DURMS
Durham Undergraduate Research in Music & Science

Musicianship, Tonality, and Subjective Time: Musical Sophistication Predicts Retrospective Time Judgements for Tonal (but not Atonal) Chord Progressions

Matt Moore, Joy Sutcliffe, & Nico Wood-Olivan
Durham University

ABSTRACT

The human experience of time is highly subjective, particularly during music listening. We use music to make time feel like it is passing more quickly, or to distort our memory of how much time has passed, but how does music alter our temporal experience so profoundly? We aimed to replicate Firmino and Bueno's (2008) observation that tonal complexity - in the form of increasingly distant modulations - was accompanied by a commensurate underestimation on a time interval reproduction task. We extended their study with an atonal stimulus and collected musical sophistication scores to investigate effects of expertise on time judgements. Here we show that time underestimation for tonally complex stimuli was much smaller than previously reported and was heavily influenced by participants' musical sophistication. The original study proposed an Expected Development Fraction (EDF) model to explain their results, based on a perceived incompatibility with existing models. However, we report no significant differences in estimations between any of the conditions; instead, expertise was a much more meaningful predictor of time judgements. Our findings appear compatible with prevailing models of retrospective timing. We interpret musical sophistication as a mediating factor for event segmentation, which in turn affected how many chords were recalled during the timing task. Musical sophistication correlated positively with time judgements only for stimuli that followed rules of functional harmony; in the absence of harmonic syntax, it provided no benefits for segmentation. The role of expertise in time judgements has not been properly explored. We provide novel and compelling evidence that expertise significantly affects retrospective timing: particularly expertise for syntactical rules and the chunking of perceptual information.

1. INTRODUCTION

How long was the last song that you heard? How do you know? We take advantage of the fallibility of human time perception every time we turn on the radio to make a long car journey seem shorter, or to minimise the drudgery of everyday housework. But what are the mechanisms underpinning our subjective experience of time, and how does music exploit them so readily?

Cognitive psychologists distinguish two distinct forms of time perception: prospective and retrospective (Block, 1990). Prospective timing refers to the ongoing experience of duration, where the observer is explicitly aware of the passage of time as it unfolds. Conversely, retrospective timing concerns the memory of duration - how much time we feel to have passed while we were not consciously attending to it. Block (1992) observed that the two paradigms appear to be qualitatively different and may react in opposite ways to increasing processing demands, with prospective time estimations becoming shorter and retrospective estimations becoming longer. These kinds of findings are not universal, however, as in the case of Brown and Stubbs (1992) who demonstrated similar effects of attentional load on timing under both prospective and retrospective conditions. Nonetheless, the literature on human timing retains a clear differentiation between the two paradigms, each with their own distinct cognitive models (Block & Zakay, 1997). This paper is concerned exclusively with the retrospective timing paradigm; specifically, in relation to music listening.

There have been several attempts to describe the nature of retrospective time estimation, though certainly not as many as for its prospective counterpart. Ornstein's (1969) storage size model was based on the observation that time judgements were longer for intervals during which participants were required to process more complex information, or a larger quantity of it. However, Block (1978) challenged this model with evidence that mere stimulus complexity is not sufficient to cause time overestimations. He found that when participants were presented with sequences of visual patterns, it was not the complexity of individual patterns within the sequence that resulted in time overestimation, but the complexity of the sequence itself. This was explained in terms of a contextual change model (Block & Reed, 1978), in which retrospective time judgements are a function of the number of contextual changes people are able to recall from a time interval. Poynter (1983) elaborated on this with his change/segmentation model, which posits that duration judgements are a function of the quantity, discreteness, and memorability of experienced events; this in turn affects the brain's ability to effectively "chunk" (see Chase and Simon (1973)) and retrieve the events for time interval reconstruction.

