frontiers

Supplementary Material

Competition Between Audiovisual Correspondences Aids Understanding of Interactions Between Auditory and Visual Perception

Laura Getz*[1]

[1]Department of Psychological Sciences, University of San Diego, San Diego, CA,

*** Correspondence:**
Corresponding Author: Laura Getz
lgetz@sandiego.edu

1 Supplementary Results

The main text focuses on responses to the auditory dimension of pitch because the purpose of this project was to put the two visual dimensions in competition with each other to "compete" for pitch congruency. Therefore, the condition that is most relevant and where the results will be most straightforward is the pitch-relevant responding condition. However, I also asked participants to respond to the height of the shape and the size/brightness/sharpness/spatial frequency of the shape in different blocks of trials in order to compare the results; those findings are discussed here.

2 Size vs. Height Supplementary Results

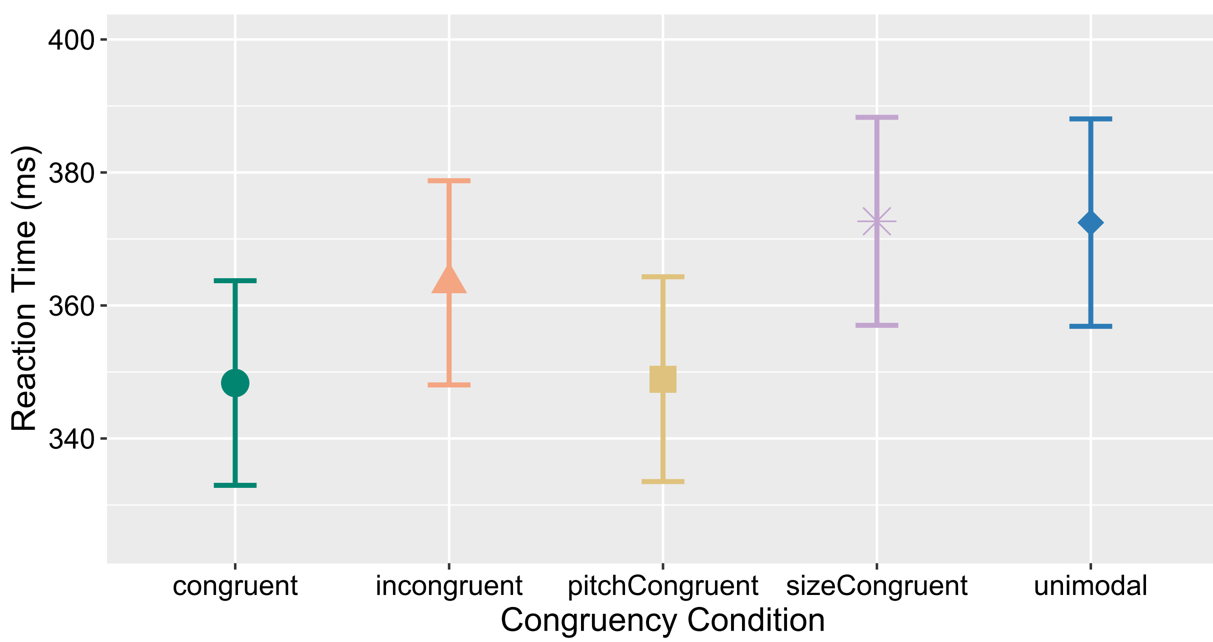
2.1 Height-Relevant Condition

We began with 12,330 observations from 56 participants; 707 trials were removed for short/long responses and 1,920 inaccurate responses were removed, leaving 10,410 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) and height difference (treated as continuous) as fixed effects and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 1** for all possible congruency condition comparisons). Although trending in the expected direction, there was not a significant congruency advantage or any other significant differences between the congruency conditions (see **Supplementary Figure 1)**. However, there was a significant effect of height difference (b = 93.85, SE = 0.06, t = -9.69, p < 0.001), meaning participants were slower to respond to height differences that were closer in pixels to the comparison circle.

Supplementary Table 1

Size experiment; height-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Incongruent – congruent	15.07	9.45	1.59	.501
Pitch congruent – congruent	0.58	9.52	0.06	1.00
Size congruent – congruent	24.31	9.92	2.45	.102
Unimodal – congruent	24.12	9.85	2.45	.103
Pitch congruent – incongruent	-14.49	9.47	-1.53	.543
Size congruent – incongruent	9.25	9.97	0.94	.882
Unimodal – incongruent	9.05	9.80	0.92	.888
Size congruent – pitch congruent	23.74	9.94	2.39	.118
Unimodal – pitch congruent	23.54	9.87	2.39	.119
Unimodal – size congruent	-0.20	10.24	-0.02	1.00



Supplementary Figure 1. Size experiment; Height-relevant task. Reaction times for the five congruency conditions. Participants responded faster (though not significantly) in the congruent and pitch congruent conditions than in the size congruent, incongruent, and unimodal conditions. Error bars represent standard error.

