Emotion and liking across sensory modalities

A study of aesthetic sensitivity and perceived emotion across visual and auditory stimuli

Amanda Andrén, Nike Backman, Celine Mileikowsky & Ella Söderberg

DM2350 Human Perception for Information Technology
KTH Royal Institute of Technology

Abstract

As introduced by Corradi et al. (2020), aesthetic sensitivity is to what extent an individual is influenced by a given feature within stimuli. Such a feature could be the visual symmetry of an image, or the melodic complexity of a musical motif. In this study, we explored the concept of sensitivity profiles: clusters of individuals that are equally sensitive to a certain feature. We wanted to find out if such sensitivity profiles within liking and perceived emotion exist, and if they exist across sensory modalities. A study was conducted where participants were asked to rate their liking and perceived emotional response towards a set of both auditory and visual stimuli. The data was then analysed with a linear mixed-effect model to find individual liking slopes for features within both modalities. A cluster analysis was performed, in order to examine the existence of sensitivity profiles. The results point to that there were no liking profiles across the visual and the auditory modalities, but that some profiles within the auditory modality exist.

1. Introduction

To understand why and how certain individuals *like* and *feel* about things in our surrounding world is of essence when trying to grasp the human mind and its perceptual functionalities. Why can one person like a piece of art that someone else finds revolting, and why can some music induce a certain emotion in one individual, but not in others? Is beauty, liking and emotion rigid and universal throughout time and space, or do they constantly alter depending on who the observer is? At least some answers to these grand questions might be found in the concept of *aesthetic sensitivity.* Aesthetic sensitivity can, according to Corradi et al. (2020) be described as "the degree to which someone's liking is influenced by a given visual feature". In their view, aesthetic sensitivity is not universal, and cannot be described with a general factor. Instead, it is dependent on an individual's sensitivity, or insensitivity, to certain features in a given stimulus. In this paper, we will continue to build on the work by Corradi et al., and investigate whether or not there are emotion and liking profiles to be found across sensory modalities, using their concept of aesthetic sensitivity as a basis.

2. Background

In a recent study by Corradi et al. (2020) a new concept for visual aesthetic sensitivity was introduced. The authors questioned the validity of Eysenck's traditional view where aesthetic sensitivity is seen as the "... biologically determined ability to appreciate objective beauty" (Eysenck, 1940), and suggested a new notion, as well as a method for how it could be measured. They mean that aesthetic sensitivity should be defined as the individual responsiveness to a particular object feature, and that it can be measured as the individual slope in liking ratings in linear mixed-effect models. By conducting several studies (Corradi et al. 2020; Clemente et al. under review) in the visual and auditory modalities they concluded the validity and usefulness of their new conception of aesthetic sensitivity. The results showed that the liking of stimuli varied on an individual level based on variations in four features: contour, symmetry, complexity or

balance. This means that, in conclusion, individuals aren't necessarily *generally* aesthetically sensitive, but instead sensitive to certain features more than others. In addition to this, they also found that there is no clear correlation between aesthetic sensitivity and intelligence, openness to experience, desire for aesthetics, art interest, or art knowledge. Overall, their findings contradict Eysenck's view on the general factor of aesthetic sensitivity. Friberg & Hedblad (2011) conducted a study where they concluded that emotional ratings of energy of valence could be estimated by perceptual audio features such as speed and rhythmic complexity.

With this study, we want to investigate aesthetic sensitivity within participants across sensory modalities, and explore the relationship between aesthetic sensitivity and perceived emotions, using the same stimuli as in previous studies on the new conception of aesthetic sensitivity. In particular, we are interested in investigating if there are any liking and emotion profiles across sensory modalities. We will do so by letting a number of participants rate visual and auditory stimuli varying in balance, contour, symmetry, or complexity. Each stimulus will be rated using three categories: liking, perceived valence, and perceived arousal. The collected data will be analysed, and the distributions of individual liking, valence and arousal slopes for the same attribute in both modalities will be examined.

