 volwassenen achtereenvolgens verschillende verkeersberichten te laten beluisteren. In totaal zijn er zes inhoudelijk verschillende berichten die elk op drie manieren geformuleerd zijn en in willekeurige volgorde worden aangeboden. Na elk bericht volgt een korte pauze, waarna de deelnemers zoveel mogelijk elementen uit de verkeersinfo proberen te reproduceren.  
De resultaten van dit onderzoek vind je in het bestand *opdr\_verkeersinfo.sav* op de website. Je vindt daarin per deelnemer het gemiddeld aantal elementen die hij of zij zich kon herinneren voor de drie verschillende formuleringen. Ga na welke formulering zorgt voor de beste reproductie.

In dit onderzoek wordt als onafhankelijke variabele de "formulering" gemanipuleerd (nominaal). Afhankelijke variabele is de reproductie, gemeten in aantal herinnerde elementen uit de berichten (interval). Aangezien er drie formuleringen zijn, bestuderen we drie populaties. De steekproeven uit deze populaties zijn afhankelijk van elkaar getrokken, want het betreft telkens dezelfde 35 personen die achtereenvolgens de verschillende berichten te horen krijgen. Er is voldaan aan de assumpties voor parametrisch toetsen, dus we kunnen een repeated measures ANOVA selecteren om deze resultaten te bekijken. De assumptie van sfericiteit zullen we tijdens de analyse bekijken.

Open het databestand en kies Analyze > General Linear Model > Repeated measures. Vul de dialoogvensters in zoals in de afbeelding.





Kies eventueel ook een plot en de nodige opties via de knoppen Plots en Options.

In de tabel met beschrijvende statistieken zien we dat formulering 2 blijkbaar een hogere reproductie oplevert dan de andere twee formuleringen. We zullen verder bekijken of dit verschil ook significant is.

#### Descriptive Statistics

	Mean	Std. Deviation	N
reproductie1	8,6691	1,53296	35
reproductie2	11,1182	1,68037	35
reproductie3	7,7949	1,76912	35

Eerst kijken we nog naar de test voor sfericiteit. Deze geeft aan dat niet voldaan is aan deze assumptie. We zullen bij de interpretatie van de ANOVA dus rekening houden met een correctie voor deze schending van de assumpties.

#### Mauchly's Test of Sphericity<sup>a</sup>

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>b</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
formulering	,804	7,212	2	,027	,836	,874	,500

In de eigenlijke ANOVA-tabel bekijken we dus het significantieniveau op de lijn "Huynh-Feldt" omdat we een correctie voor sfericiteit willen toepassen. Er is blijkbaar een significant effect van de formulering op de reproductie van de inhoud van de berichten.

### Tests of Within-Subjects Effects

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
formulering	Sphericity Assumed	207,742	2	103,871	39,698	,000	,539
	Greenhouse-Geisser	207,742	1,672	124,261	39,698	,000	,539
	Huynh-Feldt	207,742	1,748	118,837	39,698	,000	,539
	Lower-bound	207,742	1,000	207,742	39,698	,000	,539
Error(formulering)	Sphericity Assumed	177,922	68	2,616			
	Greenhouse-Geisser	177,922	56,842	3,130			
	Huynh-Feldt	177,922	59,436	2,994			
	Lower-bound	177,922	34,000	5,233			

In de tabel Pairwise Comparisons vinden we terug tussen welke formuleringen er precies een verschil is. Blijkbaar verschillen formulering 1 en 2, alsook 2 en 3 van elkaar. Formulering 1 verschilt niet significant van formulering 3.

### Pairwise Comparisons

Measure: MEASURE\_1

(I) formulering	(J) formulering	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Lower Bound	Upper Bound
1	2	-2,449 <sup>*</sup>	,289	,000	-3,176	-1,722
	3	,874	,428	,147	-,203	1,952
2	1	2,449 <sup>*</sup>	,289	,000	1,722	3,176
	3	3,323 <sup>*</sup>	,427	,000	2,248	4,398
3	1	-,874	,428	,147	-1,952	,203
	2	-3,323 <sup>*</sup>	,427	,000	-4,398	-2,248

De effectgrootte lezen we ook af uit de ANOVA-tabel, nl .539.

