

Complete Solutions

Analysis of Variance

Ratiomera Statistics

These complete solutions use the same identifiers and order as the Exercise Sheet. Intermediate values are retained until the stated rounding step, so small differences caused by earlier rounding are acceptable where noted. All settings, values, data, and software outputs are constructed teaching material; they are not empirical findings.

Part I: Theory

A01: The Question Answered by One-Way ANOVA

T08-A01-V01: Reading formats and comprehension

Identify the issue

The quantitative outcome is comprehension score, measured in points.

The factor is reading format, with levels print, tablet, audio.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because cases were randomly assigned, a well-conducted study can support a causal interpretation for these conditions, subject to its design and assumptions.

T08-A01-V02: Museum routes and visit duration

Identify the issue

The quantitative outcome is visit duration, measured in minutes.

The factor is route type, with levels free route, numbered route, guided route.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because cases were randomly assigned, a well-conducted study can support a causal interpretation for these conditions, subject to its design and assumptions.

T08-A01-V03: Study locations and concentration**Identify the issue**

The quantitative outcome is concentration score, measured in points.

The factor is usual study location, with levels home, library, shared workspace.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because the groups were observed rather than randomly assigned, a difference describes association and does not by itself identify a causal effect.

T08-A01-V04: Reminder schedules and response delay**Identify the issue**

The quantitative outcome is response delay, measured in hours.

The factor is reminder schedule, with levels none, one reminder, three reminders.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because cases were randomly assigned, a well-conducted study can support a causal interpretation for these conditions, subject to its design and assumptions.

T08-A01-V05: Archive interfaces and retrieval accuracy**Identify the issue**

The quantitative outcome is retrieval accuracy, measured in points.

The factor is interface version, with levels standard, compact, guided.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because cases were randomly assigned, a well-conducted study can support a causal interpretation for these conditions, subject to its design and assumptions.

T08-A01-V06: Travel modes and commuting time**Identify the issue**

The quantitative outcome is commuting time, measured in minutes.

The factor is usual travel mode, with levels walking, public transport, car.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because the groups were observed rather than randomly assigned, a difference describes association and does not by itself identify a causal effect.

T08-A01-V07: Practice routines and delayed recall**Identify the issue**

The quantitative outcome is delayed-recall score, measured in points.

The factor is practice routine, with levels rereading, self-testing, mixed practice.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because cases were randomly assigned, a well-conducted study can support a causal interpretation for these conditions, subject to its design and assumptions.

T08-A01-V08: Workshop tracks and confidence

Identify the issue

The quantitative outcome is confidence score, measured in points.

The factor is chosen workshop track, with levels methods, writing, presentation.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because the groups were observed rather than randomly assigned, a difference describes association and does not by itself identify a causal effect.

T08-A01-V09: Caption styles and tutorial understanding

Identify the issue

The quantitative outcome is understanding score, measured in points.

The factor is caption style, with levels none, verbatim, edited.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because cases were randomly assigned, a well-conducted study can support a causal interpretation for these conditions, subject to its design and assumptions.

T08-A01-V10: Neighborhood types and park use

Identify the issue

The quantitative outcome is weekly park visits, measured in visits.

The factor is neighborhood type, with levels central, suburban, rural.

Write the population means as μ_1 , μ_2 , and μ_3 .

Reason through the evidence

The null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$.

The alternative is that at least two means differ.

It does not say that every pair differs.

State the conclusion and its limits

Beginning with three separate tests would create a family of opportunities for a Type I error, while the omnibus F test asks the single global question first.

A significant result would provide evidence against equality of all three population means, but it would not identify which pairs differ.

Because the groups were observed rather than randomly assigned, a difference describes association and does not by itself identify a causal effect.

A06: Defining a Comparison Family Before Looking at Results

T08-A06-V01: Study routines

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V02: Reading layouts

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V03: Archive prompts

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V04: Museum routes

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V05: Reminder schedules

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	none
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	none
Analyst 5	5	0.010	none

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V06: Note templates

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V07: Practice intervals

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V08: Sound settings

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V09: Navigation aids

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

T08-A06-V10: Feedback schedules

Identify the issue, part (a)

and

Reason through the evidence, part (b)

A family contains every comparison made available by the decision process, including comparisons examined and later hidden. The resulting family sizes and thresholds are:

Analyst	Family size	Bonferroni threshold	Results meeting threshold
Analyst 1	2	0.025	A versus B, A versus C
Analyst 2	5	0.010	A versus B
Analyst 3	1	0.050	A versus D
Analyst 4	5	0.010	A versus B
Analyst 5	5	0.010	A versus B

Reason through the evidence, part (c)

Each result in the final column is obtained by comparing only the p-values in that analyst's honest family with $0.05/m$. Analyst 1 has two prospectively named comparisons, Analyst 2 has five, and Analyst 3 has one. Analysts 4 and 5 also have five because their reduction happened after the outcomes were visible.

Reason through the evidence, part (d)

Removing the largest result or highlighting the smallest result after inspection is data-dependent selection. It cannot erase the other opportunities for a Type I error, and it cannot make the highlighted question planned retrospectively.

State the conclusion and its limits, part (e)

Before viewing outcomes, each analyst needed to record the scientific comparison, its direction or contrast weights where relevant, the complete family of primary and secondary comparisons, and the chosen multiplicity rule.

Part II: Calculator Practice

A02: Group Means, Sum-of-Squares Partition, and the One-Way F Test

T08-A02-V01: Study routines and learning score

Set up the calculation

The group means are 12, 16, 20, and the grand mean is 16.0000 points.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 128.0000$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 188.0000$, and $188.0000 = 128.0000 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 128.0000/2 = 64.0000$, $MS_e = 60.0000/9 = 6.6667$, and $F = 64.0000/6.6667 = 9.6000$.

Since 9.6000 is greater than 4.26, we reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V02: Reading layouts and speed

Set up the calculation

The group means are 20, 22, 24, and the grand mean is 22.0000 words per minute above baseline.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 32.0000$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 92.0000$, and $92.0000 = 32.0000 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 32.0000/2 = 16.0000$, $MS_e = 60.0000/9 = 6.6667$, and $F = 16.0000/6.6667 = 2.4000$.