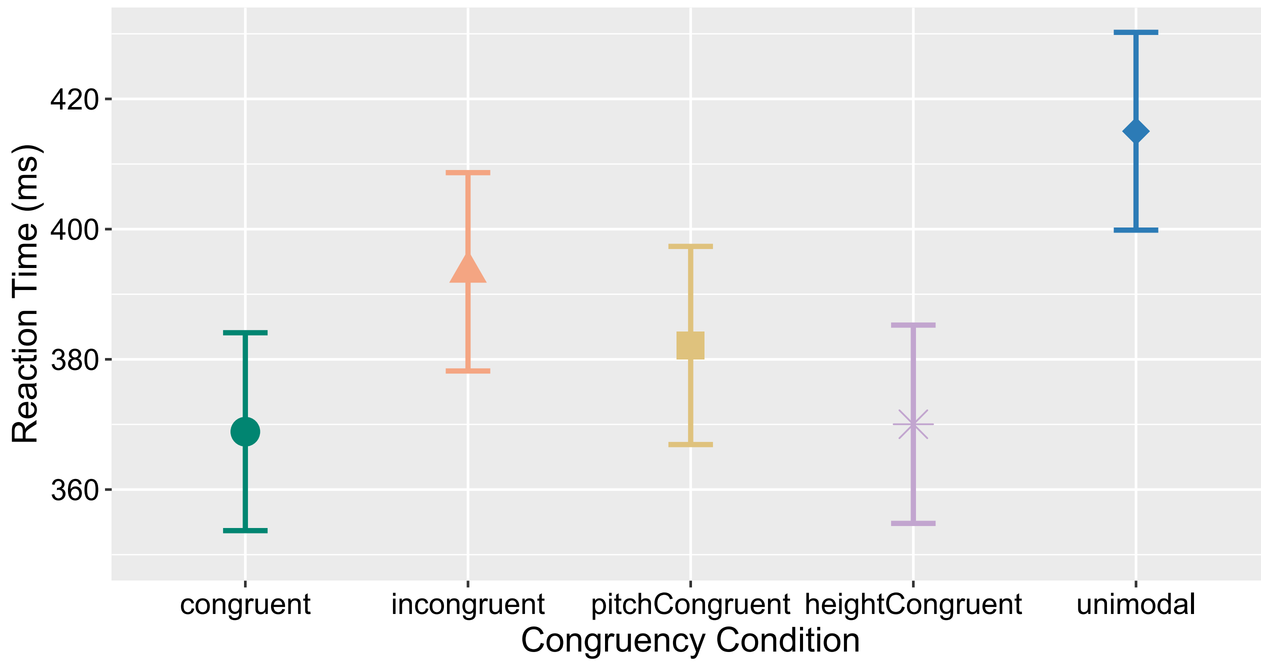
2.2 Size-Relevant Condition

We began with 11,760 observations from 56 participants; 336 trials were removed for short/long responses and 588 inaccurate responses were removed, leaving 10,836 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) and size difference (treated as continuous) as fixed effects and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 2** for all possible congruency condition comparisons). As shown in **Supplementary Figure 2**, we found a marginal congruency advantage ($b = 24.56$, $p = .072$). We also found a significant effect of size difference ($b = -69.94$, $SE = 3.69$, $t = -18.94$, $p < .001$), meaning participants were slower to respond to circles that were closer in size to the comparison circle.

Supplementary Table 2

Size experiment; size-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Height-congruent – congruent	1.14	9.47	0.12	.999
Incongruent – congruent	**24.56**	9.48	2.59	.072
Pitch congruent – congruent	13.24	9.45	1.40	.627
Unimodal – congruent	46.15	9.40	4.91	< .001
Incongruent – height congruent	23.41	9.53	2.46	.100
Pitch congruent – height congruent	12.10	9.50	1.27	.708
Unimodal – height congruent	45.01	9.46	4.76	< .001
Pitch congruent – incongruent	-11.32	9.51	-1.19	.757
Unimodal – incongruent	21.59	9.46	2.28	.151
Unimodal – pitch congruent	32.91	9.44	3.49	.004



Supplementary Figure 2. Size experiment; Size-relevant task. Reaction times for the five congruency conditions. Participants responded significantly slower in the unimodal condition than the congruent, pitch congruent, and height congruent conditions. Error bars represent standard error.

