# Downbeat delays are a key component of the swing feel in jazz

**Corentin Nelias**
Max-planck institute for dynamics and self-organization
**Eva Marit Sturm**
Georg-Elias-Mueller Institute of Psychology
**Thorsten Albrecht**
Georg-Elias-Mueller Institute of Psychology
**York Hagmayer**
Georg-Elias-Mueller Institute of Psychology
**Theo Geisel** ( geisel@ds.mpg.de )
Max Planck Institute for Dynamics and Self Organization

**Article**

**Keywords:** jazz music, synchronization, swing, downbeats, microtiming deviations (MTD), nonlinear dynamics, complex systems

**Posted Date:** October 25th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-830993/v1



**Version of Record:** A version of this preprint was published at Communications Physics on October 6th, 2022. See the published version at https://doi.org/10.1038/s42005-022-00995-z.

# Downbeat delays are a key component of the swing feel in jazz

Corentin Nelias[1,2], Eva Marit Sturm[3], Thorsten Albrecht[3], York Hagmayer[3], and Theo Geisel[1,2,4*]

[1]Max Planck Institute for Dynamics and Self-Organization, 37077 Göttingen, Germany
[2]Department of Physics, Georg-August-University Göttingen, 37077 Göttingen, Germany
[3]Georg-Elias-Müller Institute of Psychology, Georg-August-University Göttingen, 37073 Göttingen, Germany
[4]Bernstein Center for Computational Neuroscience, 37077 Göttingen, Germany
[*]Corresponding author: Theo Geisel, theo.geisel@ds.mpg.de

October 15, 2021

## Abstract

The swing feel is a salient feature of jazz music, yet its main psychoacoustical and musical components have remained elusive - save the obvious long-short subdivision of quarter notes. In particular, the possible role of microtiming deviations for the swing feel has been a subject of long-standing controversy. Adopting an operational definition of swing we present a study which ultimately demonstrates a positive effect of certain microtiming deviations on the swing feel. We manipulate the timing of original piano recordings to carry out an experiment with expert jazz musicians measuring the swing feel of different timing conditions. Thereby we prove that slightly delayed downbeats and synchronized offbeats of a soloist with respect to a rhythm section enhance the swing feel. Analyzing a set of 456 full solo performances we find that many jazz musicians do use minute downbeat delays and we characterize the dependence of the average downbeat delay on tempo and swing ratio. These results show that systematic microtiming deviations in the form of downbeat delays are a key component of the swing feel in jazz.

## Introduction

"What is this thing called swing", is a question raised already by Louis Armstrong in a well-known song. The swing feel certainly is one of the most salient features of jazz music and is considered an essential ingredient of jazz performances. Yet astonishingly, a century after jazz musicians like Armstrong and Ellington came on stage, it is still controversial what is the nature of the swing feel and what are its main musical and psychoacoustical components. It was even argued in the past that *"you can feel it but you just can't explain it"* [1], or similarly, according to The New Harvard Dictionary of Music, that swing is *"an intangible rhythmic momentum in jazz. Specifically manifested in a variety of relationships between long and short notes, or in the presentation of single notes, swing defies analysis"* [2].

Among the components of the swing feel only one is established unambiguously so far, the conspicuous uneven subdivision of quarter notes (i.e. the succession of long and short eighth notes mentioned above). It is measured by the so-called swing ratio, i.e. the length ratio of consecutive long and short eighth notes known as *"downbeats"* and *"offbeats"*. Listening to computer-generated jazz music that was "swingified" by merely implementing a swing ratio [3], it is obvious that this is not sufficient and that there must be other components. But which are these components, and which ones are important?