The capacity for music to cause subjective time distortion has been the subject of growing scholarly attention, primarily for its applications in consumer psychology and retail contexts (see Bailey and Areni (2006), for review). Kellaris and Kent (1992) observed longer time estimates in the presence of positively valenced (major key) music in comparison to minor or atonal music, contrary to conventional wisdom (i.e., "time flies when you're having fun"). Likewise, Hul, Dube, and Chebat (1997) reported a positive correlation between time estimates and liking ratings for musical stimuli, and Gulas and Schewe (1994) reported longer time estimates for familiar rather than unfamiliar music. Bailey and Areni (2006) interpret these findings as supporting a discrete event model of timing; musical characteristics facilitating the retention of more events appear to result in longer estimations of time. These findings are in broad agreement with the models of Block and Reed (1978) and Poynter (1983). However, there are some studies presenting opposing evidence. Bueno, Firmino, and Engelman (2002) reported longer time estimates for music high in generalised complexity, despite the lower probability of event retention. Similarly, Yalch and Spangenberg (2000) demonstrated that customers believed they had shopped for a longer duration when exposed to unfamiliar background music compared to familiar foreground music. These findings are somewhat more challenging to explain within the prevailing models of subjective time estimation.

This study will attempt to replicate the findings of Firmino and Bueno (2008), who observed significant time underestimations for chord sequences that modulate suddenly and to distant keys. They propose the Expected Development Fraction (EDF) model to explain this dynamic. The EDF model effectively treats tonal modulation as an interruption to the expected temporal development of a harmonic progression, which is stored in semantic memory. The perceived duration of the stimulus, represented in implicit working memory, is shorter than the temporal expectation held in semantic memory, resulting in an underestimation on a retrospective time interval reproduction task.

We found elements of this model to be problematic. Firstly, it accounts for time distortion strictly in relation to tonal modulation, despite the fact that the same phenomenon occurs in response to many other musical characteristics; we believe it is unlikely that there would be radically different mechanisms for every musical characteristic producing the same outcome. Secondly, it attempts to apply elements of Newtonian classic mechanics to the entirely intangible concept of musical tonality. Whilst metaphors of gravitation are often used when discussing harmonic function, the authors appear to have taken this notion rather literally. Thirdly, and partially pursuant to the previous point, the EDF model is largely obfuscated behind dense mathematical language and perplexing diagrams of cognitive circuitry, which makes it difficult to rigorously assess the appropriateness of the model for what it attempts to predict.

For main our hypothesis, we determined that the EDF model would predict a drastic time underestimation for the atonal condition. Therefore, successful replication would entail a descending pattern of time estimations according to increasing tonal complexity. As a secondary hypothesis, we predicted that musical sophistication would interact with time judgements, but made no specific predictions concerning the nature of such an interaction due to the exploratory nature of this aspect of our study.

2. METHOD

Design. The experiment used a between-subjects design, which was a necessity of the retrospective timing paradigm. The independent variable was tonal condition (diatonic, modulating, or atonal) and the dependent variable was retrospective time judgement. Allocation to tonal condition groups was determined by Qualtrics' built-in randomisation function, although manual adjustment of group quotas was performed in order to balance actual group sizes in cases of survey non-completion (which would still count towards Qualtrics group totals).

Participants. We received 94 anonymous responses to the online Qualtrics survey, of which 40 contained useable data; the rest were either incomplete or had implausibly short response times on the reproduction task. Participants were recruited via opportunity sampling, primarily using word-of-mouth and social media. Since age and gender were not implicated in the effect size for any other studies utilising the EDF model, we opted not to collect this information to keep the online survey as short as possible. A selection of questions (see Appendix A) from the Goldsmiths Musical Sophistication Index (Gold-MSI) (Müllensiefen, Gingras, Musil, & Stewart, 2014) revealed that the sample ranged from 14 to 42 ($M = 26.85$, $SD = 8.20$) in terms of musical sophistication, with a bimodal distribution.

Materials. Each 20-second stimulus consisted of 29 chords, which were restricted to simple major and minor triads (see Appendix B). The tonal stimuli used harmonic progressions identical to Firmino and Bueno's (2008) original study, whilst the atonal stimulus was composed by applying the chord qualities (major/minor) from the tonal stimuli to a randomly generated tone row. Of the four chord progressions used in the original study, only the CC (no modulation) and CGb (sudden distant) progressions were used here. The CF (sudden close) and CEbGb (gradual distant) modulations were dropped; since our intent was simply to reproduce the effect itself, we chose to only use the progression that resulted in the largest time underestimation in the original study in order to maximise the likelihood of detecting an effect.