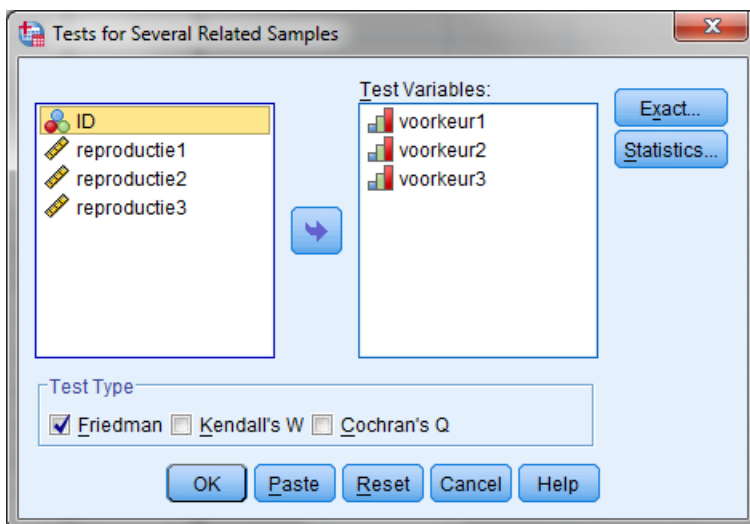
Rapportering:

Om na te gaan of de formulering van een verkeersbericht een invloed heeft op het reproduceren van de inhoud ervan, werd een repeated measures ANOVA uitgevoerd. Mauchly's test geeft aan dat de assumptie van sfericiteit geschonden is,  $\chi^2(2) = 7.21, p = .027$ . Daarom werd de Huynh-Feldt correctie gebruikt bij het interpreteren van het effect. Hieruit bleek dat er een significant effect was van formulering op reproductie,  $F(1.75, 59.44) = 39.70, p < .001, \eta^2 = .54$ . Formulering 2 zorgt voor meer reproductie ( $M = 11.12, SD = 1.68$ ) dan formulering 1 ( $M = 8.66, SD = 1.53, p < .001$ ) en formulering 3 ( $M = 7.80, SD = 1.77, p = .001$ ). Het verschil tussen formulering 1 en formulering 3 was niet significant,  $p = .147$ .

- In het onderzoek uit opdracht 1 werd aan de deelnemers ook gevraagd om de drie verschillende formuleringen te rangschikken volgens hun eigen voorkeur. De formulering die een deelnemer het beste vindt krijgt een 3, de volgende een 2 en de minst goede formulering krijgt een 1. De data staan in hetzelfde bestand. Kan je besluiten dat één van de formuleringen meer de voorkeur wegdraagt dan de andere? Stemt dit overeen met de formulering die de beste reproductie genereert?

Dit onderzoek is gelijkaardig aan het onderzoek in opdracht één, behalve wat het meetniveau van de afhankelijke variabele betreft. De voorkeur van de deelnemers wordt immers gemeten met een ordinale schaal. Dat betekent dat we met de nonparametrische variant van de repeated measures ANOVA zullen werken, met name Friedman's ANOVA.

Kies in het menu Analyze > Nonparametric tests > Legacy Dialogs > k related samples. Vul het dialoogvenster in zoals hieronder en vink Descriptives aan achter de knop Statistics.



We stellen vast dat er een significant verschil is tussen de drie formuleringen:

**Test Statistics<sup>a</sup>**

N	35
Chi-Square	10,000
df	2
Asymp. Sig.	,007

a. Friedman Test

Aan de hand van enkele Wilcoxon Signed Rank tests zullen we nagaan waar de verschillen zich precies voordoen. We zullen alle condities onderling vergelijken en laten dus het alpha-niveau zakken tot  $\alpha/3 = .017$ . We stellen vast dat het verschil tussen formulering 1 en 2 en tussen 1 en 3 significant is.

**Test Statistics<sup>a</sup>**

	voorkeur2 - voorkeur1	voorkeur3 - voorkeur1	voorkeur3 - voorkeur2
Z	-2,442 <sup>b</sup>	-2,618 <sup>b</sup>	-,852 <sup>c</sup>
Asymp. Sig. (2-tailed)	,015	,009	,394

- a. Wilcoxon Signed Ranks Test
- b. Based on positive ranks.
- c. Based on negative ranks.