Thus, the research question at hand is ***are there any liking and emotion profiles across sensory modalities?*** In other words, do people tend to cluster together into sensitivity profiles across sensory modalities and liking and emotion scales? We hypothesize we will find such sensitivity profiles: that it's likely that individuals cluster together based on how aesthetically and/or emotionally sensitive they are to certain features within stimuli, and that these clusters also exist across sensory modalities.

3. Method

3.1 Participants

A total of 60 subjects took part in this study. In the end the data collected from two participants was lost or deemed unusable, meaning that data from 58 persons was collected. Participants were recruited by students partaking in the course *Human Perception for Information Technology* given at KTH Royal Institute of Technology. All participants were asked to confirm that they were at least 18 years of age, had normal or corrected-to-normal vision and hearing and no cognitive impairments before taking part in the study. They were unaware of the aim of study, and were treated in accordance with local regulations and the Declaration of Helsinki.[1] Due to the circumstances of the Covid-19-pandemic, participants were asked to disinfect and were also offered the use of masks and gloves. Areas and materials used were disinfected between participants, and only a maximum of five people was allowed in the experiment room in order to ensure that a safe distance was kept.

3.2 Materials

A web experiment interface was created using LabJS, an online experiment builder.

[1]https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

Once completed, the interface was hosted on the web. The interface contained 92 auditory stimuli, from a set of stimuli created by Clemente et al. (2020), intended to be used in research music psychology, empirical aesthetics, music information retrieval, and musicology. They were divided into four subsets with varying values in either contour, symmetry, complexity or balance. In this way, the set of stimuli can be used to investigate how the variation in different features impact the aesthetic sensitivity in the auditory modality. Four of these were practice stimuli, where the participant was allowed to get familiar with the task without any significant data being collected. The interface also contained 70 visual stimuli, where four again were for practice. The visual stimuli was the same used by Corradi et al. (2020) and just like the auditory stimuli, they were divided into four subsets that had varying values in contour, symmetry, complexity or balance. The stimuli was presented in their subsets, with one stimulus, either visual or auditory, displayed at a time. Each stimulus included three ratings: liking, perceived valence, and perceived arousal.

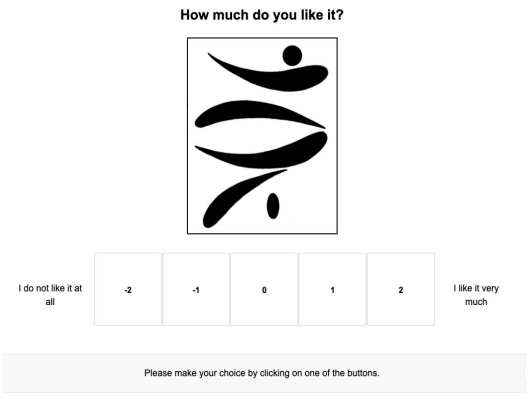


Fig. 1 *Rating of liking of a visual stimulus*

Liking was rated using a five-point Likert scale, rating from -2 (I do not like it at all) to 2 (I like it very much). The user was prompted to make their choice by clicking the button that corresponded most to their liking.