The chords in the original study were effectively made up of Shepard tones (Shepard, 1964) - complex tones comprising multiple octaves, which decrease in volume towards the extreme high and low frequencies. Firmino and Bueno (2016) have since demonstrated that the effect persists even for real modulating music, so using Shepard tones for this replication was deemed an unnecessary extra step. We instead used simple piano sounds, which would have sounded less artificial and more like "real" music. The chord progressions were created in Sibelius 7.5 and exported as .mp3 files to be most suitable for use in an online study.

Twelve items from the Gold-MSI were selected in order to obtain an approximate measure of musical sophistication whilst maximising participant completion rates. The questions covered the full range of Gold-MSI subscales, however, certain subscales (such as Musical Training and Perceptual Abilities) were more relevant to the present study than others due to the nature of the task, and thus comprised a greater share of the questions asked.

Procedure. Participants were instructed to complete the survey on a PC or laptop. This was primarily due to limitations of the Qualtrics platform, but carried the secondary benefit of ensuring the study was carried out in reasonably comparable environments (to the maximum extent permitted by an entirely online study). After providing informed consent, participants would hear one of the three possible 20s chord progressions. They were then asked to read the following instructions carefully:

> Your task is to reproduce the time duration of the chord sequence you just listened to.
>
> When you are ready, click next, then wait. Once you feel that the amount of time you have waited equals the length of the chord sequence you heard, click next again. There will be no visible timer, simply click next to begin, then click again to stop.

These instructions were adapted from Firmino and Bueno (2008) for both greater task clarity and better suitability for the online platform. Finally, participants completed the items selected from the Gold-MSI.

3. RESULTS

We used R (Version 3.6.0; R Core Team, 2018) and the R-packages afex (Version 0.23.0; Singmann, Bolker, Westfall, & Aust, 2019), car (Version 3.0.2; Fox & Weisberg, 2011), and multcomp (Version 1.4.10; Hothorn, Bretz, & Westfall, 2008) for all our analyses. This report was compiled using the R-packages knitr (Version 1.22; Xie, 2015), and papaja (Version 0.1.0.9842; Aust & Barth, 2018).

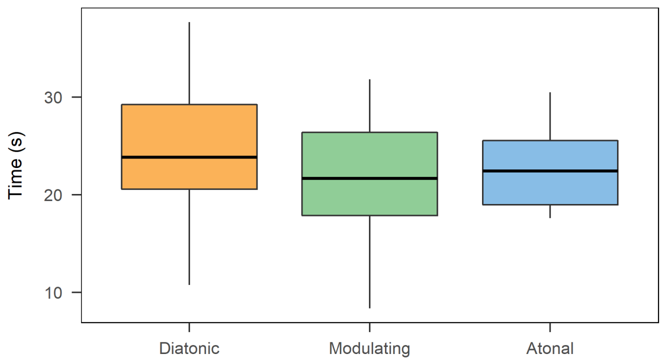

Figure 1. Retrospective time estimations by tonal condition

Initial box plots of time judgements by condition indicated a general trend as predicted, with increasing tonal complexity resulting in lower retrospective time estimations (see Figure 1). However, a one-way ANOVA showed no significant effect of tonal condition on retrospective time judgement (F(2, 37) = 0.36, MSE = 46.52, p = .701). Variance for the atonal condition appeared to be much lower than for the other two conditions; whilst Levene's test returned a non-significant p-value (p = .113), it was nonetheless deemed worthy of further investigation.

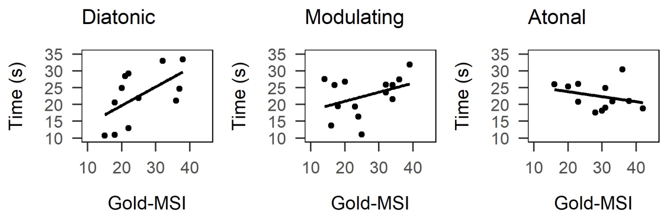

Figure 2. Scatterplots of Gold-MSI (General Sophistication) scores against retrospective time estimations

Plotting Gold-MSI (General Sophistication) scores against retrospective time judgements for each condition (see Figure 2) indicated correlations between Gold-MSI scores and time judgements that were confirmed to be significant or marginally significant in both diatonic (r(11) = .66, p = .015) and modulating (r(13) = .46, p = .086) conditions, but not in the atonal condition (r(10) = -.29, p = .369). Based on the observed correlations, we decided to pool the diatonic and modulating conditions into a single "tonal" group, and conducted an ANCOVA using general sophistication as the covariate. However, even controlling for general sophistication in this manner revealed no significant effect of tonal condition on retrospective time judgement (see Table 1).