It has long been speculated that rhythmic effects, in particular *microtiming deviations* (MTD), play the other major role for the swing feel. However, while the importance of the *swing ratio* is generally accepted, the role of rhythmic microtiming deviations has been a subject of controversy for many decades. Various authors have emphasized the importance of participatory discrepancies; i.e. *"little discrepancies within a jazz drummer's beat, between bass and drums, between rhythm section and soloists, that create 'swing' and invite us to participate"* [4–10]; or other kinds of microtiming deviations [11–14]. Others, however, contested their influence and rather stressed the importance of rhythmic accuracy [15–20][‡]. Many of these claims were based on observational analyses of extracts from individual jazz musicians, which may explain the origin of the controversial claims, as MTD are not used equally by all musicians. It is not clear, however, that MTD - even if they occur - are specifically related to and enhance the swing feel. Is there a way to prove that MTD do contribute substantially to the swing feel?

Adopting an operational definition of swing the present paper presents the first study which is able to clarify the controversy and to rigorously demonstrate a positive effect of certain microtiming deviations on the swing feel. By manipulating the timing of original piano recordings and measuring the swing feel of different manipulated versions as rated by jazz musicians we demonstrate that a playing style with *systematic* microtiming deviations, slightly *delaying downbeats* of the soloist with respect to the rhythm section while *synchronizing offbeats*, considerably enhances the swing feel. As the soloist's offbeats need to remain synchronized with the rhythm section, this playing style has an influence on the *swing ratio*. If the downbeat onsets of the soloist are delayed (their durations thus shortened) and offbeats remain synchronized, this implies a somewhat smaller swing ratio for the soloist than for the rhythm section and may create a perceived friction between them.

Analyzing short musical extracts in six recorded solos of different jazz musicians Friberg and Sundström observed that such downbeat delays did show up in a majority (not all) of their extracts [13]. As the variation in their measurements was quite large and some musicians did not make use of such delays, they called for a substantially larger data set to confirm these anecdotal observations [13]. We therefore also analyze a large set (456) of full solo performances using the Weimar Jazz Database (WJD) [21] and determine the average downbeat delays. We find downbeat delays of jazz soloists as a general trend and we find that their magnitude decreases with tempo.

Taken together the results of our experimental and our observational study lead to the conclusion that *downbeat delays are a key component of the swing feel* in jazz. They underline the general importance of timing and rhythmic effects for the swing feel and resolve the long-standing controversy on the role of microtiming deviations by demonstrating that certain *systematic microtiming deviations* enhance the swing feel, while involuntary *random microtiming deviations* do not, as we showed in previous work [20]. That downbeat delays could play such an important role for the swing feel was widely unknown. Professional and semiprofessional jazz musicians participating in our online experiment reported a pleasant friction between soloist and rhythm section, but were unaware of the effect and could not determine its nature (cf. "*you can feel it but you just can't explain it*" [1]. We emphasize that the phenomenon reported here (with downbeat delays of the order of 30 ms or 9% of a quarter note for intermediate tempi) is not identical to the well known laid-back mode, where musicians play with much larger and easily perceivable delays.

# Results

## Timing analysis of jazz solos

We begin by an in-depth analysis of onset timing in a large set of jazz recordings. As outlined above, our main goal is to prove that there is a positive effect of downbeat delays on the swing feel, but we first want to clarify the question, whether or not and to which extent soloists tend to delay their downbeats with respect to the rhythm section. We evaluated data from the Weimar Jazz Database [21], which contains accurately labeled transcriptions of 456 jazz solos of various artists, and gives access to several quantities like note positions or rhythmic value. We want to stress that our general analysis, which does not consider individual differences and different playing styles, can only have a limited accuracy with a large scatter of data. Nevertheless it is able to reveal general trends, which is the goal of this section.

[‡]Strictly speaking, refs. [15] to [19] have investigated the possible impact of MTD on groove, a concept which in the literature is defined by the impetus of wanting to move along with the rhythm. As there is no swing without groove, the (negative) conclusions of these references also extend to swing. Our previous paper ref. [20], however, did investigate swing and found that involuntary *random* MTD did not enhance the swing feel.