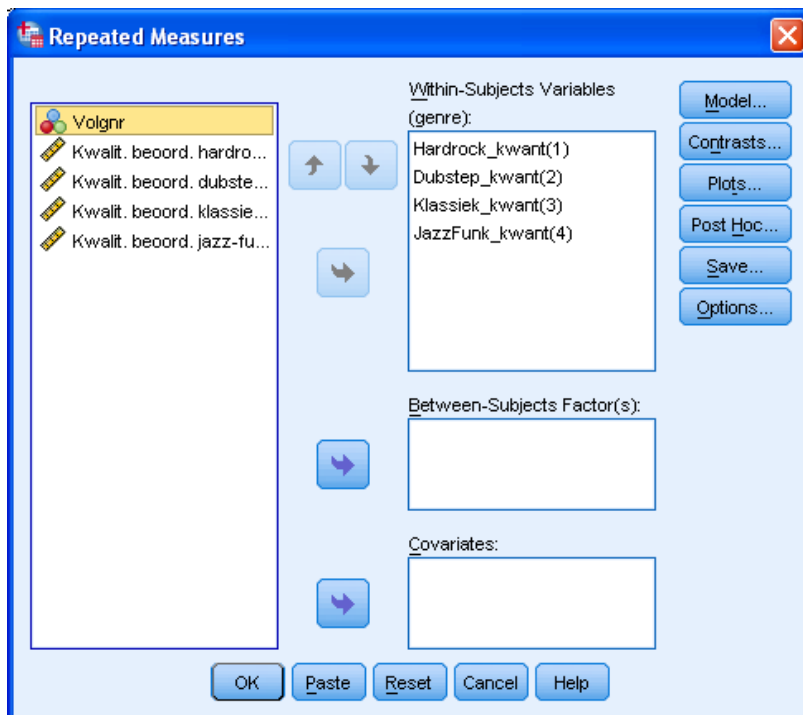
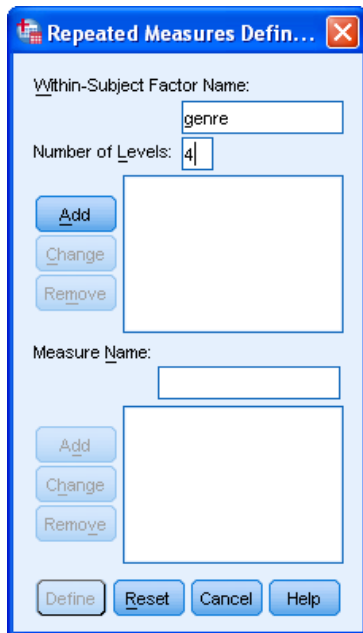
#### Rapportering:

Om de voorkeur van luisteraars voor een bepaalde formulering van verkeersberichten na te gaan werd Friedman's ANOVA uitgevoerd. Hieruit bleek de formulering een significant effect te hebben op de voorkeur,  $F = 10, p = .007$ . Bijkomend werden paarsgewijze Wilcoxon signed-rank toetsen uitgevoerd om de voorkeur voor formulering 1 (mean rank = 3.43), formulering 2 (mean rank = 1.71) en formulering 3 (mean rank = 1.86) onderling te vergelijken. Hierbij werd een gecorrigeerd significantieniveau van  $\alpha = .017$  gehanteerd. Uit deze post hoc toetsen bleken significante verschillen tussen de voorkeuren voor formulering 1 en 2 ( $z = -2.44, p = .015, r = -.29$ ) alsook tussen formulering 1 en 3 ( $z = -2.62, p = .009, r = -.31$ ). Er was geen significant verschil tussen de voorkeuren voor formulering 2 en 3 ( $z = -.85, p = .39, r = -.10$ ).