Table 1. Analysis of Covariance in Retrospective Time Estimation by Tonal Condition and Gold-MSI (General Sophistication) Score

Effect	F	df_1	df_2	MSE	p	$\hat{\eta}^2_G$
Condition	0.47	1	24	42.48	.499	.019
General Sophistication	11.69	1	24	42.48	.002	.328
Condition × General Sophistication	0.92	1	24	42.48	.348	.037

As an alternative analysis, we took advantage of the bimodal distribution present in the general sophistication scores to conduct a 2 x 3 ANOVA (General Sophistication (High/Low) x Tonal Condition). This again confirmed an overall effect of general sophistication on time judgement, and appeared to show a significant effect of tonal condition and general sophistication on time judgement (F(2, 34) = 3.94, MSE = 35.09, p = .029, η^2G = .188), but post-hoc comparisons using the Tukey HSD test revealed only one statistically significant interaction (see Table 2 in Appendix C).

4. DISCUSSION

The aim of this study was to replicate the findings of Firmino and Bueno (2008) whilst testing their proposed EDF model with an additional atonal condition. We determined that the EDF model would predict a drastic underestimation of retrospective time judgements in the atonal condition.

Additionally, we collected a basic measure of musical sophistication using the Gold-MSI to look for possible effects of musical expertise, which was not measured in the original study or any related studies containing the EDF model.

Our initial results did show a trend in the predicted direction for the modulating condition (see Figure 1); however, this was much smaller than the effect reported by the authors of the original study. Plotting Gold-MSI (General Sophistication) scores against retrospective time judgements showed a positive relationship between the two for the diatonic and modulating conditions, but not for the atonal condition (see Figure 2). This strongly suggests that expertise mediates retrospective time judgements for tonal progressions. Moreover, the fact that there was no correlation between Gold-MSI scores and time judgements in the atonal condition indicates that this mediating effect is due to an implicit understanding of the rules governing harmonic progression. In the absence of discernible harmonic function, there was no difference in time judgements between musicians and non-musicians.

We believe our findings are incompatible with the EDF model for a number of reasons. Primarily, we did not observe the significant underestimation of time judgements present in the original study for either of the modulating or atonal conditions. Furthermore, the EDF model does not incorporate any effect of musical expertise into its predictions, which we have shown to explain a significant amount of the variance in time judgements between subjects.

Our findings can be interpreted more parsimoniously by returning to the models initially discarded by Firmino and Bueno (2008). They justified the need for a new model based on the fact that more tonally complex stimuli resulted in time underestimations, which could not be accounted for by models proposing that greater cognitive processing activity results in the stretching of subjective time. However, our results paint a broader picture: processing activity is mediated by an implicit understanding of harmonic syntax - the greater the understanding, the greater the amount of activity. This interpretation is entirely compatible with prevailing models of retrospective timing, particularly Poynter's (1983) segmentation model. Interpreting our results within the segmentation model, we argue that the number of changes (chords) retrieved is linearly related to the expertise of the listener via their ability to efficiently segment this specific type of perceptual information. This is consistent with Block (1974), who found that word strings blocked by category result in retrospective time overestimations compared to random word strings; in effect, retrieval of contextual changes was facilitated by linguistic expertise - an understanding of the semantic relationship between words in a string. We speculate that sudden tonal modulation simply represents lost information due to the unexpected violation of harmonic syntax, which would explain the slight tendency towards underestimation that we observed.