Since 2.4000 is not greater than 4.26, we do not reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V03: Archive prompts and correct records

Set up the calculation

The group means are 15, 15, 15, and the grand mean is 15.0000 records.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 0.0000$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 60.0000$, and $60.0000 = 0.0000 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 0.0000/2 = 0.0000$, $MS_e = 60.0000/9 = 6.6667$, and $F = 0.0000/6.6667 = 0.0000$.

Since 0.0000 is not greater than 4.26, we do not reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V04: Museum maps and route confidence

Set up the calculation

The group means are 45, 50, 47, and the grand mean is 47.3333 points.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 50.6667$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 110.6667$, and $110.6667 = 50.6667 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 50.6667/2 = 25.3333$, $MS_e = 60.0000/9 = 6.6667$, and $F = 25.3333/6.6667 = 3.8000$.

Since 3.8000 is not greater than 4.26, we do not reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V05: Reminder timing and completion

Set up the calculation

The group means are 30, 35, 39, and the grand mean is 34.6667 points.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 162.6667$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 222.6667$, and $222.6667 = 162.6667 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 162.6667/2 = 81.3333$, $MS_e = 60.0000/9 = 6.6667$, and $F = 81.3333/6.6667 = 12.2000$.

Since 12.2000 is greater than 4.26, we reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V06: Note formats and argument detection

Set up the calculation

The group means are 52, 56, 61, and the grand mean is 56.3333 points.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 162.6667$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 222.6667$, and $222.6667 = 162.6667 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 162.6667/2 = 81.3333$, $MS_e = 60.0000/9 = 6.6667$, and $F = 81.3333/6.6667 = 12.2000$.

Since 12.2000 is greater than 4.26, we reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V07: Practice spacing and recall

Set up the calculation

The group means are 40, 43, 48, and the grand mean is 43.6667 points.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 130.6667$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 190.6667$, and $190.6667 = 130.6667 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 130.6667/2 = 65.3333$, $MS_e = 60.0000/9 = 6.6667$, and $F = 65.3333/6.6667 = 9.8000$.

Since 9.8000 is greater than 4.26, we reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V08: Sound settings and focus

Set up the calculation

The group means are 70, 68, 63, and the grand mean is 67.0000 points.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 104.0000$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 164.0000$, and $164.0000 = 104.0000 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 104.0000/2 = 52.0000$, $MS_e = 60.0000/9 = 6.6667$, and $F = 52.0000/6.6667 = 7.8000$.

Since 7.8000 is greater than 4.26, we reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V09: Route instructions and errors

Set up the calculation

The group means are 18, 14, 10, and the grand mean is 14.0000 errors.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 128.0000$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 188.0000$, and $188.0000 = 128.0000 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 128.0000/2 = 64.0000$, $MS_e = 60.0000/9 = 6.6667$, and $F = 64.0000/6.6667 = 9.6000$.

Since 9.6000 is greater than 4.26, we reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

T08-A02-V10: Feedback timing and revision quality

Set up the calculation

The group means are 64, 67, 65, and the grand mean is 65.3333 points.

The between-group calculation gives $SS_A = \sum_i n_i (\bar{y}_i - \bar{y})^2 = 18.6667$.

Work through the calculation

Summing squared deviations inside the three groups gives $SS_e = 60.0000$.

Around the grand mean, $SS_{total} = 78.6667$, and $78.6667 = 18.6667 + 60.0000$, so the partition checks.

The degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 12 - 3 = 9$, and $df_{total} = 11$.

Interpret and check the result

Thus $MS_A = 18.6667/2 = 9.3333$, $MS_e = 60.0000/9 = 6.6667$, and $F = 9.3333/6.6667 = 1.4000$.

Since 1.4000 is not greater than 4.26, we do not reject the equal-means null at the 5% level.

The decision concerns the global mean pattern, not every pair separately.

A03: Reconstructing an ANOVA Table and Reading the Design

T08-A03-V01: Randomized reading prompts

Reason before calculating

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $96 + 144 = 240$.

Work through the calculation

With $N = 24$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 24 - 3 = 21$, and $df_{total} = 24 - 1 = 23$.

Then $MS_A = 96/2 = 48.0000$, $MS_e = 144/21 = 6.8571$, and $F = 7.0000$.

Interpret and check the result

Because 7.0000 is greater than 3.44, we reject the null at the 5% level.

The equal group sizes make the design balanced.

Random assignment can support a causal interpretation for the assigned conditions if the implementation and model assumptions are credible.

T08-A03-V02: Observed commuting modes

Reason before calculating

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $45+210=255$.

Work through the calculation

With $N = 26$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 26 - 3 = 23$, and $df_{total} = 26 - 1 = 25$.

Then $MS_A = 45/2 = 22.5000$, $MS_e = 210/23 = 9.1304$, and $F = 2.4643$.

Interpret and check the result

Because 2.4643 is not greater than 3.42, we do not reject the null at the 5% level.

The unequal group sizes make the design unbalanced.

Without random assignment, the result describes group association and cannot by itself rule out preexisting group differences.

T08-A03-V03: Randomized archive interfaces

Reason before calculating

For $k = 4$ groups, $H_0 : \mu_1 = \dots = \mu_4$; the alternative is that at least two means differ.

The total sum of squares is $180+220=400$.

Work through the calculation

With $N = 24$, the degrees of freedom are $df_A = 4 - 1 = 3$, $df_e = 24 - 4 = 20$, and $df_{total} = 24 - 1 = 23$.

Then $MS_A = 180/3 = 60.0000$, $MS_e = 220/20 = 11.0000$, and $F = 5.4545$.

Interpret and check the result

Because 5.4545 is greater than 3.10, we reject the null at the 5% level.

The equal group sizes make the design balanced.

Random assignment can support a causal interpretation for the assigned conditions if the implementation and model assumptions are credible.

T08-A03-V04: Self-selected workshop tracks

Reason before calculating

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $30+330=360$.

Work through the calculation

With $N = 36$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 36 - 3 = 33$, and $df_{total} = 36 - 1 = 35$.

Then $MS_A = 30/2 = 15.0000$, $MS_e = 330/33 = 10.0000$, and $F = 1.5000$.

Interpret and check the result

Because 1.5000 is not greater than 3.28, we do not reject the null at the 5% level.

The equal group sizes make the design balanced.

Without random assignment, the result describes group association and cannot by itself rule out preexisting group differences.