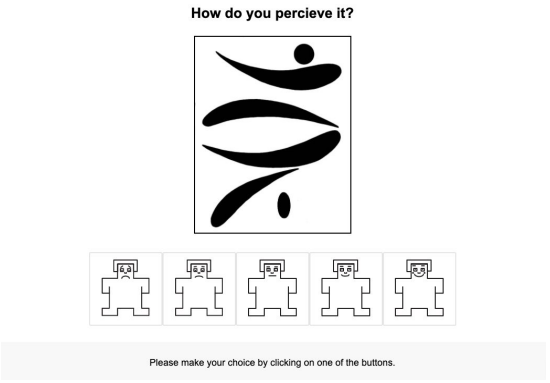


Fig. 2 *Rating of perceived valence of a visual stimulus*

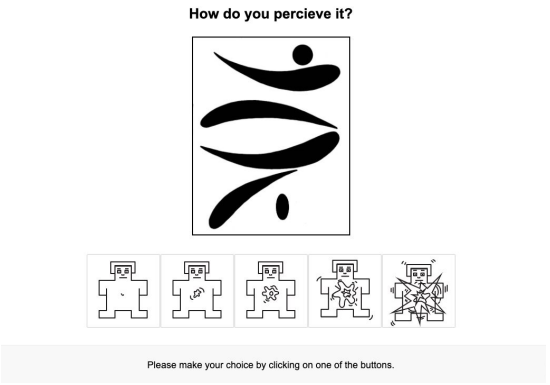


Fig. 3 *Rating of perceived arousal of a visual stimulus*

In order for participants to rate the perceived emotion, the Self-Assessment Manikin (SAM) was used. In particular, the two SAM-scales that measure valence and arousal were used. SAM consists of five figures which portray the level of affective response, and is a simple and scientifically tested method to measure these types of responses (Bradley & Lang, 1994). In the same manner as for the liking, users were prompted to make their choice by clicking the button that best corresponded to their perceived emotional response. See fig. 2 and 3.

The experiment interface was presented to

participants on a computer screen. They used a mouse to navigate around the interface, and to select ratings of stimulus. Participants wore headphones during the full duration of the auditory task of the experiment. All participants used a similar screen size and resolution, keyboard, mouse and headphones, and were exposed to the same illumination conditions.

3.3 Procedure

One to two participants at a time conducted the experiment while under supervision. Before the experiment started participants were informed that they would be exposed to a number of visual and auditory stimuli and asked to rate them according to their subjective liking of each given stimulus, as well as the perceived valence and arousal for each given stimulus. All participants were given the same standard oral and written instructions, and supervisors could at any point give further instructions to participants if the tasks were unclear. Then participants were seated by their screen and asked to begin the experiment and follow the instructions given in the interface. The experimental paradigm was structured as follows:

1. **Landing page**
   The landing page contained information about the background and purpose of the study. It informed participants that the study would take approximately 40 minutes to conduct, and they could take breaks between rating blocks of stimuli. Participants were also informed about terms of participation and the storing and use of collected data. Participants were asked to confirm that they were at least 18 years of age, had normal or corrected-to-normal vision and no cognitive impairments before being allowed to start the experiment, as well as giving informed consent about their participation.

2. **Sound calibration**
   In the following page participants were prompted that they needed to use headphones during the experiment. Once they had headphones on, they were asked to calibrate the sound level, by listening to a calibrating sound, and turning up the volume to a comfortable level. This sound level remained constant throughout the experiment.

3. **Behavioral tasks**
   The behavioral tasks of the experiment were divided into two parts: one visual, and one auditory. The order in which the tasks were given was randomized for each participant. In turn, the visual task was divided into three blocks, and the auditory task was divided into four blocks. The order of the blocks were individually randomized, as were the stimuli within each block. Between each completed block participants were allowed to take a break before continuing with the next one. In each block either a visual or auditory stimulus was presented to the participant. If the stimulus was auditory, the participant was allowed to listen to it for as many times as they wanted. The participant was asked to rate each given stimulus using

three categories: liking, perceived valence, and perceived arousal. The participant rated each stimulus by clicking a button by mouse, and no time constraints were given for the rating.

4. **Questionnaires**
   Once the behavioral tasks were finished, the participants were asked to fill out two questionnaires in order to evaluate their cognitive and emotional needs: the Need for Cognition Scale and the Need for Emotion Scale (Cacioppo, 1984; Raman, 1995). These scales are used to capture the participants' need for seeking out both thinking and emotion in stimuli. The Need for Cognition Scale was rated using a 9-point Likert scale, rating from -4 (Very strong disagreement) to 4 (Very strong agreement). The Need for Emotion Scale was rated using a 5-point Likert scale, rating from -2 (Strongly disagree) to 2 (Strongly agree).