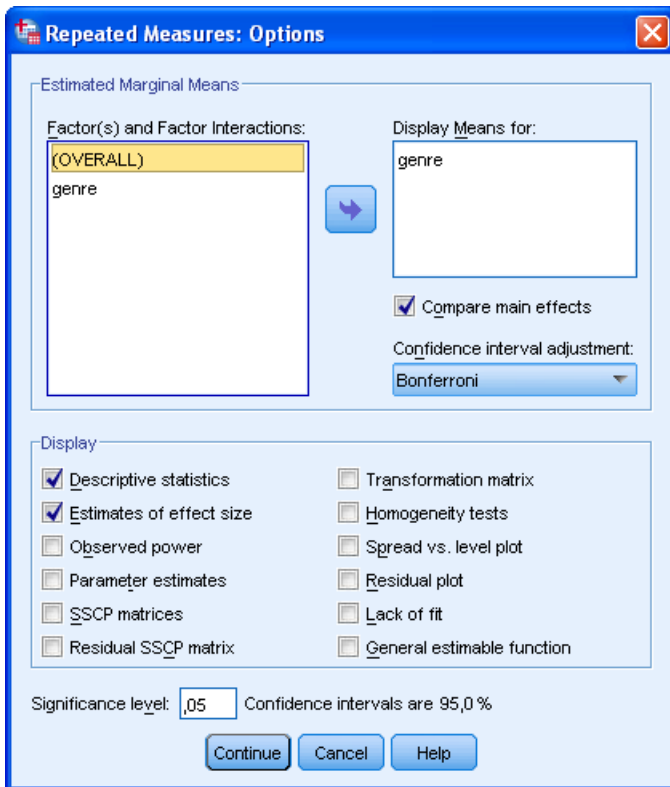
3. Het schrijven van een boek is een aangename en intellectuele bezigheid die bijdraagt tot de zelfverwezenlijking, maar neigt al eens uit te draaien op oncontroleerbare paniek naarmate het einde nadert. Gezien de onwrikbaarheid van de deadline van de uitgever komen er dan immers vaak lange dagen en nachtelijke sessies aan te pas. Sommige auteurs beroepen zich dan op muziek om de motivatie en concentratie te bewaren. Gezien het belang van teksten die geproduceerd worden kan je je afvragen in hoeverre de aard van die muziek een invloed heeft op de kwaliteit van het werk. We testen dit uit door een aantal personen vier dagen op rij te laten schrijven aan een essay over uiteenlopende onderwerpen (geweld in videogames, de teloorgang van het kapitalisme, de wenselijkheid van kookprogramma's op televisie en het nut van lange onderbroeken), die over de dagen gecontrabalanceerd worden. De deelnemers schrijven de hele dag door waarbij ze telkens één van vier muziekgenres te horen krijgen: hard rock, dub step, klassiek en jazzfunk. Op het einde van elke dag dienen ze het essay in, waarvan het aantal woorden geregistreerd wordt en dat vervolgens kwalitatief beoordeeld wordt door enkele onafhankelijke beoordelaars. De deelnemers krijgen zodoende elke dag twee scores: een aantal woorden en een kwaliteitsbeoordeling op 100 punten. Je mag ervan uitgaan dat beide variabelen normaal verdeeld zijn in de populatie. Alle resultaten staan in het bestand *opdr\_essay.sav* op de website. Ga na welke muziek de beste invloed heeft op de kwantiteit en kwaliteit van het werk.

We bestuderen het verband tussen het muziekgenre en twee afhankelijke variabelen: kwantiteit van het essay en kwaliteit van het essay. De onafhankelijke variabele is van nominaal meetniveau en bevat vier niveaus (genres) terwijl de afhankelijke variabelen worden gemeten op intervalniveau. De onderzoekseenheden zijn de deelnemers aan het experiment, die de verschillende condities doorlopen. Het gaat dus om herhaalde metingen of afhankelijke steekproeven.

Bij het checken van de assumptie van normaliteit met het Explore commando stellen we vast dat we mogen aannemen dat de variabelen normaal verdeeld zijn in de populatie. Op die manier komen we terecht bij een repeated measures ANOVA, die we starten via Analyze > General Linear Model > Repeated Measures.



Onder Options:



Eventueel ook nog een grafiek opvragen via Plots.

In de output lezen we een aantal beschrijvende gegevens af:

#### Descriptive Statistics

	Mean	Std. Deviation	N
Aantal woorden hardrock	4772,36	490,149	18
Aantal woorden dubstep	5117,94	355,234	18
Aantal woorden klassiek	3703,20	410,168	18
Aantal woorden jazz-funk	3770,88	420,255	18

De test voor sfericiteit wijst uit dat de assumptie geschonden is. We zullen dus een correctie moeten toepassen.

