Aims and objectives
Retard in upbeats is a common effect in some musical situations; in other contexts retarded anacrusis is regarded as inappropriate. The effect is often used to create musical tension and expectation of the coming downbeat, and the magnitude varies substantially depending on the context and also from one interpreter to another in the same piece. The aim of the study was to see if the magnitude of upbeat retard can be changed without distorting the rest of the performance, and also if the magnitude of retard influenced the listeners’ appreciation of the performance.

Context
Three pieces of music (Bach Bourrée for solo cello, Schubert string quintet 4th mvt. and the start of Mahler’s Fourth Symphony) were selected, all of which had an anacrusis motive and in which more or less anacrusis retard was used in all known interpretations. Two original performances of each piece were selected, one with much retard and one with relatively little. From each of these six performances three manipulations were made, increasing or decreasing the magnitude of anacrusis retards by 15 and 30 percent but otherwise leaving the performances unchanged. From the “shallow” performances two “deeper” and one even “flatter” versions was made, and opposite with the “deep” originals. In all six original performances and eighteen manipulations were used in the test. 49 participants, music students and professors at the University of Stavanger listened to the 24 performances and evaluated them answering three questions for each example:
1. Has this performance too little (1) suitable (5) or too much tempo variation (10) in the marked upbeats? [Score was given] (10 point Likert scale)
2. Rate the overall quality of the performance (1 poor to 10 excellent)
3. Do you think this is an original or a manipulated performance?
Results

Identification of original/manipulated was low (59%). The identification was lower for originals than for manipulations (48% versus 63%). Highest correct identification was for the manipulations with the deepest retards, whereas the flattened manipulations (reduced retards) had a significantly lower correct identification. The “deepest” originals (using much retard) were by the majority regarded as manipulations, the shallow originals were more correctly identified.

In the ratings there was a general preference for the “shallow” performances. This was somewhat surprising, and different from Johnson’s (2003) findings where the ‘deepened’ performances were preferred. Fig. 1 illustrates the different mean ratings for each performance. Notice the large standard deviations indicating substantial individual differences among listeners.

Figure 1 Mean rating for the 24 performances. Bars = sd. Red are originals.

Figure 2 shows the ratings collapsed by type of manipulation. The shallow performances are rated clearly higher than the deeper, regardless of whether they were originals or manipulations.

---

1 The model used by Johnson is an average of 15 different performances, and is thus a performance ‘evened out’ from the start. This could be at least part of an explanation for this difference.
In the evaluating of tempo variation magnitude (question one), there was the same diversity in the evaluation as seen in the rating. However, the majority of answers gave a score of around 5 indicating that the participants found most of the performances acceptable or fine (fig 3).

There is a markedly skewed distribution with a mean at 6.0 indicating too much tempo variation. The participants seem to prefer rather flat and straightforward performances. Collapsing the magnitude evaluations shows this tendency clearly (fig 4). We see that the mean for each category of “deepness” varies from 4.3 (slightly too little) to 7.8 (way too much). In this graph both originals and manipulations are included, showing the four versions of each performance from the shallow originals and the deep originals. The difference between these
evaluations is highly significant ($F_{3,1144}=28.75 \ p<0.001$). It is worth noticing that the evaluation of tempo variation magnitude is nearly identical for the shallow originals “deepened” by 30% and the deep originals, and likewise for the “flattened” deep and the flat originals. This similarity was intended.

![Figure 4 Evaluation of tempo variation magnitude for each manipulation. 5 (line)= fine, >5=too much  <5 too little. Arrows indicate originals.](image)

When splitting this for each performance the picture is even clearer. Fig 5 shows a high number giving five (= fine) for the shallowest versions of both deep and shallow originals. The evaluations of the deeper versions are more spread.
Figure 21 Distribution of magnitude evaluation for each type of manipulation (total count each group = 49x3=147)

**Key Contribution**

Manipulated piano performances was used in a study by Repp (1997) where judges rated both original piano performances by experts and students and averaged performances. The highest rating was given to the expert average, the student average also scored high, while only one expert and four student performances were in the top list.

The tendency in the present test is more in line with Repp’s findings (preference for standard, evened-out performances) than with Johnson’s. The participants in this test clearly preferred the “easy-going”, straightforward and “simplified” performances rather than the deeper and more “resistant” and laborious ones with much tempo variation. We can only speculate why this is so. One explanation could be that the repeated hearing of the same piece creates a preference for the averaged. Also, the short excerpts do not constitute a condition for the involvement from listeners necessary for appreciation of the deeper performances. Other possible explanations will be discussed.

**Key words**

Tempo variation, interpretation, performance evaluation, upbeats

**Bibliography**