5. **Demographics**
   Finally, the participants were asked to answer some demographic questions. They were asked about their age, gender identity, highest educational degree, highest degree of formal education in music, and highest degree of formal education in visual art, architecture or design. See appendix 1 for the full demographic questions.

3.4 Data analysis

Since the collected data includes both fixed effects, which are the same across an entire population, and random effects, which vary across individuals, the data was modeled using a linear mixed-effects model.[2] A linear mixed-effects model, or LMEM, has the advantage of working even with sets of data where there is dependence. For this project, the function lmer which fits a LMEM to the data was used to model the effects. This function can be found in the R package lme4. (Bates et al., 2015)

To determine if cross-modality profiles exist, a cluster analysis was performed, using the R package mclust. (Scrucca et al., 2016). Clusters describe groupings of objects in such a way that objects within a cluster are more like other objects in the same cluster in a certain aspect than the objects outside of the cluster. If a cluster analysis of the perceived emotion and liking of stimuli across modalities results in more than one cluster, it can be concluded that there are ways to categorize liking and perception profiles across modalities. (Clemente et al. under review). A cluster analysis can be performed using many different methods, in the data analysis for this study, a Gaussian finite mixture model fitted by Expectation-Maximization algorithm was used.

4. Results

The results section is divided into two parts: the general effects of features on aesthetic sensitivity and perceived emotion, and the emotion and liking profiles across sensory modalities.

4.1 Effects of features on aesthetic

[2] http://www.stat.columbia.edu/~gelman/research/published/banova7.pdf

sensitivity and perceived emotion

The collected data was analysed using a linear mixed effects model. The distributions of individual liking, valence, and arousal slopes were plotted for both modalities. Overall, we found a wide variability of sensitivities for each feature (contour, symmetry, complexity, or balance) within both the visual and auditory modality, which supports the findings of Corradi et al. (2020) and Clemente et al. (under review). See appendix 2 for the full set of plots.

4.1.1 Visual stimuli

The visual stimuli varied in visual contour from jagged to smooth, in visual symmetry from asymmetrical to symmetrical, in visual complexity from simple to complex, and in visual balance from unbalanced to balanced. Results were plotted using histograms for aesthetic and emotional sensitivity for each feature, where positive slopes indicate a higher liking for curved, symmetric, complex, and balanced images, and negative slopes indicate a higher liking for sharp-angled, asymmetric, simple, and unbalanced visual designs.

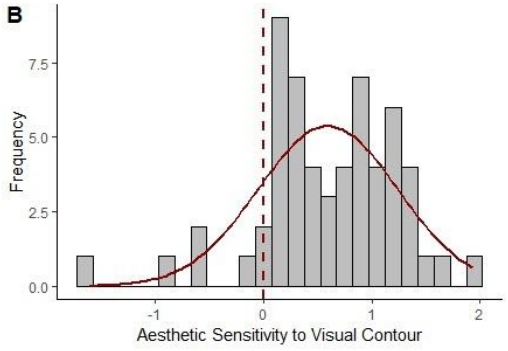


Fig. 4 *A histogram of the slope for aesthetic sensitivity to visual contour. The positive slope indicates a liking for smooth contours.*

See appendix 2 for the full set of plotted slopes.

4.1.2 Auditory stimuli

The auditory stimuli varied in melodic and rhythmic contour from smooth to jagged, in musical symmetry from symmetrical to asymmetrical, in melodic and rhythmic complexity from simple to complex, and in musical balance from balanced to unbalanced. Just as for the visual stimuli, the results were plotted using histograms, where negative aesthetic sensitivities to the musical features mean preference for balanced, smooth, symmetric, and simple musical motifs, and positive aesthetic sensitivities mean preference for unbalanced, jagged, asymmetric, and complex musical motifs.