T08-A03-V05: Randomized reminder schedules

Reason before calculating

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $120+180=300$.

Work through the calculation

With $N = 27$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 27 - 3 = 24$, and $df_{total} = 27 - 1 = 26$.

Then $MS_A = 120/2 = 60.0000$, $MS_e = 180/24 = 7.5000$, and $F = 8.0000$.

Interpret and check the result

Because 8.0000 is greater than 3.35, we reject the null at the 5% level.

The equal group sizes make the design balanced.

Random assignment can support a causal interpretation for the assigned conditions if the implementation and model assumptions are credible.

T08-A03-V06: Observed neighborhood types

Reason before calculating

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $75+270=345$.

Work through the calculation

With $N = 30$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 30 - 3 = 27$, and $df_{total} = 30 - 1 = 29$.

Then $MS_A = 75/2 = 37.5000$, $MS_e = 270/27 = 10.0000$, and $F = 3.7500$.

Interpret and check the result

Because 3.7500 is greater than 3.35, we reject the null at the 5% level.

The unequal group sizes make the design unbalanced.

Without random assignment, the result describes group association and cannot by itself rule out preexisting group differences.

T08-A03-V07: Randomized caption styles

Reason before calculating

For $k = 4$ groups, $H_0 : \mu_1 = \dots = \mu_4$; the alternative is that at least two means differ.

The total sum of squares is $210+252=462$.

Work through the calculation

With $N = 28$, the degrees of freedom are $df_A = 4 - 1 = 3$, $df_e = 28 - 4 = 24$, and $df_{total} = 28 - 1 = 27$.

Then $MS_A = 210/3 = 70.0000$, $MS_e = 252/24 = 10.5000$, and $F = 6.6667$.

Interpret and check the result

Because 6.6667 is greater than 3.01, we reject the null at the 5% level.

The equal group sizes make the design balanced.

Random assignment can support a causal interpretation for the assigned conditions if the implementation and model assumptions are credible.

T08-A03-V08: Chosen study locations

Reason before calculating

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $54+243=297$.

Work through the calculation

With $N = 30$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 30 - 3 = 27$, and $df_{total} = 30 - 1 = 29$.

Then $MS_A = 54/2 = 27.0000$, $MS_e = 243/27 = 9.0000$, and $F = 3.0000$.

Interpret and check the result

Because 3.0000 is not greater than 3.35, we do not reject the null at the 5% level.

The unequal group sizes make the design unbalanced.

Without random assignment, the result describes group association and cannot by itself rule out preexisting group differences.

T08-A03-V09: Randomized route maps**Reason before calculating**

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $160+240=400$.

Work through the calculation

With $N = 30$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 30 - 3 = 27$, and $df_{total} = 30 - 1 = 29$.

Then $MS_A = 160/2 = 80.0000$, $MS_e = 240/27 = 8.8889$, and $F = 9.0000$.

Interpret and check the result

Because 9.0000 is greater than 3.35, we reject the null at the 5% level.

The equal group sizes make the design balanced.

Random assignment can support a causal interpretation for the assigned conditions if the implementation and model assumptions are credible.

T08-A03-V10: Observed employment sectors**Reason before calculating**

For $k = 3$ groups, $H_0 : \mu_1 = \dots = \mu_3$; the alternative is that at least two means differ.

The total sum of squares is $40+260=300$.

Work through the calculation

With $N = 30$, the degrees of freedom are $df_A = 3 - 1 = 2$, $df_e = 30 - 3 = 27$, and $df_{total} = 30 - 1 = 29$.

Then $MS_A = 40/2 = 20.0000$, $MS_e = 260/27 = 9.6296$, and $F = 2.0769$.

Interpret and check the result

Because 2.0769 is not greater than 3.35, we do not reject the null at the 5% level.

The unequal group sizes make the design unbalanced.

Without random assignment, the result describes group association and cannot by itself rule out preexisting group differences.

A04: Simple Pairwise and Pooled Complex Contrasts**T08-A04-V01: Four study routines****Set up the calculation**

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 4.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(25/8)[(-1)^2 + 1^2]} = 2.5000$, giving $t_s = 4.0000/2.5000 = 1.6000$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 16.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(25/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 3.5355$, so $t_c = 4.5255$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V02: Four reading layouts

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 5.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(16/8)[(-1)^2 + 1^2]} = 2.0000$, giving $t_s = 5.0000/2.0000 = 2.5000$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 13.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(16/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 2.8284$, so $t_c = 4.5962$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V03: Four archive prompts

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 3.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(9/8)[(-1)^2 + 1^2]} = 1.5000$, giving $t_s = 3.0000/1.5000 = 2.0000$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 8.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(9/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 2.1213$, so $t_c = 3.7712$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V04: Four museum routes

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = -3.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(36/8)[(-1)^2 + 1^2]} = 3.0000$, giving $t_s = -3.0000/3.0000 = -1.0000$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 15.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(36/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 4.2426$, so $t_c = 3.5355$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V05: Four reminder schedules

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 5.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(20/8)[(-1)^2 + 1^2]} = 2.2361$, giving $t_s = 5.0000/2.2361 = 2.2361$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 14.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(20/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 3.1623$, so $t_c = 4.4272$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V06: Four note templates

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 3.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(24/8)[(-1)^2 + 1^2]} = 2.4495$, giving $t_s = 3.0000/2.4495 = 1.2247$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 17.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(24/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 3.4641$, so $t_c = 4.9075$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V07: Four practice intervals

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 5.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(30/8)[(-1)^2 + 1^2]} = 2.7386$, giving $t_s = 5.0000/2.7386 = 1.8257$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 17.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(30/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 3.8730$, so $t_c = 4.3894$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V08: Four sound settings

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = -4.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(18/8)[(-1)^2 + 1^2]} = 2.1213$, giving $t_s = -4.0000/2.1213 = -1.8856$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = -11.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(18/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 3.0000$, so $t_c = -3.6667$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V09: Four navigation aids

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 5.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(22/8)[(-1)^2 + 1^2]} = 2.3452$, giving $t_s = 5.0000/2.3452 = 2.1320$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 17.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(22/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 3.3166$, so $t_c = 5.1257$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

T08-A04-V10: Four feedback schedules

Set up the calculation

Both weight vectors sum to zero, so both are valid contrasts.