2.3 Size Discussion

Although trending in the expected direction, there was no congruency advantage when height was the relevant dimension and no difference between the pitch congruent and size congruent conditions. When size was the relevant dimension, there was a marginal congruency advantage, but no difference between the pitch congruent and height congruent conditions. These results show the same general trend as the pitch-relevant responses, but are not significant when responding to the visual dimensions instead.

3 Sharp vs. Height Supplementary Results

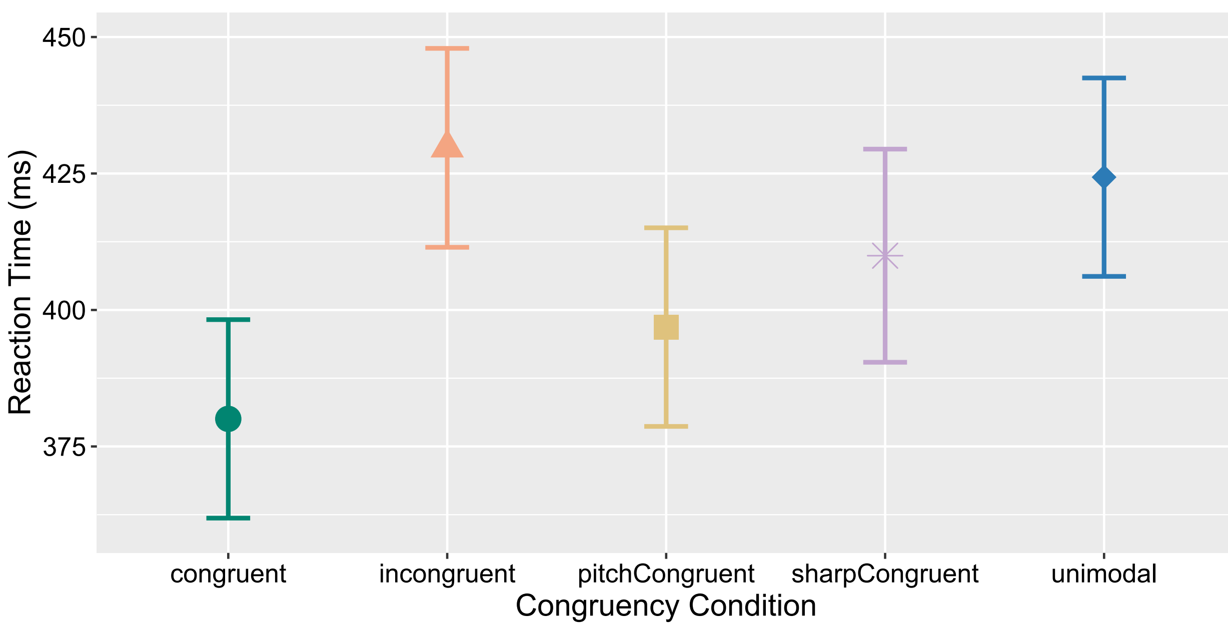
3.1 Height-Relevant Condition

We began with 12,960 observations from 53 participants; 557 trials were removed for short/long responses and 1,868 inaccurate responses were removed, leaving 10,535 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) as fixed effects and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 3** for all possible congruency condition comparisons). As shown in **Supplementary Figure 3**, we found a significant congruency advantage (b = 49.65, p < .001). We also found that participants were faster to respond in the pitch congruent condition than the both incongruent condition (b = -32.84, p = .013), whereas the sharp congruent condition did not significantly differ from either the both congruent (b = 29.89, p = .117) or both incongruent (b = -19.76, p = .511) conditions.

Supplementary Table 3

Sharp experiment; height-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Incongruent – congruent	49.65	10.33	4.81	< .001
Pitch congruent – congruent	16.81	10.29	1.64	.472
Sharp congruent – congruent	29.89	12.50	2.39	.117
Unimodal – congruent	44.27	10.26	4.32	< .001
Pitch congruent – incongruent	-32.84	10.35	-3.17	.013
Sharp congruent – incongruent	-19.76	12.55	-1.57	.511
Unimodal – incongruent	-5.38	10.31	-0.52	.985
Sharp congruent – pitch congruent	13.08	12.52	1.05	.833
Unimodal – pitch congruent	27.46	10.27	2.67	.057
Unimodal – sharp congruent	14.38	12.49	1.15	.777



Supplementary Figure 3. Sharp experiment; Height-relevant task. Reaction times for the five congruency conditions. Participants responded faster in the congruent and pitch congruent conditions than in the incongruent and unimodal conditions. Error bars represent standard error.