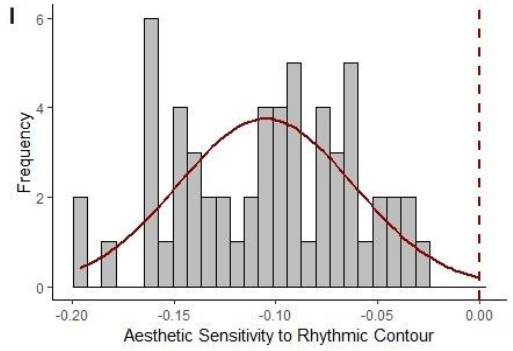


Fig. 5 *Histogram of the slopes for aesthetic sensitivity rhythmic contour. The negative slope indicates a liking for rhythmically smooth musical motifs.*

4.2 Emotion and liking profiles across sensory modalities

Clustering analyses were run within and across features and sensory modalities. The analyses showed if individuals tend to cluster together into emotion and liking profiles based on their sensitivities to certain features. In general, the analyses showed that no emotion profiles could be identified within and across features and sensory modalities. However, as for liking, the results showed some aesthetic sensitivity profiles were found. The tables in the following section shows the number

of individuals that clustered based on their sensitivity to certain features, and around which value on the individual liking slopes this clustering appeared.

Number of individuals	**34**	**24**
Visual balance	-0.011158376	-0.015454099
Visual contour	0.594306335	0.595540837
Visual symmetry	0.630290904	0.791838556
Visual complexity	0.007663943	0.006500535
Musical balance	-0.104349634	-0.208487703
Melodic contour	0.264192612	-0.001968322
Rhythmic contour	-0.074401238	-0.146795959
Musical symmetry	0.199316114	0.174359799
Melodic complexity	0.340990254	0.127134812
Rhythmic complexity	-0.035298151	-0.037874190

Table 1 *Cluster analysis results for aesthetic sensitivity across features and modalities*

The cluster analysis for aesthetic sensitivity across features and modalities (see table 1) shows two clusters for each feature. However, the clustering is weak for most features: there isn't much difference in preferred value between one cluster and the other. The most significant clustering is found in terms of melodic contour, where one cluster shows a slight preference for melodically smooth musical motifs (-0.001968322) and the other cluster shows a preference for melodically jagged musical motifs (0.264192612). Some clustering is also found for melodic complexity, where one cluster shows preference for somewhat complex musical motifs (0.127134812) and one cluster shows preference for more complex musical motifs (0.340990254).

In order to determine whether both modalities contributed to the clustering, separate tests on aesthetic sensitivity within each modality was run.

Number of individuals	**58**
Visual balance	-0.01297218
Visual contour	0.59482759
Visual symmetry	0.69850213
Visual complexity	0.00717271

Table 2 *Cluster analysis for aesthetic sensitivity in the visual modality*

For the visual modality (see table 2) no clustering was found. Each feature only generated one preferred value. This indicates that there are no sensitivity profiles in regards to these features.

Number of individuals	**28.718855445**	**30.07706160**
Musical balance	-0.206793915	-0.10497559
Melodic contour	-0.001094719	0.26515421
Rhythmic contour	-0.146539990	-0.07415329
Musical symmetry	0.174049526	0.19969699
Melodic complexity	0.126835355	0.34250518

Rhythmic complexity	-0.038349436	-0.03493029

Table 3 *Cluster analysis for aesthetic sensitivity in the auditory modality*

As for the auditory liking slopes (see table 3), each feature generates two clusters, where the most notable differences in liking once again can be found in melodic contour and melodic complexity.