The simple contrast compares group 2 with group 1: $D_s = 4.0000$.

Work through the calculation

Its standard error is $SE_s = \sqrt{(28/8)[(-1)^2 + 1^2]} = 2.6458$, giving $t_s = 4.0000/2.6458 = 1.5119$.

The complex contrast compares the sum of groups 3 and 4 with the sum of groups 1 and 2: $D_c = 8.0000$.

Interpret and check the result

Its standard error is $SE_c = \sqrt{(28/8)[(-1)^2 + (-1)^2 + 1^2 + 1^2]} = 3.7417$, so $t_c = 2.1381$.

Dividing every complex weight by 2 would express the difference between the two pooled averages and would leave its t statistic unchanged because its estimate and standard error scale together.

Raw contrast estimates use different scales when their weights differ, so compare the stated question and standardized statistic, not magnitude alone.

A05: All Pairwise Comparisons and Bonferroni Protection

T08-A05-V01: All pairs among 3 levels

Set up the calculation

The number of distinct unordered pairs is $m = k(k-1)/2 = 3(3-1)/2 = 3$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/3 = 0.0167$.

If all 3 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^3 = 0.1426$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V02: All pairs among 4 levels

Set up the calculation

The number of distinct unordered pairs is $m = k(k-1)/2 = 4(4-1)/2 = 6$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/6 = 0.0083$.

If all 6 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^6 = 0.2649$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V03: All pairs among 5 levels**Set up the calculation**

The number of distinct unordered pairs is $m = k(k - 1)/2 = 5(5 - 1)/2 = 10$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/10 = 0.0050$.

If all 10 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^{10} = 0.4013$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V04: All pairs among 6 levels**Set up the calculation**

The number of distinct unordered pairs is $m = k(k - 1)/2 = 6(6 - 1)/2 = 15$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/15 = 0.0033$.

If all 15 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^{15} = 0.5367$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V05: All pairs among 7 levels**Set up the calculation**

The number of distinct unordered pairs is $m = k(k - 1)/2 = 7(7 - 1)/2 = 21$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/21 = 0.0024$.

If all 21 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^{21} = 0.6594$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V06: All pairs among 8 levels**Set up the calculation**

The number of distinct unordered pairs is $m = k(k - 1)/2 = 8(8 - 1)/2 = 28$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/28 = 0.0018$.

If all 28 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^{28} = 0.7622$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V07: All pairs among 4 levels

Set up the calculation

The number of distinct unordered pairs is $m = k(k - 1)/2 = 4(4 - 1)/2 = 6$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/6 = 0.0083$.

If all 6 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^6 = 0.2649$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V08: All pairs among 5 levels

Set up the calculation

The number of distinct unordered pairs is $m = k(k - 1)/2 = 5(5 - 1)/2 = 10$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/10 = 0.0050$.

If all 10 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^{10} = 0.4013$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V09: All pairs among 6 levels

Set up the calculation

The number of distinct unordered pairs is $m = k(k - 1)/2 = 6(6 - 1)/2 = 15$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/15 = 0.0033$.

If all 15 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^{15} = 0.5367$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

T08-A05-V10: All pairs among 7 levels

Set up the calculation

The number of distinct unordered pairs is $m = k(k - 1)/2 = 7(7 - 1)/2 = 21$.

Work through the calculation

Bonferroni uses $\alpha_{test} = 0.05/21 = 0.0024$.

If all 21 tests were independent and each used 0.05, the probability of at least one Type I error under the complete null would be $1 - (1 - 0.05)^{21} = 0.6594$.

Interpret and check the result

Actual pairwise comparisons often share groups and are therefore dependent, so that last expression is an illustration rather than their general exact familywise rate.

Bonferroni relies on an upper bound for the probability of a union of errors, so it controls the family even without independence, although it can be conservative.

A07: A Prespecified Trend Contrast

T08-A07-V01: Practice sessions and recall

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = 57.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(25/10)(9 + 1 + 1 + 9)} = 7.0711$. Thus $t = 57.0000/7.0711 = 8.0610$.

Interpret and check the result, part (e)

Since $|8.0610|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The positive sign means the weighted pattern tends to rise across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V02: Reminder intensity and response

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = 36.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(16/10)(9 + 1 + 1 + 9)} = 5.6569$. Thus $t = 36.0000/5.6569 = 6.3640$.

Interpret and check the result, part (e)

Since $|6.3640|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The positive sign means the weighted pattern tends to rise across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V03: Reading guidance and comprehension

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = 40.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(20/10)(9 + 1 + 1 + 9)} = 6.3246$. Thus $t = 40.0000/6.3246 = 6.3246$.

Interpret and check the result, part (e)

Since $|6.3246|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The positive sign means the weighted pattern tends to rise across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V04: Archive examples and accuracy

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = 31.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(9/10)(9 + 1 + 1 + 9)} = 4.2426$. Thus $t = 31.0000/4.2426 = 7.3068$.

Interpret and check the result, part (e)

Since $|7.3068|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The positive sign means the weighted pattern tends to rise across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V05: Route rehearsal and confidence

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = 51.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(24/10)(9 + 1 + 1 + 9)} = 6.9282$. Thus $t = 51.0000/6.9282 = 7.3612$.

Interpret and check the result, part (e)

Since $|7.3612|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The positive sign means the weighted pattern tends to rise across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V06: Note structure and reasoning

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = 42.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(18/10)(9 + 1 + 1 + 9)} = 6.0000$. Thus $t = 42.0000/6.0000 = 7.0000$.

Interpret and check the result, part (e)

Since $|7.0000|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The positive sign means the weighted pattern tends to rise across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V07: Feedback frequency and revision

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = 31.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(22/10)(9 + 1 + 1 + 9)} = 6.6332$. Thus $t = 31.0000/6.6332 = 4.6734$.

Interpret and check the result, part (e)

Since $|4.6734|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The positive sign means the weighted pattern tends to rise across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V08: Ambient noise and focus

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = -39.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(16/10)(9 + 1 + 1 + 9)} = 5.6569$. Thus $t = -39.0000/5.6569 = -6.8943$.

Interpret and check the result, part (e)

Since $|-6.8943|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The negative sign means the weighted pattern tends to fall across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V09: Navigation support and errors

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = -37.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(12/10)(9 + 1 + 1 + 9)} = 4.8990$. Thus $t = -37.0000/4.8990 = -7.5526$.