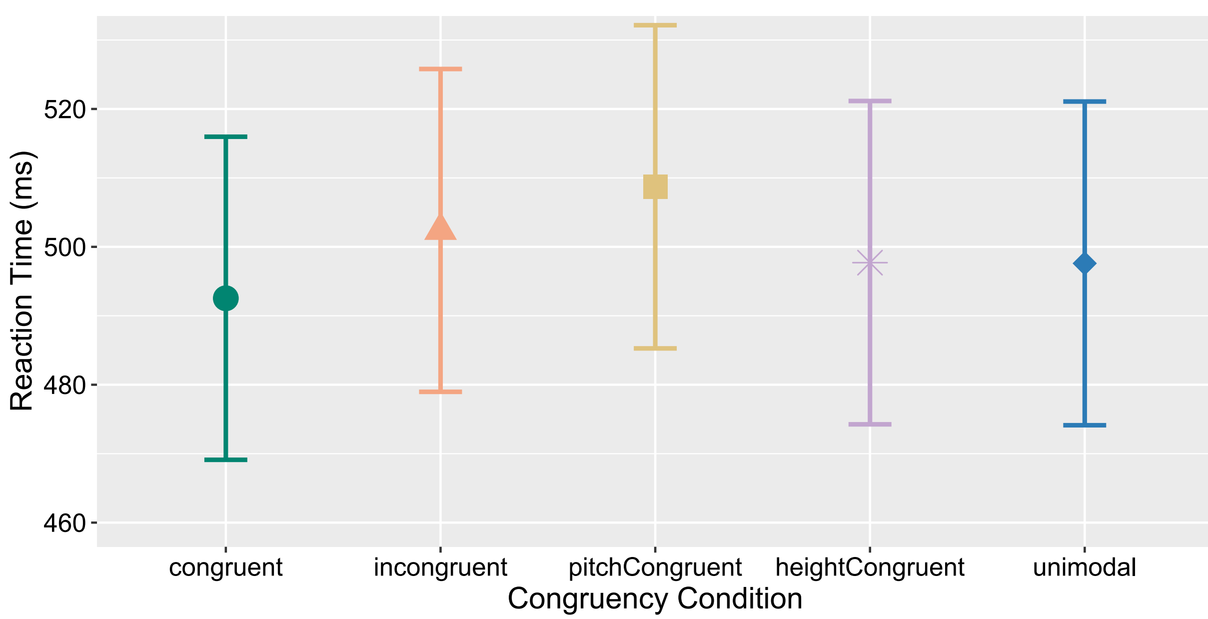
3.2 Sharp-Relevant Condition

We began with 12,952 observations from 53 participants; 569 trials were removed for short/long responses and 671 inaccurate responses were removed, leaving 11,712 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) as a fixed effect and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 4** for all possible congruency condition comparisons). There was not a significant congruency advantage or any other significant differences between the congruency conditions (see **Supplementary Figure 4)**.

Supplementary Table 4

Sharp experiment; sharp-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Height-congruent – congruent	5.17	11.94	0.43	.993
Incongruent – congruent	9.84	11.88	0.83	.922
Pitch congruent – congruent	16.17	11.92	1.36	.656
Unimodal – congruent	5.06	11.99	0.42	.993
Incongruent – height congruent	4.68	11.91	0.39	.995
Pitch congruent – height congruent	11.00	11.96	0.92	.889
Unimodal – height congruent	-0.10	12.02	-0.01	1.00
Pitch congruent – incongruent	6.33	11.90	0.54	.984
Unimodal – incongruent	-4.78	11.97	-0.40	.995
Unimodal – pitch congruent	-11.11	12.01	-0.93	.888



Supplementary Figure 4. Sharp experiment; Sharp-relevant task. Reaction times for the five congruency conditions showed no significant differences. Error bars represent standard error.

3.3 Sharp Discussion

When height was the relevant dimension, there was a significant congruency advantage and participants were faster to respond in the pitch congruent (sharp incongruent) condition than in the both incongruent condition, again showing the importance of the pitch-height congruency over other visual dimensions. However, there were no differences between conditions when sharp was the relevant dimension, again highlighting that congruency effects are stronger when responding to the pitch dimension.