5. Discussion

The results in general showed no support for emotional sensitivity profiles across sensory modalities. However, as for aesthetic sensitivity some interesting results were found: individuals do tend to cluster together somewhat in regards to their liking for melodic contour and complexity, and show different preferences for values in these features. The clustering was not very significant, though, and there was no indication that there are profiles with strongly contrasting preferences. The results also showed that the liking slopes for visual aesthetic sensitivity didn't contribute to this clustering. With that in mind, we have found no aesthetic sensitivity profiles that exist across sensory modalities. The only profiles found existed within the auditory modality. Conclusively, this disproves our hypothesis that we would find emotion and liking profiles across sensory modalities. The hypothesis is only partly correct in that some aesthetic sensitivity profiles were found, however, these were auditory-specific. In general, this speaks for individuals being more prone to cluster together into liking profiles as for their taste in music, than their taste in visual stimuli. This could be interpreted as a phenomenon that we often witness in everyday life: that people are generally more certain of their taste in music than their taste in artwork. You could expect almost everyone to have an answer of some sort when asked about their music preferences, but assumptively, not as many individuals would have a clear answer about their artwork preferences. Future research should continue to explore the possible sensitivity profiles across sensory modalities, and extend these past visual and auditory. What about taste, smell, or touch - are there any profiles to be found within these modalities?

One limitation in this study was the current circumstances due to the ongoing covid-19 pandemic. If the study had not been conducted during a pandemic it's possible that a higher number of participants could have been recruited. One alternative discussed was to conduct the study online. Due to the lack of control of what equipment remote participants would use while partaking in the study, as well as lack of supervision of participants completing the tasks, this option was deemed not appropriate. Actually conducting the study in person made sure that we could be certain that all participants received the same standard instructions, that they could ask questions about possible uncertainties, and that they kept their focus on the given tasks and that they completed them within a reasonable timeframe. This allows us to conclude that the data collected was accurate to a higher degree than if the study had been online. However, it's possible that we could have collected data from a significantly higher number of participants if the study would have been conducted remotely.

When participants were debriefed about the study some expressed that it took a long time to complete, and that they lost interest or focus in the task, especially towards the end. It's possible that this could have affected the quality of the ratings, and in the end, the outcome of the study. To fully eliminate the possibility that the data was affected by participants losing interest in the given task, the study should be replicated with a smaller number of visual and auditory tasks.

6. Bibliography

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1 - 48. doi:http://dx.doi.org/10.18637/jss.v067.i01

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of personality assessment*, 48(3), 306-307. https://doi.org/10.1207/s15327752jpa4803_13

Clemente, A., Vila-Vidal, M., Pearce, M. T., Aguiló, G., Corradi, G., & Nadal, M. (2020). A Set of 200 Musical Stimuli Varying in Balance, Contour, Symmetry, and Complexity: Behavioral and Computational Assessments. *Behavior Research Methods*. https://doi.org/10.3758/s13428-019-01329-8

Clemente, A., Pearce, M. T., and Nadal, M. (under review). Musical Aesthetic Sensitivity.

Corradi, G., Chuquichambi, E. G., Barrada, J. R., Clemente, A., & Nadal, M. (2020). A new conception of visual aesthetic sensitivity. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12427

Eysenck, H. J. (1940). The general factor in aesthetic judgments. *British Journal of Psychology, 31,* 94–102.

Friberg, A. & Hedblad, A. (2011). A Comparison of Perceptual Ratings and Computed Audio Features. *Proc. 8th Sound and Music Computing Conference.*

Raman, N. V., Chattopadhyay, P., & Hoyer, W. D. (1995). Do consumers seek emotional situations: The need for emotion scale. *NA - Advances in Consumer Research*, 22, 537-542. https://www.acrwebsite.org/volumes/7804

Scrucca, L., Fop, M., Murphy, T., & Raftery, A. (2016). mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. The R journal, 8 1, 289-317 .

Appendix 1: Demographic questions

These were the demographic questions asked to participants, and the alternatives given.