#### Mauchly's Test of Sphericity<sup>a</sup>

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>b</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
genre	,454	12,410	5	,030	,754	,876	,333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept

Within Subjects Design: genre

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

### Tests of Within-Subjects Effects

Measure: MEASURE\_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	
genre	Sphericity Assumed	27387806,588	3	9129268,863	52,692	,000	,756
	Greenhouse-Geisser	27387806,588	2,262	12109302,418	52,692	,000	,756
	Huynh-Feldt	27387806,588	2,627	10427278,455	52,692	,000	,756
	Lower-bound	27387806,588	1,000	27387806,588	52,692	,000	,756
Error(genre)	Sphericity Assumed	8836182,093	51	173258,472			
	Greenhouse-Geisser	8836182,093	38,449	229814,596			
	Huynh-Feldt	8836182,093	44,651	197892,555			
	Lower-bound	8836182,093	17,000	519775,417			

Als we de Huynh-Feldt correctie toepassen, stellen we vast dat er een significant effect is van Genre.

Om meer gedetailleerd te bekijken tussen welke genres dit effect zich voordoet kijken we naar deze tabel:

### Pairwise Comparisons

Measure: MEASURE\_1

(I) genre	(J) genre	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Lower Bound	Upper Bound
1	2	-345,577	119,048	,059	-700,823	9,668
	3	1069,167*	142,736	,000	643,237	1495,096
	4	1001,487*	176,959	,000	473,435	1529,540
2	1	345,577	119,048	,059	-9,668	700,823
	3	1414,744*	141,528	,000	992,420	1837,068
	4	1347,064*	105,750	,000	1031,502	1662,627
3	1	-1069,167*	142,736	,000	-1495,096	-643,237
	2	-1414,744*	141,528	,000	-1837,068	-992,420
	4	-67,680	135,765	1,000	-472,807	337,448
4	1	-1001,487*	176,959	,000	-1529,540	-473,435
	2	-1347,064*	105,750	,000	-1662,627	-1031,502
	3	67,680	135,765	1,000	-337,448	472,807

Based on estimated marginal means

\*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.



Blijkbaar zijn er verschillen tussen 1 en 3, tussen 1 en 4, tussen 2 en 3, tussen 2 en 4.

Dezelfde werkwijze dienen we nog eens te volgen voor de tweede afhankelijke variabele, kwaliteit van het essay.

Rapportering:

Om na te gaan of het muziekgenre waarnaar mensen luisteren tijdens het schrijven van een essay een invloed heeft op het aantal woorden in het essay en de inhoudelijke kwaliteit van het essay, werden twee repeated measures ANOVA's uitgevoerd.

Uit de eerste ANOVA bleek dat er een effect is van het muziekgenre op het aantal woorden gebruikt in het essay,  $F(2.63, 44.65) = 52.69$ ,  $p < .001$ ,  $\eta^2 = .76$ . Hiervoor werd een Huynh-Feldt correctie toegepast bij gebrek aan sfericiteit in de data. De essays bevatten gemiddeld meer woorden bij het luisteren naar hardrock ( $M = 4772.36$ ,  $SD = 490.15$ ) dan het luisteren naar klassieke muziek ( $M = 3703.20$ ,  $SD = 410.168$ ,  $p < .001$ ) en jazz-funk ( $M = 3770.88$ ,  $SD = 420.26$ ,  $p < .001$ ). Ook het luisteren naar dubstep ( $M = 5117.94$ ,  $SD = 355.234$ ) leverde meer woorden op dan het luisteren naar klassieke muziek ( $p < .001$ ) of jazz-funk ( $p < .001$ ). De andere onderlinge verschillen waren niet significant.

De tweede ANOVA leverde evidentie voor een effect van muziekgenre op de kwaliteit van de essays,  $F(3, 51) = 5.54$ ,  $p = .002$ ,  $\eta^2 = .25$ . Meer bepaald werden de essays uit de jazz-funk conditie hoger beoordeeld ( $M = 82.55$ ,  $SD = 18.38$ ) dan de essays uit de hardrock conditie ( $M = 62.41$ ,  $SD = 11.41$ ,  $p = .014$ ) en de dubstep conditie ( $M = 62.87$ ,  $SD = 21.29$ ,  $p = .035$ ). De verschillen tussen deze condities en de conditie met klassieke muziek ( $M = 73.47$ ,  $SD = 18.80$ ) waren niet significant.