4 Spatial Frequency vs. Height Supplementary Results

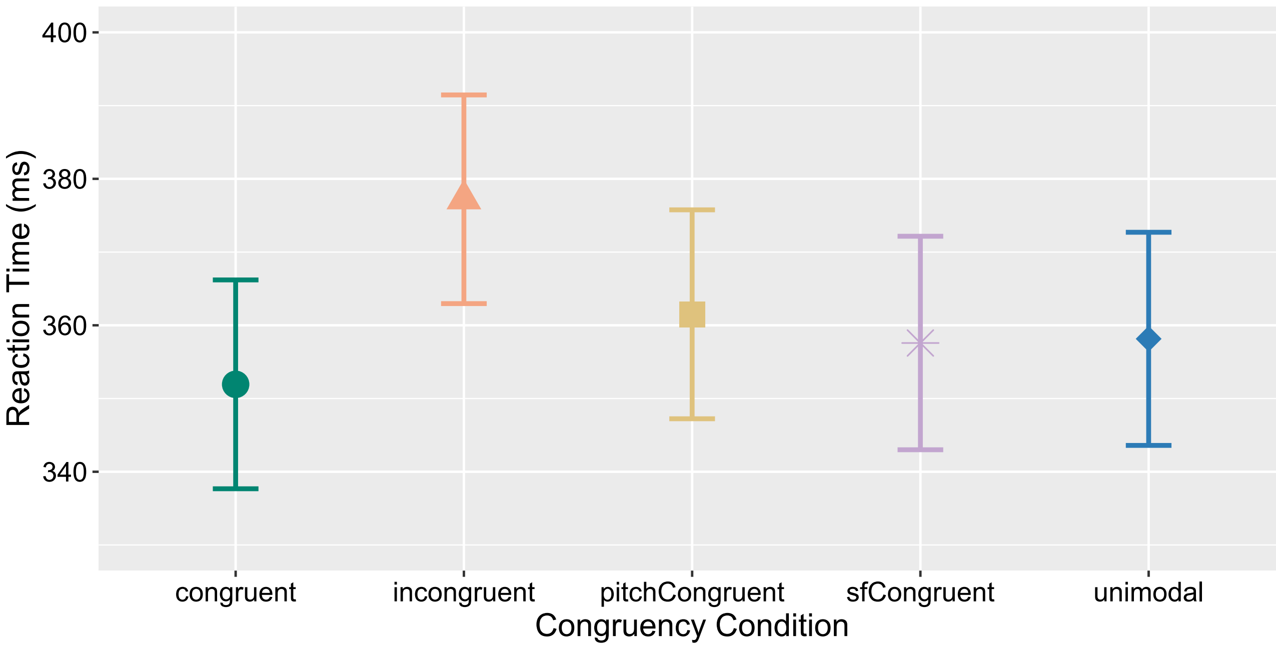
4.1 Height-Relevant Condition

We began with 12,610 observations from 55 participants; 830 trials were removed for short/long responses and 1,338 inaccurate responses were removed, leaving 11,272 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) and height difference (treated as continuous) as fixed effects and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 5** for all possible congruency condition comparisons). As shown in **Supplementary Figure 5**, we found a marginally significant congruency advantage ($b = 25.27$, $p = .093$). We also found a significant effect of height difference ($b = 75.36$, $SE = 0.07$, $t = -8.68$, $p < 0.001$), meaning participants were slower to respond to height differences that were closer in pixels to the comparison circle.

Supplementary Table 5

Spatial frequency experiment; height-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Incongruent – congruent	25.27	10.15	2.49	.093
Pitch congruent – congruent	9.56	10.18	0.94	.881
Spatial frequency congruent – congruent	5.64	10.62	0.53	.984
Unimodal – congruent	6.21	10.58	0.59	.977
Pitch congruent – incongruent	-15.71	10.17	-1.55	.533
Spatial frequency congruent – incongruent	-19.63	10.61	-1.85	.344
Unimodal – incongruent	-19.06	10.57	-1.80	.371
Spatial frequency congruent – pitch congruent	-3.92	10.63	-0.37	.996
Unimodal – pitch congruent	-3.35	10.59	-0.32	.998
Unimodal – spatial frequency congruent	0.57	11.01	0.05	1.00



Supplementary Figure 5. Spatial frequency experiment; Height-relevant task. Reaction times for the five congruency conditions. Participants were only marginally faster to respond in the congruent condition than in the incongruent condition. Error bars represent standard error.