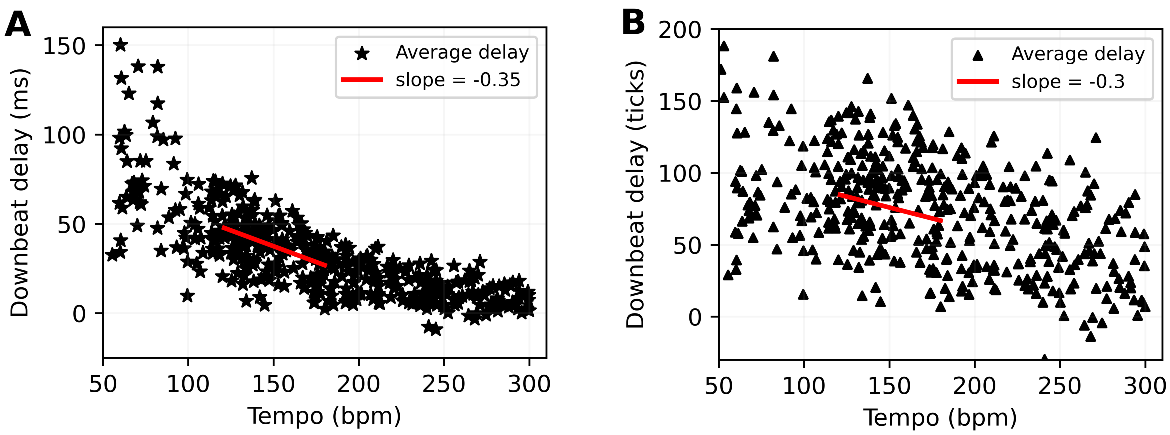


Figure 1: **Average downbeat delays of soloists as a function of tempo.** Each point in the scatter plots corresponds to a piece of the Weimar Jazz Database [21]. **(A)** Average downbeat delays in milliseconds as a function of tempo in beats per minute. The red line delimits the tempo range of pieces used in our experiment (see Methods section) and corresponds to the fit in Fig. 1b. The scattered data exhibit mostly positive delays with generally a nonlinear trend, which is nearly linear in this restricted tempo range. **(B)** Average downbeat delays expressed in ticks as a function of tempo. The red line shows a linear fit to the data.

For each given piece in the database [21], we isolated every downbeat-offbeat pair of the solo to compute the average downbeat delay and swing ratio (averaged over each solo) as a function of tempo using the downbeats of the drums as a reference. The results presented in Fig. 1 show the existence of finite downbeat delays in most cases (with the exception of a few negative and a few very small delays). The data show some variation, probably reflecting individual preferences, but there is a clear trend for decreasing delays with increasing tempo (Fig. 1a). The trend becomes nearly linear, if the downbeat delays are measured in *ticks* (Fig. 1b). Ticks represent fractions of quarter notes (which are subdivided into 960 ticks) and are not an absolute measurement of time (see equation (1)). The figure demonstrates that many soloists are using *systematic* MTD, i.e. positive downbeat delays, which typically are *of the order of 30 ms or 85 ticks for intermediate tempi of about 150 bpm*. (This value in ticks corresponds to delays of about 9% of a quarter note). While this is true for a majority of jazz soloists, it should be mentioned that a few soloists use only small or no downbeat delays at all.

This trend did not change, when we considered jazz sub-genres ("bebop", "swing" or "hardbop") separately (Supplementary Fig. 6). Of course the magnitude of downbeat delays may vary within a solo or a whole piece and it makes sense to also look at individual delays in their musical context. Here, however, we want to detect general trends and are therefore studying average quantities. It is important to note that the standard deviation of the downbeat delays is smaller than their mean (for tempi below 200 bpm, see Supplementary Fig. 1).

The *swing ratio* is another important parameter. Albeit not in the focus of the present paper, we determine it here, as it is also relevant for the swing feel. In particular, we realized in the course of our experimental study that it was crucial to choose a suitable swing ratio before applying systematic timing manipulations (for details see the end of the results section and section 3 of the Supplementary Information). The swing ratio is a measure of non-isochronous metrical subdivisions. Non-isochronous rhythmical patterns are prominent in jazz music, but are found also in other cultures, e.g. in Malian jembe drumming and Uruguayan candombe drumming [22–24]. While the swing ratio has been extensively studied for drummers [13, 14, 25, 26], the swing ratio of soloists and in particular how its optimal value varies with tempo is still not unambiguously established [27].