Interpret and check the result, part (e)

Since $|-7.5526|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The negative sign means the weighted pattern tends to fall across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

T08-A07-V10: Delay before feedback and retention

Reason before calculating, part (a)

The weights c_1 rise by one unit at every level, so their sketch is a straight increasing sequence. The weights c_2 also rise by equal increments and are centered around zero. The weights c_3 compare the two endpoints with the two middle levels, so they represent curvature rather than a linear trend.

Work through the calculation, part (b)

The sums are $0 + 1 + 2 + 3 = 6$, $-3 - 1 + 1 + 3 = 0$, and $0.5 - 0.5 - 0.5 + 0.5 = 0$. Thus c_2 and c_3 are contrasts, but only c_2 is the stated linear-trend contrast.

Work through the calculation, part (c)

The mean of c_1 is 1.5. Subtracting it gives $(-1.5, -0.5, 0.5, 1.5)$, which is exactly one half of c_2 . Multiplying every contrast weight by the same positive constant changes the numerical scale of the estimate and standard error together but not the tested direction or t statistic.

Work through the calculation, part (d)

With c_2 , the weighted estimate is $D = \sum c_i \bar{y}_i = -40.0000$. Because the groups are balanced, $SE(D) = \sqrt{(MS_e/n) \sum c_i^2} = \sqrt{(20/10)(9 + 1 + 1 + 9)} = 6.3246$. Thus $t = -40.0000/6.3246 = -6.3246$.

Interpret and check the result, part (e)

Since $|-6.3246|$ is greater than 2.028, the linear trend meets the two-sided 5% criterion. The negative sign means the weighted pattern tends to fall across the ordered levels. It does not prove equal adjacent changes or rule out curvature.

A08: Cell Means, Marginal Means, and Interaction in a Two-Factor ANOVA

T08-A08-V01: Captioning and practice

Reason before calculating, part (a)

The factor-A marginal means are (65.0000, 73.0000); the factor-B marginal means are (66.0000, 72.0000); and the grand mean is 69.0000.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B is the same at both levels of factor A, so these cell means show no interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 384.0000$, $SS_B = 216.0000$, and $SS_{AB} = 0.0000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 16(20) = 320.0000$. Therefore $F_A = 24.0000$, $F_B = 13.5000$, and $F_{AB} = 0.0000$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 62.0000)$ and $(B_1, 68.0000)$. Plot A_1 through $(B_0, 70.0000)$ and $(B_1, 76.0000)$. The two changes are equal, so the lines are parallel and the plot displays no interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V02: Map and route rehearsal

Reason before calculating, part (a)

The factor-A marginal means are (57.0000, 64.0000); the factor-B marginal means are (56.0000, 65.0000); and the grand mean is 60.5000.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 294.0000$, $SS_B = 486.0000$, and $SS_{AB} = 54.0000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 20(20) = 400.0000$. Therefore $F_A = 14.7000$, $F_B = 24.3000$, and $F_{AB} = 2.7000$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 54.0000)$ and $(B_1, 60.0000)$. Plot A_1 through $(B_0, 58.0000)$ and $(B_1, 70.0000)$. The two changes differ, so the lines are nonparallel and the plot displays the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V03: Quiet room and checklist

Reason before calculating, part (a)

The factor-A marginal means are (68.0000, 72.0000); the factor-B marginal means are (67.0000, 73.0000); and the grand mean is 70.0000.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B is the same at both levels of factor A, so these cell means show no interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 96.0000$, $SS_B = 216.0000$, and $SS_{AB} = 0.0000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 18(20) = 360.0000$. Therefore $F_A = 5.3333$, $F_B = 12.0000$, and $F_{AB} = 0.0000$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 65.0000)$ and $(B_1, 71.0000)$. Plot A_1 through $(B_0, 69.0000)$ and $(B_1, 75.0000)$. The two changes are equal, so the lines are parallel and the plot displays no interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V04: Prompt and feedback

Reason before calculating, part (a)

The factor-A marginal means are (54.5000, 58.5000); the factor-B marginal means are (53.0000, 60.0000); and the grand mean is 56.5000.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 96.0000$, $SS_B = 294.0000$, and $SS_{AB} = 24.0000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 15(20) = 300.0000$. Therefore $F_A = 6.4000$, $F_B = 19.6000$, and $F_{AB} = 1.6000$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 50.0000)$ and $(B_1, 59.0000)$. Plot A_1 through $(B_0, 56.0000)$ and $(B_1, 61.0000)$. The two changes differ, so the lines are nonparallel and the plot displays the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V05: Icons and examples

Reason before calculating, part (a)

The factor-A marginal means are (73.0000, 79.5000); the factor-B marginal means are (74.0000, 78.5000); and the grand mean is 76.2500.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 253.5000$, $SS_B = 121.5000$, and $SS_{AB} = 37.5000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 24(20) = 480.0000$. Therefore $F_A = 10.5625$, $F_B = 5.0625$, and $F_{AB} = 1.5625$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 72.0000)$ and $(B_1, 74.0000)$. Plot A_1 through $(B_0, 76.0000)$ and $(B_1, 83.0000)$. The two changes differ, so the lines are nonparallel and the plot displays the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V06: Planning and self-testing

Reason before calculating, part (a)

The factor-A marginal means are (63.5000, 72.0000); the factor-B marginal means are (63.0000, 72.5000); and the grand mean is 67.7500.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 433.5000$, $SS_B = 541.5000$, and $SS_{AB} = 37.5000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 21(20) = 420.0000$. Therefore $F_A = 20.6429$, $F_B = 25.7857$, and $F_{AB} = 1.7857$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 60.0000)$ and $(B_1, 67.0000)$. Plot A_1 through $(B_0, 66.0000)$ and $(B_1, 78.0000)$. The two changes differ, so the lines are nonparallel and the plot displays

the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V07: Lighting and background sound

Reason before calculating, part (a)