There are some limitations to the present study. Whilst the overall sample size was reasonably adequate, the ANCOVA for the pooled tonal results had a sample size of only 28. We believe that a larger sample size may have produced a more robust model to control for the effect of general sophistication, potentially leading to a statistically significant difference in means between diatonic and modulating conditions. This would replicate the findings of the original study, albeit with an alternative explanation. Indeed, we conjecture that the statistically significant findings of Firmino and Bueno (2008) were likely due to a sample that was relatively homogeneous in terms of musical sophistication, resulting in less overall variance in their results; given that our results did at least trend in the same direction as theirs, it seems reasonable to speculate that the lack of statistical significance was due to the confounding effect of our musically diverse sample. A further limitation to our study arises when contextualising our selection of Gold-MSI questions within the theory we used to explain our results. The Emotion subscale of the Gold-MSI was the most underrepresented, comprising only a single question in our survey - this was because our initial hypothesis did not anticipate any effect of emotion on retrospective time judgement using the EDF model. However, emotional arousal could easily be implicated in the successful encoding and recollection of discrete events (e.g., Carr and Rickard (2015)); as such, repeating our study in controlled conditions with an expanded selection of Gold-MSI questions would be a worthwhile endeavour.

To summarise, our study failed to replicate the findings of Firmino and Bueno (2008). We hypothesised that the addition of an atonal stimulus would result in a drastic time underestimation according to the EDF model; this was not the case. Instead, musical sophistication emerged as a far more suitable predictor of retrospective time judgements, but only for stimuli adhering to conventional harmonic syntax. We take this as evidence in favour of Poynter's (1983) segmentation model of retrospective timing on the basis that expertise appears to play a crucial role in the effective segmentation and recall of tonal chunks. Potential follow-up studies might seek to explore similar mediating effects of musical sophistication on retrospective timing across more generalised measures of musical complexity, such as event density and harmonic vocabulary. Ultimately, we demonstrate the need for further research specifically on the topic of task expertise in relation to retrospective timing, for which there appears to be no known literature.

REFERENCES

Aust, F., & Barth, M. (2018). papaja: Create APA manuscripts with R Markdown. Retrieved from https://github.com/crsh/papaja

Bailey, N., & Areni, C. S. (2006). When a few minutes sound like a lifetime: Does atmospheric music expand or contract perceived time? *Journal of Retailing, 82*(3), 189–202. doi:10.1016/j.jretai.2006.05.003

Block, R. A. (1974). Memory and the experience of duration in retrospect. *Memory & Cognition*, *2*(1), 153–160. doi:10.3758/BF03197508

Block, R. A. (1978). Remembered duration: Effects of event and sequence complexity. *Memory & Cognition, 6*(3), 320–326. doi:10.3758/BF03197462

Block, R. A. (Ed.). (1990). *Cognitive models of psychological time.* Hillsdale, N.J: L. Erlbaum Associates.

Block, R. A. (1992). Prospective and Retrospective Duration Judgment: The Role of Information Processing and Memory. In F. Macar, V. Pouthas, & W. J. Friedman (Eds.), *Time, Action and Cognition: Towards Bridging the Gap* (pp. 141–152). Dordrecht: Springer Netherlands. doi:10.1007/978-94-017-3536-0_16

Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 656–665. doi:10.1037/0278-7393.4.6.656

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review, 4*(2), 184–197. doi:10.3758/BF03209393

Brown, S. W., & Stubbs, D. A. (1992). Attention and Interference in Prospective and Retrospective Timing. *Perception*, *21*(4), 545–557. doi:10.1068/p210545

Bueno, J. L. O., Firmino, É. A., & Engelman, A. (2002). Influence of Generalized Complexity of a Musical Event on Subjective Time Estimation. *Perceptual and Motor Skills*, *94*(2), 541–547. doi:10.2466/pms.2002.94.2.541

Carr, S., & Rickard, N. (2015). The use of emotionally arousing music to enhance memory for subsequently presented images. *Psychology of Music, 44*. doi:10.1177/0305735615613846

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55–81. doi:10.1016/0010-0285(73)90004-2

Firmino, É. A., & Bueno, J. L. O. (2008). Tonal Modulation and Subjective Time. *Journal of New Music Research, 37*(4), 275–297. doi:10.1080/09298210802711652

Firmino, É. A., & Bueno, J. L. O. (2016). Interkey Distances Also Shorten Subjective Time Reproductions in Real Modulating Tonal Music. *Music Perception: An Interdisciplinary Journal, 33*(5), 613–630. doi:10.1525/mp.2016.33.5.613

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression (Second Edition)*. Thousand Oaks CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Gulas, C., & Schewe, C. (1994). Atmospheric Segmentation: Managing Store Image With Background Music. 1994 Ama Educators' Proceedings: Enhancing Knowledge Development in Marketing, 325–330. Retrieved from https://corescholar.libraries.wright.edu/marketing/17

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363.