1. How old are you, in years?

2. What is your current gender identity?

- Female
- Male
- Other

3. What is the highest educational degree you have ever achieved?

- None
- Primary
- Secondary
- Bachelor
- Masters
- Doctorate

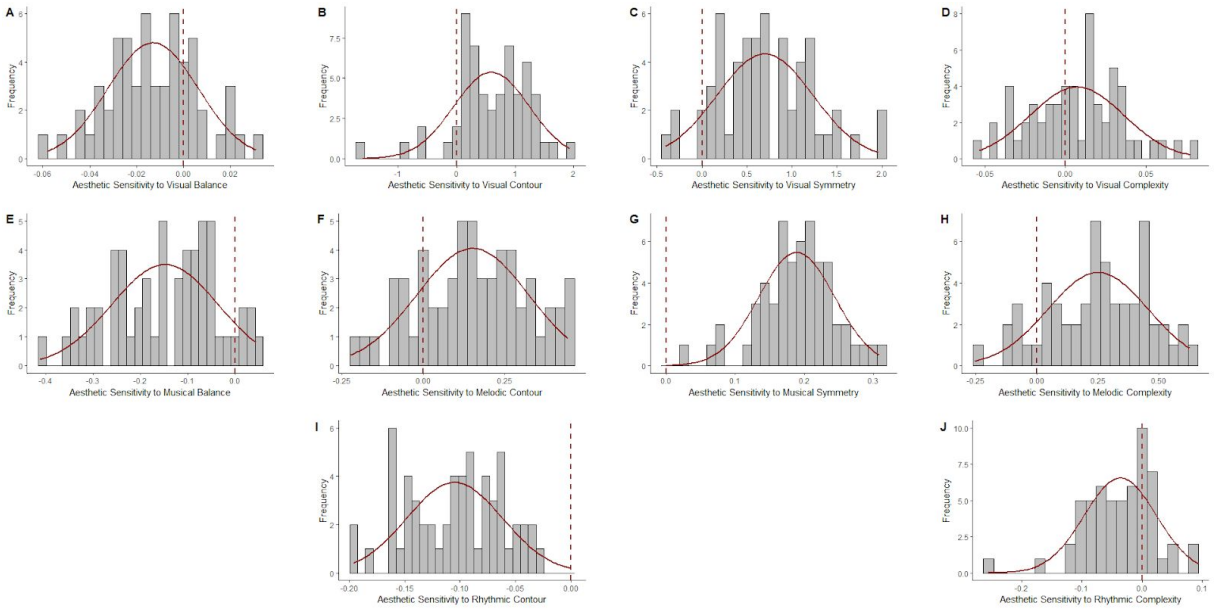
4. What is the highest degree of formal education in music that you have ever attained?

- None
- Primary
- Secondary
- Bachelor
- Masters
- Doctorate

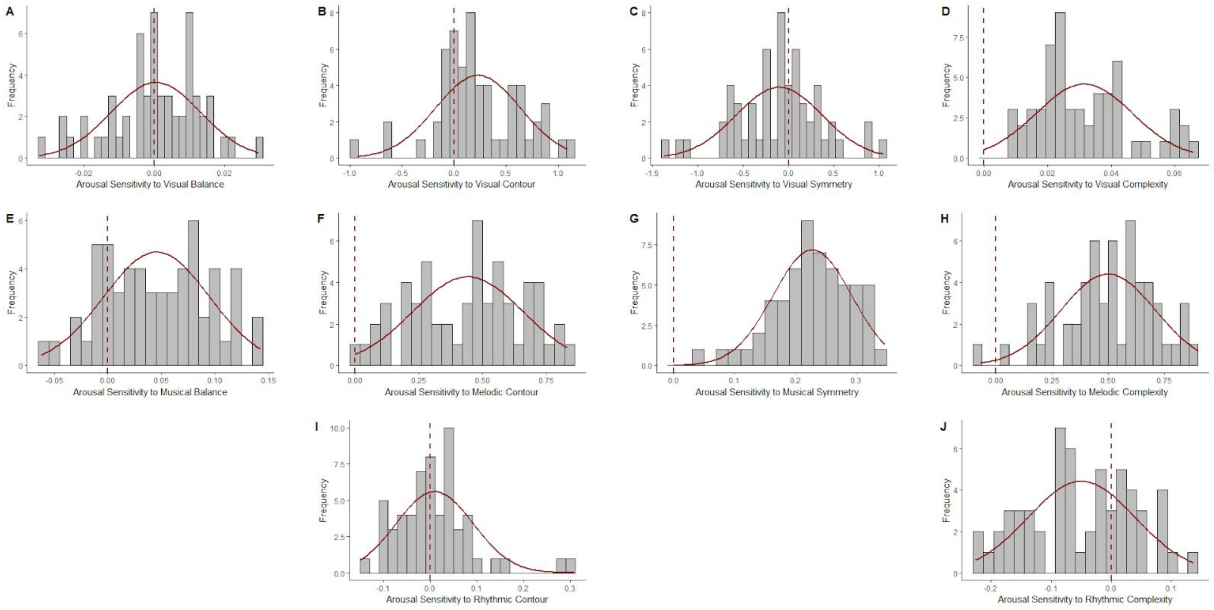
5. What is the highest degree of formal education in visual art, architecture, or design that you have ever attained?