The factor-A marginal means are (70.5000, 67.5000); the factor-B marginal means are (72.0000, 66.0000); and the grand mean is 69.0000.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 54.0000$, $SS_B = 216.0000$, and $SS_{AB} = 6.0000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 17(20) = 340.0000$. Therefore $F_A = 3.1765$, $F_B = 12.7059$, and $F_{AB} = 0.3529$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 74.0000)$ and $(B_1, 67.0000)$. Plot A_1 through $(B_0, 70.0000)$ and $(B_1, 65.0000)$. The two changes differ, so the lines are nonparallel and the plot displays the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V08: Orientation and signs

Reason before calculating, part (a)

The factor-A marginal means are (51.5000, 63.5000); the factor-B marginal means are (53.5000, 61.5000); and the grand mean is 57.5000.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 864.0000$, $SS_B = 384.0000$, and $SS_{AB} = 6.0000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 19(20) = 380.0000$. Therefore $F_A = 45.4737$, $F_B = 20.2105$, and $F_{AB} = 0.3158$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 48.0000)$ and $(B_1, 55.0000)$. Plot A_1 through $(B_0, 59.0000)$ and $(B_1, 68.0000)$. The two changes differ, so the lines are nonparallel and the plot displays the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V09: Spacing and retrieval cues

Reason before calculating, part (a)

The factor-A marginal means are (68.0000, 74.5000); the factor-B marginal means are (66.5000, 76.0000); and the grand mean is 71.2500.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 253.5000$, $SS_B = 541.5000$, and $SS_{AB} = 13.5000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 23(20) = 460.0000$. Therefore $F_A = 11.0217$, $F_B = 23.5435$, and $F_{AB} = 0.5870$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 64.0000)$ and $(B_1, 72.0000)$. Plot A_1 through $(B_0, 69.0000)$ and $(B_1, 80.0000)$. The two changes differ, so the lines are nonparallel and the plot displays the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

T08-A08-V10: Template and peer review

Reason before calculating, part (a)

The factor-A marginal means are (61.5000, 68.5000); the factor-B marginal means are (60.0000, 70.0000); and the grand mean is 65.0000.

Work through the calculation, part (b)

The A main effect compares its two marginal means, and the B main effect compares its two marginal means. The change across factor B differs between the two levels of factor A, so the cell pattern contains an interaction.

Work through the calculation, part (c)

The three null hypotheses are no population A main effect, no population B main effect, and no population $A \times B$ interaction.

Work through the calculation, part (d)

Balanced-design calculations give $SS_A = 294.0000$, $SS_B = 600.0000$, and $SS_{AB} = 6.0000$. With $df_A = df_B = df_{AB} = 1$ and $df_e = 4(6 - 1) = 20$, $SS_e = MS_e df_e = 20(20) = 400.0000$. Therefore $F_A = 14.7000$, $F_B = 30.0000$, and $F_{AB} = 0.3000$.

Interpret and check the result, part (e)

Plot the factor-A level A_0 through the coordinates $(B_0, 57.0000)$ and $(B_1, 66.0000)$. Plot A_1 through $(B_0, 63.0000)$ and $(B_1, 74.0000)$. The two changes differ, so the lines are nonparallel and the plot displays the interaction. Main effects summarize margins, while interaction asks whether one factor's pattern changes across the other factor.

A09: Fixed and Random Factors, Variance Components, and the ICC

T08-A09-V01: Libraries sampled from a regional population

Reason before calculating

The library levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, three deliberately chosen interface designs names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of library levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (18 - 6)/5 = 2.4000$ and $\hat{\sigma}_e^2 = MS_e = 6.0000$.

Interpret and check the result

Hence $ICC = 2.4000/[2.4000 + 6.0000] = 0.2857$.

The model attributes about 28.6% of its variance to differences among random library levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V02: Schools sampled from a district

Reason before calculating

The school levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, two named teaching programs names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of school levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (20 - 5)/6 = 2.5000$ and $\hat{\sigma}_e^2 = MS_e = 5.0000$.

Interpret and check the result

Hence $ICC = 2.5000/[2.5000 + 5.0000] = 0.3333$.

The model attributes about 33.3% of its variance to differences among random school levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V03: Interviewers sampled from a trained pool

Reason before calculating

The interviewer levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, three fixed questionnaire versions names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of interviewer levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (15 - 7)/4 = 2.0000$ and $\hat{\sigma}_e^2 = MS_e = 7.0000$.

Interpret and check the result

Hence $ICC = 2.0000/[2.0000 + 7.0000] = 0.2222$.

The model attributes about 22.2% of its variance to differences among random interviewer levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V04: Neighborhoods sampled from a city

Reason before calculating

The neighborhood levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, two selected outreach messages names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of neighborhood levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (24 - 8)/8 = 2.0000$ and $\hat{\sigma}_e^2 = MS_e = 8.0000$.

Interpret and check the result

Hence $ICC = 2.0000/[2.0000 + 8.0000] = 0.2000$.

The model attributes about 20.0% of its variance to differences among random neighborhood levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V05: Museum guides sampled from the staff roster

Reason before calculating

The guide levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, four fixed tour scripts names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of guide levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (21 - 6)/5 = 3.0000$ and $\hat{\sigma}_e^2 = MS_e = 6.0000$.

Interpret and check the result

Hence $ICC = 3.0000/[3.0000 + 6.0000] = 0.3333$.

The model attributes about 33.3% of its variance to differences among random guide levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V06: Archive boxes sampled from a collection

Reason before calculating

The archive box levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, three chosen scanning settings names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of archive box levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (19 - 5)/7 = 2.0000$ and $\hat{\sigma}_e^2 = MS_e = 5.0000$.

Interpret and check the result

Hence $ICC = 2.0000/[2.0000 + 5.0000] = 0.2857$.

The model attributes about 28.6% of its variance to differences among random archive box levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V07: Tutorial groups sampled from a program

Reason before calculating

The tutorial group levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, two fixed practice schedules names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of tutorial group levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (17 - 7)/6 = 1.6667$ and $\hat{\sigma}_e^2 = MS_e = 7.0000$.

Interpret and check the result

Hence $ICC = 1.6667/[1.6667 + 7.0000] = 0.1923$.

The model attributes about 19.2% of its variance to differences among random tutorial group levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V08: Routes sampled from a transport network

Reason before calculating

The route levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, three selected sign designs names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of route levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (27 - 9)/9 = 2.0000$ and $\hat{\sigma}_e^2 = MS_e = 9.0000$.