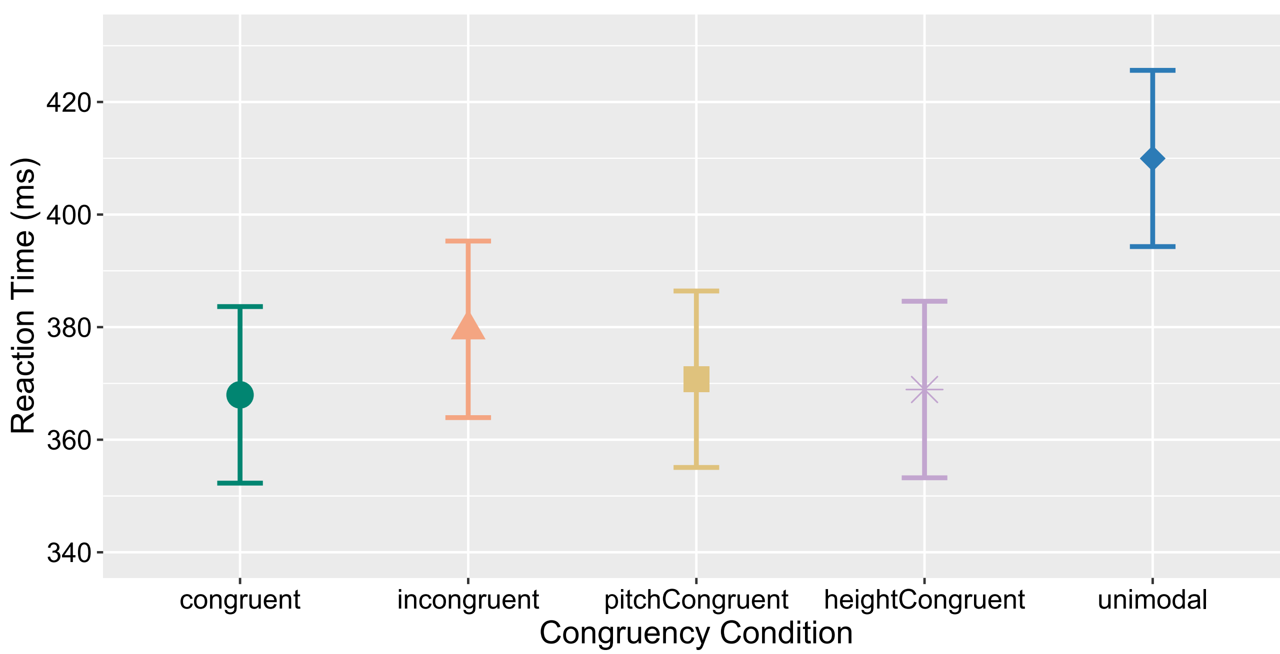
4.2 Spatial Frequency-Relevant Condition

We began with 13,920 observations from 57 participants; 413 trials were removed for short/long responses and 503 inaccurate responses were removed, leaving 13,004 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) and width difference (treated as continuous) as fixed effects and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 6** for all possible congruency condition comparisons). As shown in **Supplementary Figure 6**, the unimodal condition was significantly slower than each audiovisual condition, but there was no congruency advantage (b = 11.64, p = .651) or differences between any other audiovisual congruency condition. However, we did find a significant effect of width difference (b = -64.92, SE = 3.30, t = -19.65, p < .001), meaning participants were slower to respond to circles whose striping was closer in spatial frequency to the comparison circle striping.

Supplementary Table 6

Spatial Frequency experiment; width-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Height-congruent – congruent	0.95	8.54	0.11	.999
Incongruent – congruent	11.64	8.54	1.36	.651
Pitch congruent – congruent	2.78	8.52	0.33	.998
Unimodal – congruent	42.00	8.49	4.95	< .001
Incongruent – height congruent	10.69	8.54	1.25	.720
Pitch congruent – height congruent	1.83	8.52	0.22	.999
Unimodal – height congruent	41.05	8.49	4.84	< .001
Pitch congruent – incongruent	-8.87	8.52	-1.04	.837
Unimodal – incongruent	30.36	8.49	3.58	.004
Unimodal – pitch congruent	39.22	8.47	4.63	< .001



Supplementary Figure 6. Spatial frequency experiment; Width-relevant task. Reaction times for the five congruency conditions. Participants were slower in the unimodal condition than any of the audiovisual congruency condition, which did not differ from each other. Error bars represent standard error.

4.3 Spatial Frequency Discussion

We found a marginal congruency advantage when height was the relevant dimension, but no differences between pitch congruent and spatial frequency congruent conditions. We also found no significant differences between audiovisual conditions when spatial frequency was the relevant condition. As with the other visual dimensions, we found little of interest here compared to the pitch relevant condition reported in the main text.

