Table S1. Estimates of effects of Chord Quality on ratings of default volume seventh chords for Experiment 1.

	Happy vs. Bittersweet			Sad vs. Bittersweet		
Variable	Estimate	SE	p	Estimate	SE	p
Intercept	-.428	.135	.002	-.752	.138	< .001
Chord Quality	-.203	.090	.024	0.105	.098	.288

The significant negative intercept estimates suggest that, when seventh chords were presented with equal volume to all of the notes, participants were more likely to label the chord bittersweet. However, there was a difference between major and minor seventh chords; minor seventh chords were more likely to be labeled happy than major seventh chords.

Table S2. Estimates of effects of Chord Quality on ratings of default volume seventh chords for Experiment 2.

	Happy vs. Bittersweet			Sad vs. Bittersweet		
Variable	Estimate	SE	p	Estimate	SE	p
Intercept	-.071	.157	.654	-.994	.164	< .001
Chord Quality	-.190	.097	.050	0.163	.122	.180

While in Experiment 1 the intercepts for both happy vs. bittersweet and sad vs. bittersweet comparisons indicated a higher likelihood of a seventh chord being labeled as bittersweet, this was not the case in Experiment 2 where unmanipulated seventh chords were just as likely to be labeled happy (although in both experiments, unmanipulated seventh chords were less likely to be labeled sad). In addition, an effect of Chord Quality suggests that minor seventh chords may have been more likely to be labeled as happy in comparison to major sevenths, as in Experiment 1.

Consistency Analyses

In Experiments 1 and 2, participants completed two blocks of chord judgments, separated by attention checks. Given the 3AFC design implemented in Experiments 1 and 2 (with response options of happy, sad, or bittersweet), we assessed participants' consistency in labeling the chords from the first to the second block. For example, if a participant labeled a particular chord as 'happy' in both the first and second block, this would receive a consistency value of 1. In contrast, if a participant labeled a particular chord as 'happy' in the first block and 'sad' or 'bittersweet' in the second block, this would receive a consistency score of 0. For each participant, a mean consistency score was calculated and then the grand mean of all participants was assessed against 33.3% using a one-sample *t*-test, as this represents chance performance given the 3AFC paradigm.

In Experiment 1, we observed mean participant consistency of 50.6%, which was well above chance with a large effect size, $t(91) = 11.87$, $p < .001$, $d = 1.24$. In Experiment 2, we observed participant consistency of 51.0%, which was well above chance with a large effect size, $t(74) = 11.79$, $p < .001$, $d = 1.36$. In sum, these additional analyses suggest that participants were statistically consistent in their responses across blocks.

Table S3 Wald χ2 tests assessing statistical significance of the fixed effects of Chord Quality and Volume on ratings of seventh chords for Experiment 3, including Pitch Height as a covariate.

Dependent measure	Predictor	χ2	*p*
Happy	Chord Quality	.06	.85
	Quiet Root	22.52	< .001
	Quiet Seventh	5.95	.064
	Pitch Height	296.87	< .001
	Chord Quality x Quiet Root	10.28	.013

	Chord Quality x Quiet Seventh	9.38	.018
Sad	Chord Quality	8.16	.011
	Quiet Root	8.98	.021
	Quiet Seventh	12.64	.006
	Pitch Height	468.29	< .001
	Chord Quality x Quiet Root	4.78	.110
	Chord Quality x Quiet Seventh	29.37	< .001
Bittersweet	Chord Quality	.45	.573
	Quiet Root	16.79	< .001
	Quiet Seventh	6.17	.061
	Pitch Height	7.65	.033
	Chord Quality x Quiet Root	7.80	.032
	Chord Quality x Quiet Seventh	6.20	.061
Nostalgic	Chord Quality	8.63	.010
	Quiet Root	10.20	.013
	Quiet Seventh	20.93	< .001
	Pitch Height	8.75	.022
	Chord Quality x Quiet Root	.89	.700
	Chord Quality x Quiet Seventh	.10	.952

Note: p-values are corrected for multiple comparisons using FDR.