Interpret and check the result

Hence $ICC = 2.0000/[2.0000 + 9.0000] = 0.1818$.

The model attributes about 18.2% of its variance to differences among random route levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V09: Workshops sampled from an annual series

Reason before calculating

The workshop levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, two named facilitation formats names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of workshop levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (16 - 4)/5 = 2.4000$ and $\hat{\sigma}_e^2 = MS_e = 4.0000$.

Interpret and check the result

Hence $ICC = 2.4000/[2.4000 + 4.0000] = 0.3750$.

The model attributes about 37.5% of its variance to differences among random workshop levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

T08-A09-V10: Days sampled from a semester

Reason before calculating

The day levels were sampled to represent a wider population of possible levels, so repeating the study could draw new levels.

By contrast, three fixed reminder messages names the exact selected conditions of interest.

Work through the calculation

The random-factor target is variation across the population of day levels, not a list of pairwise differences among only the sampled labels.

In this balanced one-way setting, $\hat{\sigma}_A^2 = (MS_A - MS_e)/n = (22 - 6)/8 = 2.0000$ and $\hat{\sigma}_e^2 = MS_e = 6.0000$.

Interpret and check the result

Hence $ICC = 2.0000/[2.0000 + 6.0000] = 0.2500$.

The model attributes about 25.0% of its variance to differences among random day levels.

This equation depends on a balanced one-way random-factor structure; crossed, nested, repeated, or unbalanced designs can require different components and denominators.

A10: Repeated Measures, Sphericity, and Greenhouse-Geisser Correction

T08-A10-V01: Reading at three occasions

Reason before calculating, part (a)

Because all three marginal variances are 32.0000, each standard deviation is $\sqrt{32.0000}$. Substitution gives difference-score variances approximately (18, 19, 20); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 18 to 20, with largest-to-smallest ratio 1.1111. This pattern is fairly similar and offers descriptive reassurance, although it does not prove sphericity.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 84/2 = 42.0000$, $MS_{person} = 176/11 = 16.0000$, and $MS_e = 132/22 = 6.0000$. Thus $F_{condition} = 7.0000$ with p-value 0.0044, and $F_{person} = 16.0000/6.0000 = 2.6667$ with p-value 0.0241. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (16.0000 - 6.0000)/3 = 3.3333$, so $ICC = 3.3333/[3.3333 + 6.0000] = 0.3571$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.82(2) = 1.6400$ and $df_e^* = 0.82(22) = 18.0400$. Using the observed $F = 7.0000$ with those reference degrees of freedom gives corrected p-value 0.0080.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V02: Focus under three sound settings

Reason before calculating, part (a)

Because all three marginal variances are 43.0000, each standard deviation is $\sqrt{43.0000}$. Substitution gives difference-score variances approximately (12, 22, 31); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 12 to 31, with largest-to-smallest ratio 2.5833. This pattern is noticeably unequal and warns that the uncorrected reference may be unreliable.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 66/2 = 33.0000$, $MS_{person} = 154/11 = 14.0000$, and $MS_e = 110/22 = 5.0000$. Thus $F_{condition} = 6.6000$ with p-value 0.0057, and $F_{person} = 14.0000/5.0000 = 2.8000$ with p-value 0.0191. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (14.0000 - 5.0000)/3 = 3.0000$, so $ICC = 3.0000/[3.0000 + 5.0000] = 0.3750$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.74(2) = 1.4800$ and $df_e^* = 0.74(22) = 16.2800$. Using the observed $F = 6.6000$ with those reference degrees of freedom gives corrected p-value 0.0125.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V03: Recall across three delays

Reason before calculating, part (a)

Because all three marginal variances are 29.0000, each standard deviation is $\sqrt{29.0000}$. Substitution gives difference-score variances approximately (16, 17, 15); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 15 to 17, with largest-to-smallest ratio 1.1333. This pattern is fairly similar and offers descriptive reassurance, although it does not prove sphericity.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 72/2 = 36.0000$, $MS_{person} = 198/11 = 18.0000$, and $MS_e = 121/22 = 5.5000$. Thus $F_{condition} = 6.5455$ with p-value 0.0059, and $F_{person} = 18.0000/5.5000 = 3.2727$ with p-value 0.0086. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (18.0000 - 5.5000)/3 = 4.1667$, so $ICC = 4.1667/[4.1667 + 5.5000] = 0.4310$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.91(2) = 1.8200$ and $df_e^* = 0.91(22) = 20.0200$. Using the observed $F = 6.5455$ with those reference degrees of freedom gives corrected p-value 0.0077.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V04: Navigation across three route trials

Reason before calculating, part (a)

Because all three marginal variances are 50.0000, each standard deviation is $\sqrt{50.0000}$. Substitution gives difference-score variances approximately (10, 25, 38); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 10 to 38, with largest-to-smallest ratio 3.8000. This pattern is noticeably unequal and warns that the uncorrected reference may be unreliable.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 90/2 = 45.0000$, $MS_{person} = 165/11 = 15.0000$, and $MS_e = 143/22 = 6.5000$. Thus $F_{condition} = 6.9231$ with p-value 0.0047, and $F_{person} = 15.0000/6.5000 = 2.3077$ with p-value 0.0457. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (15.0000 - 6.5000)/3 = 2.8333$, so $ICC = 2.8333/[2.8333 + 6.5000] = 0.3036$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.68(2) = 1.3600$ and $df_e^* = 0.68(22) = 14.9600$. Using the observed $F = 6.9231$ with those reference degrees of freedom gives corrected p-value 0.0130.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V05: Confidence at three course points

Reason before calculating, part (a)

Because all three marginal variances are 33.0000, each standard deviation is $\sqrt{33.0000}$. Substitution gives difference-score variances approximately (20, 21, 19); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 19 to 21, with largest-to-smallest ratio 1.1053. This pattern is fairly similar and offers descriptive reassurance, although it does not prove sphericity.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 78/2 = 39.0000$, $MS_{person} = 187/11 = 17.0000$, and $MS_e = 126/22 = 5.7273$. Thus $F_{condition} = 6.8095$ with p-value 0.0050, and $F_{person} = 17.0000/5.7273 = 2.9683$ with p-value 0.0144. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (17.0000 - 5.7273)/3 = 3.7576$, so $ICC = 3.7576/[3.7576 + 5.7273] = 0.3962$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.88(2) = 1.7600$ and $df_e^* = 0.88(22) = 19.3600$. Using the observed $F = 6.8095$ with those reference degrees of freedom gives corrected p-value 0.0073.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V06: Accuracy under three interfaces