Hul, M. K., Dube, L., & Chebat, J.-C. (1997). The impact of music on consumers' reactions to waiting for services. *Journal of Retailing, 73*(1), 87–104. doi:10.1016/S0022-4359(97)90016-6

Kellaris, J. J., & Kent, R. J. (1992). The Influence of Music on Consumers' Temporal Perceptions: Does Time Fly When You're Having Fun? *Journal of Consumer Psychology*, (4), 365–376. Retrieved from http://www.jstor.org/stable/1480665

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLOS ONE*, *9*(2), e89642. doi:10.1371/journal.pone.0089642

Ornstein, R. E. (1969). *On the experience of time.* Boulder, Colo: WestviewPress.

Poynter, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition, 11*(1), 77–82. doi:10.3758/BF03197664

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Shepard, R. N. (1964). Circularity in Judgments of Relative Pitch. *The Journal of the Acoustical Society of America, 36*(12), 2346–2353. doi:10.1121/1.1919362

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2019). Afex: Analysis of factorial experiments. Retrieved from https://CRAN.R-project.org/package=afex

Xie, Y. (2015). Dynamic documents with R and knitr (2nd ed.). Boca Raton, Florida: Chapman; Hall/CRC. Retrieved from https://yihui.name/knitr/

Yalch, R. F., & Spangenberg, E. R. (2000). The Effects of Music in a Retail Setting on Real and Perceived Shopping Times. *Journal of Business Research, 49*(2), 139–147. doi:10.1016/S0148-2963(99)00003-X

APPENDIX A

List of Items Selected from the Gold-MSI Alongside Relevant Subscales

Item	Subscale
I spend a lot of my free time doing music-related activities.	Active Engagement, General Sophistication
I can sing or play music from memory.	Singing Ability, General Sophistication
Pieces of music rarely evoke emotions for me.	Emotions
I can compare and discuss differences between two performances or versions of the same piece of music.	Perceptual Abilities, General Sophistication
I often read or search the internet for things related to music.	Active Engagement, General Sophistication
When I hear a music I can usually identity its genre.	Perceptual Abilities
I only need to hear a new tune once and I can sing it back hours later.	Singing Ability
I engaged in regular, daily practice of a musical instrument (including voice) for __ years.	Musical Training, General Sophistication
I have attended __ live music events as an audience member in the past twelve months.	Active Engagement
I have had __ years of formal training on a musical instrument (including voice) during my lifetime.	Musical Training
I can play __ musical instruments.	Musical Training, General Sophistication
I listen attentively to music for __ per day.	Active Engagement

APPENDIX B

Chord Progressions used for Stimuli

Diatonic

Modulating

Atonal

APPENDIX C

Table 2. Post-hoc Comparisons for Alternative 2 x 3 ANOVA Using Tukey HSD Test

	diff	lwr	upr	p adj
Diatonic:Low-Modulating:Low	1.25	-7.43	9.94	1.00
Atonal:High-Modulating:Low	2.63	-6.05	11.32	0.94
Atonal:Low-Modulating:Low	5.86	-4.89	16.60	0.58
Modulating:High-Modulating:Low	7.31	-2.11	16.73	0.21
Diatonic:High-Modulating:Low	11.26	1.29	21.23	0.02
Atonal:High-Diatonic:Low	1.38	-7.56	10.32	1.00
Atonal:Low-Diatonic:Low	4.61	-6.34	15.55	0.80
Modulating:High-Diatonic:Low	6.06	-3.60	15.71	0.42
Diatonic:High-Diatonic:Low	10.01	-0.19	20.20	0.06
Atonal:Low-Atonal:High	3.22	-7.72	14.17	0.95
Modulating:High-Atonal:High	4.68	-4.98	14.33	0.69
Diatonic:High-Atonal:High	8.62	-1.57	18.82	0.14
Modulating:High-Atonal:Low	1.45	-10.09	12.99	1.00
Diatonic:High-Atonal:Low	5.40	-6.59	17.39	0.75
Diatonic:High-Modulating:High	3.95	-6.88	14.77	0.88