- None
- Primary
- Secondary
- Bachelor
- Masters
- Doctorate

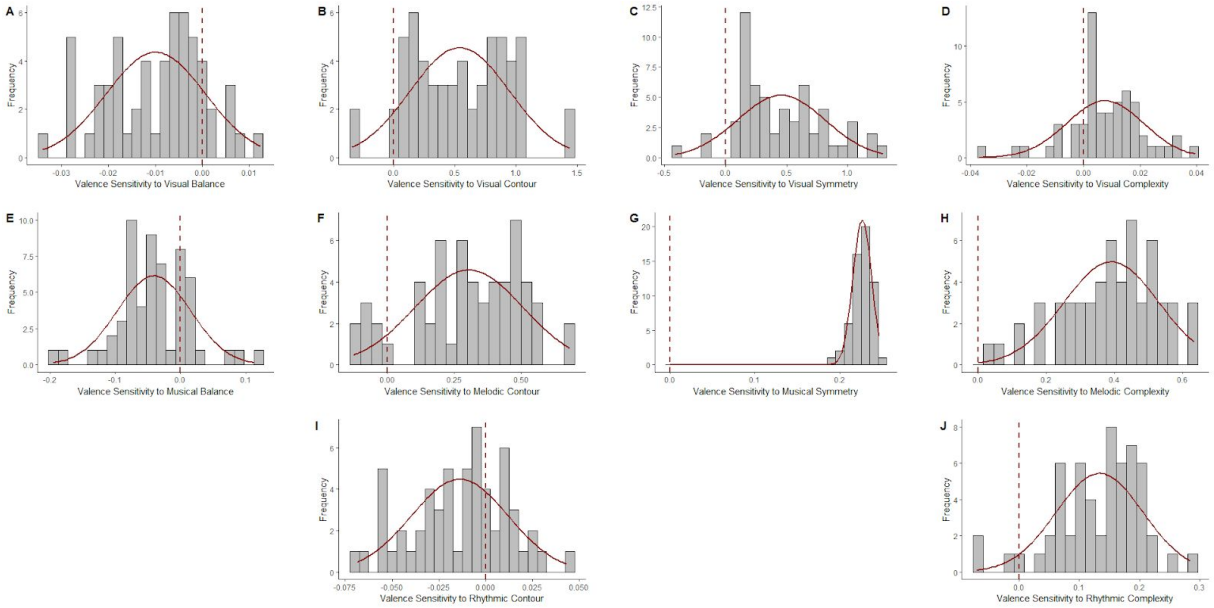
Appendix 2: Individual liking and perceived emotion slopes



Aesthetic sensitivity to features



Arousal sensitivity to features



Valence sensitivity to features