Reason before calculating, part (a)

Because all three marginal variances are 47.0000, each standard deviation is $\sqrt{47.0000}$. Substitution gives difference-score variances approximately (14, 28, 35); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 14 to 35, with largest-to-smallest ratio 2.5000. This pattern is noticeably unequal and warns that the uncorrected reference may be unreliable.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 81/2 = 40.5000$, $MS_{person} = 143/11 = 13.0000$, and $MS_e = 119/22 = 5.4091$. Thus $F_{condition} = 7.4874$ with p-value 0.0033, and $F_{person} = 13.0000/5.4091 = 2.4034$ with p-value 0.0385. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (13.0000 - 5.4091)/3 = 2.5303$, so $ICC = 2.5303/[2.5303 + 5.4091] = 0.3187$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.71(2) = 1.4200$ and $df_e^* = 0.71(22) = 15.6200$. Using the observed $F = 7.4874$ with those reference degrees of freedom gives corrected p-value 0.0092.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V07: Response time across three reminders

Reason before calculating, part (a)

Because all three marginal variances are 38.0000, each standard deviation is $\sqrt{38.0000}$. Substitution gives difference-score variances approximately (24, 23, 26); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pair-

wise condition difference are equal. The reconstructed values range from 23 to 26, with largest-to-smallest ratio 1.1304. This pattern is fairly similar and offers descriptive reassurance, although it does not prove sphericity.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 63/2 = 31.5000$, $MS_{person} = 209/11 = 19.0000$, and $MS_e = 138/22 = 6.2727$. Thus $F_{condition} = 5.0217$ with p-value 0.0160, and $F_{person} = 19.0000/6.2727 = 3.0290$ with p-value 0.0130. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (19.0000 - 6.2727)/3 = 4.2424$, so $ICC = 4.2424/[4.2424 + 6.2727] = 0.4035$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.95(2) = 1.9000$ and $df_e^* = 0.95(22) = 20.9000$. Using the observed $F = 5.0217$ with those reference degrees of freedom gives corrected p-value 0.0178.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V08: Comprehension across three formats

Reason before calculating, part (a)

Because all three marginal variances are 46.0000, each standard deviation is $\sqrt{46.0000}$. Substitution gives difference-score variances approximately (11, 20, 34); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 11 to 34, with largest-to-smallest ratio 3.0909. This pattern is noticeably unequal and warns that the uncorrected reference may be unreliable.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 96/2 = 48.0000$, $MS_{person} = 176/11 = 16.0000$, and $MS_e = 154/22 = 7.0000$. Thus $F_{condition} = 6.8571$ with p-value 0.0048, and $F_{person} = 16.0000/7.0000 = 2.2857$ with p-value 0.0476. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (16.0000 - 7.0000)/3 = 3.0000$, so $ICC = 3.0000/[3.0000 + 7.0000] = 0.3000$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.65(2) = 1.3000$ and $df_e^* = 0.65(22) = 14.3000$. Using the observed $F = 6.8571$ with those reference degrees of freedom gives corrected p-value 0.0147.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V09: Revision quality at three drafts

Reason before calculating, part (a)

Because all three marginal variances are 30.0000, each standard deviation is $\sqrt{30.0000}$. Substitution gives difference-score variances approximately (17, 18, 16); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 16 to 18, with largest-to-smallest ratio 1.1250. This pattern is fairly similar and offers descriptive reassurance, although it does not prove sphericity.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 75/2 = 37.5000$, $MS_{person} = 220/11 = 20.0000$, and $MS_e = 132/22 = 6.0000$. Thus $F_{condition} = 6.2500$ with p-value 0.0071, and $F_{person} = 20.0000/6.0000 = 3.3333$ with p-value 0.0078. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (20.0000 - 6.0000)/3 = 4.6667$, so $ICC = 4.6667/[4.6667 + 6.0000] = 0.4375$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.90(2) = 1.8000$ and $df_e^* = 0.90(22) = 19.8000$. Using the observed $F = 6.2500$ with those reference degrees of freedom gives corrected p-value 0.0094.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.

T08-A10-V10: Search skill at three practice points

Reason before calculating, part (a)

Because all three marginal variances are 52.0000, each standard deviation is $\sqrt{52.0000}$. Substitution gives difference-score variances approximately (13, 26, 40); the correlations were displayed to four decimals, so tiny reconstruction differences are rounding only. Sphericity asks whether the population variances of every pairwise condition difference are equal. The reconstructed values range from 13 to 40, with largest-to-smallest ratio 3.0769. This pattern is noticeably unequal and warns that the uncorrected reference may be unreliable.

Work through the calculation, part (b)

For conditions, $H_0 : \mu_1 = \mu_2 = \mu_3$; the alternative is that at least two condition means differ. For people, the random-person null is $H_0 : \sigma_{person}^2 = 0$ against positive between-person variation. The mean squares are $MS_{condition} = 87/2 = 43.5000$, $MS_{person} = 187/11 = 17.0000$, and $MS_e = 143/22 = 6.5000$. Thus $F_{condition} = 6.6923$ with p-value 0.0054, and $F_{person} = 17.0000/6.5000 = 2.6154$ with p-value 0.0264. The condition test rejects equal means at 5%; the person test supports between-person variation at 5%.

Work through the calculation, part (c)

$\hat{\sigma}_{person}^2 = (17.0000 - 6.5000)/3 = 3.5000$, so $ICC = 3.5000/[3.5000 + 6.5000] = 0.3500$. The ICC describes similarity among measurements from the same person under this model.

Work through the calculation, part (d)

Greenhouse-Geisser gives $df_{condition}^* = 0.70(2) = 1.4000$ and $df_e^* = 0.70(22) = 15.4000$. Using the observed $F = 6.6923$ with those reference degrees of freedom gives corrected p-value 0.0135.

Interpret and check the result, part (e)

The correction changes the reference degrees of freedom and therefore the p-value or critical value. It does not change the observed F , the fitted means, or the dependence among repeated rows. Measurements from the same person remain linked.