5 Bright vs. Height Supplementary Results

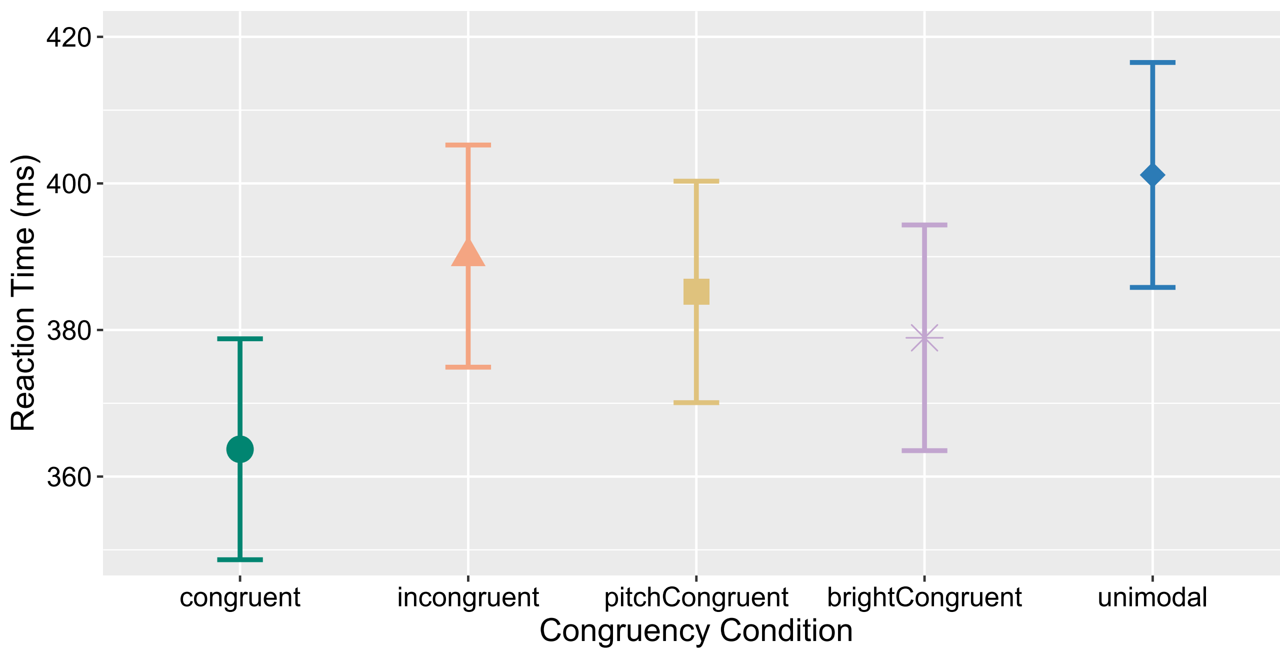
5.1 Height-Relevant Condition

We began with 14,100 observations from 58 participants; 741 trials were removed for short/long responses and 1,817 inaccurate responses were removed, leaving 11,542 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) and height difference (treated as continuous) as fixed effects and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 7** for all possible congruency condition comparisons). As shown in **Supplementary Figure 7**, we found a marginal congruency advantage (b = 26.36, p = .080), but no other significant differences between the audiovisual congruency conditions. We did find a significant effect of height difference (b = 77.31, SE = 0.07, t = -8.79, p < 0.001), meaning participants were slower to respond to height differences that were closer in pixels to the comparison circle.

Supplementary Table 7

Bright experiment; height-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Congruent – bright congruent	-15.21	10.71	-1.42	.615
Incongruent – bright congruent	11.15	10.81	1.03	.841
Pitch congruent – bright congruent	6.26	10.77	0.58	.978
Unimodal – bright congruent	22.22	11.09	2.00	.264
Incongruent – congruent	26.36	10.34	2.55	.080
Pitch congruent – congruent	21.47	10.26	2.09	.223
Unimodal – congruent	37.43	10.64	3.52	.004
Pitch congruent – incongruent	-4.89	10.37	-0.47	.990
Unimodal – incongruent	11.07	10.75	1.03	.841
Unimodal – pitch congruent	15.96	10.68	1.49	.566



Supplementary Figure 7. Bright experiment; Height-relevant task. Reaction times for the five congruency conditions. Participants were only marginally faster to respond in the congruent condition than in the incongruent condition. Error bars represent standard error.

