
Guideline for Subjective Evaluation

Thank you very much for participating in this evaluation! Be sure to read the full article carefully before starting.

Task

The current folder contains several groups of samples (i.e., the .mid files), and samples in each group are generated by different models. You need to score each sample and fill in the blanks on score.csv.

Listening to Music Samples

To ensure the same sound effect, please download Musescore 3 (click here to go to the official website), and use the software to open the samples. Simply click the button  to locate the beginning, and then click the Play button  to play.

Fill in Score Sheet

You are advised to open score.csv using Excel and other spreadsheet tool. Please fill in an integer **from 1 to 10** for each of the metrics:

1. **Musicality:** whether the music is beautiful, pleasant, and interesting, like the music made by human.
2. **Short-Term Structure:** in a relatively short range (generally 1-4 bars), whether the music has obvious structures and regularities.
3. **Long-Term Structure:** whether bars are associated with previous bars at a long distance (usually 8 or more bars apart), like repeated melodies, similar rhythms, etc.
4. **Overall:** Take all aspects together and give an overall score.

For the 1st – 4th metrics, please try to make their scores **independent**, that is, the score of one metric does not affect the scores of the others. Also, if you have anything to tell us, you can put it in the comment field (optional).


Please make sure that the scores for the samples within a group can **reflect the orders**. For example, if sample 1 is better than sample 2, then the score of sample 1 should be larger than that of sample 2. Of course, if they are too close to call, you can score the same.

Please **always complete the scoring for all samples in a group within a short time** to ensure the stability of the scoring. You can do this at different times for different groups, but try to keep the scoring criteria stable.

It is suggested that the standard be likened to a school exam: 6 is passing, 1 is totally unacceptable, and 10 is perfect.

Other Suggestions

- Please score the whole music piece, not just the beginning or a short segment.
- The score is not affected by how good the cadence is (that is, whether the music ends in a proper way).
- Please score mainly by what you hear. Musescore shows the music score to help you understand the structures of the music and help you make a more comprehensive evaluation, but focusing too much on the music score can cause you to ignore the music itself.
- All the samples use the instruments including square synthesizer, bass, piano, guitar, violin, drum, with the square synthesizer playing the lead. Please do not score from the perspective of the tones of the instruments.

Please feel free to ask any questions! Thank you .