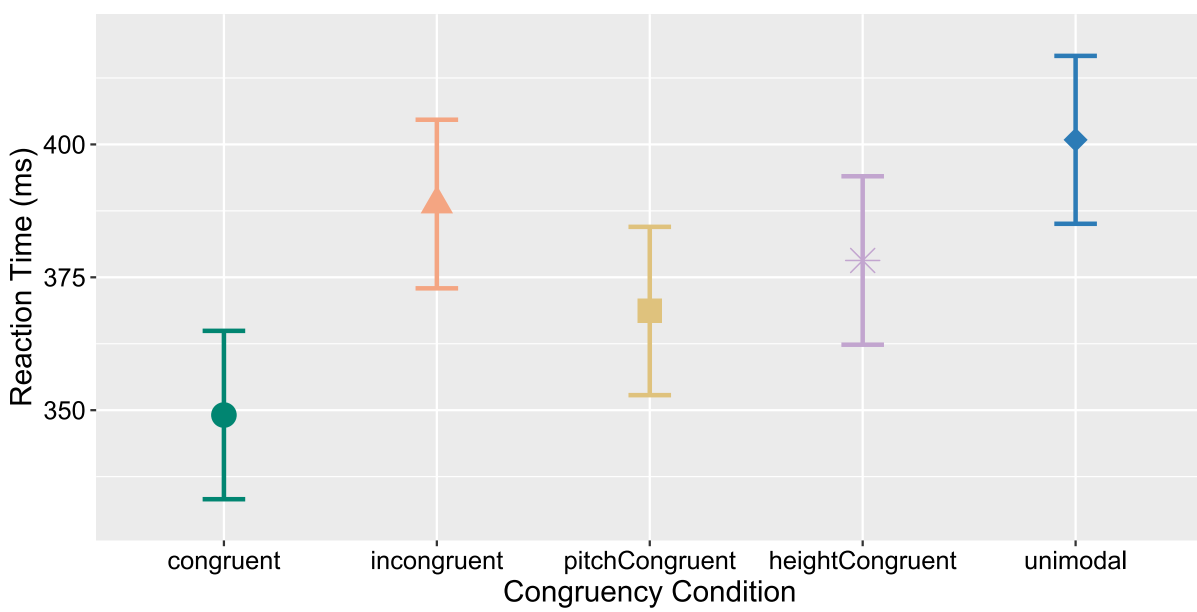
5.2 Bright-Relevant Condition

We began with 14,160 observations from 58 participants; 599 trials were removed for short/long responses and 405 inaccurate responses were removed, leaving 13,156 trials. We modeled the RT data with linear mixed-effects models (LMMs) with congruency (5 levels) and brightness difference (treated as continuous) as fixed effects and subject-by-subject variations as a random effect. A Tukey test for multiple comparisons designed for linear mixed effects models was performed on the congruency levels (see **Supplementary Table 8** for all possible congruency condition comparisons). As shown in **Supplementary Figure 8**, we found a significant congruency advantage (b = 39.70, p < .001) and found that the height congruent condition was significantly slower than the both congruent condition (b = 29.08, p = .015). We also found a significant effect of bright difference (b = 77.31, SE = 3.59, t = -7.77, p < 0.001), meaning participants were slower to respond to brightness differences that were closer in RBG scale to the comparison circle's brightness.

Supplementary Table 8

Bright experiment; bright-relevant condition. Tukey's post-hoc test for multiple comparisons of RT values across congruency conditions.

Comparison	*Estimate*	*SE*	*z*	*p*
Height-congruent – congruent	29.08	9.29	3.13	.015
Incongruent – congruent	39.70	9.31	4.26	< .001
Pitch congruent – congruent	19.57	9.26	2.11	.214
Unimodal – congruent	51.78	9.20	5.63	< .001
Incongruent – height congruent	10.62	9.33	1.14	.786
Pitch congruent – height congruent	-9.51	9.28	-1.02	.844
Unimodal – height congruent	22.70	9.22	2.46	.099
Pitch congruent – incongruent	-20.13	9.30	-2.16	.194
Unimodal – incongruent	12.08	9.24	1.31	.687
Unimodal – pitch congruent	32.20	9.19	3.50	.004



Supplementary Figure 8. Bright experiment; Bright-relevant task. Reaction times for the five congruency conditions. Participants were faster to respond in the congruent condition than in the incongruent condition. Error bars represent standard error.

5.3 Bright Discussion

We found a marginal congruency advantage when height was the relevant dimension, but no differences between pitch congruent and bright congruent conditions. When brightness was the relevant dimension, there was a significant congruency advantage, but there were no noticeable differences between the pitch congruent and height congruent conditions. So again, although the general congruency advantage exists here, the results are not as clear as in the pitch-relevant responding condition reported in the main text.