We determined the mean swing ratio of the soloists using the definition of Eq. (2) for each of the 456 pieces of the Weimar Jazz Database as described in the Methods section. The results are shown as a function of tempo in Fig. 2. Remarkably we find that the mean swing ratios of the soloists are much smaller than generally believed and also smaller than reported in early observational studies [11, 27]. In particular, the figure demonstrates that the noted triplet feel (or ternary feel, i.e. a swing ratio of 2:1) is rather a myth as far as soloists are concerned. Most of them use swing ratios that are below 1.5. For fast tempi (more than 160 bpm), one finds a decreasing trend of the soloists' swing ratio with increasing tempo. So far this is similar to the trend reported for the swing ratio of drummers [13, 28]. On the other hand the trend is reversed for medium to slow tempi (below 160 bpm), where the soloists' swing ratio tends to decrease with decreasing tempo. This means that drummers and soloists follow two opposing trends regarding the swing ratio in this tempo range and that the swing ratio of soloists tends to be smaller than that of the rhythm section. We also studied other characteristics of the recordings such as the position of individual triplets as a function of tempo. These additional findings are included in the Supplementary Information.

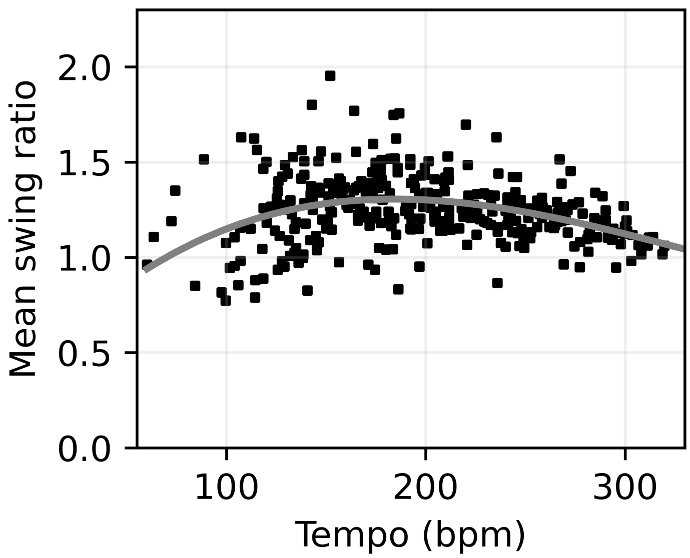


Figure 2: **Mean swing ratios of soloists as a function of tempo.** Each point corresponds to a piece of the WJD database [21] and represents the soloist's averaged swing ratio as a function of tempo. A quadratic fit to the data (grey line) as an indicator of preferential swing ratios reveals an increasing trend as a function of tempo up to 160 bpm and a decreasing trend above 160 bpm. The swing ratio of most soloists lies below 1.5, thus is much smaller than generally believed, and does not correspond to a triplet feel (i.e. swing ratio 2:1).

## Experiment investigating the perceived swing feel

The above empirical observations indicate that a large fraction of jazz musicians play jazz solos with downbeats slightly delayed with respect to the rhythm section. Nevertheless the question remains, whether these delays are related to the swing feel and whether they are essential for the swing feel, in particular as not all jazz musicians use them. To address this question we adopted an experimental approach, which we developed for a previous microtiming study on swing [20]. Manipulating the onset timing in MIDI recordings of piano jazz performances and letting expert jazz musicians rate the swing feel of different manipulations gives us the possibility to clarify whether different ways of microtiming have a positive effect on the swing feel. In that previous study we investigated the impact of *random* microtiming deviations by amplifying them, deleting them, and inverting them. We showed that random microtiming

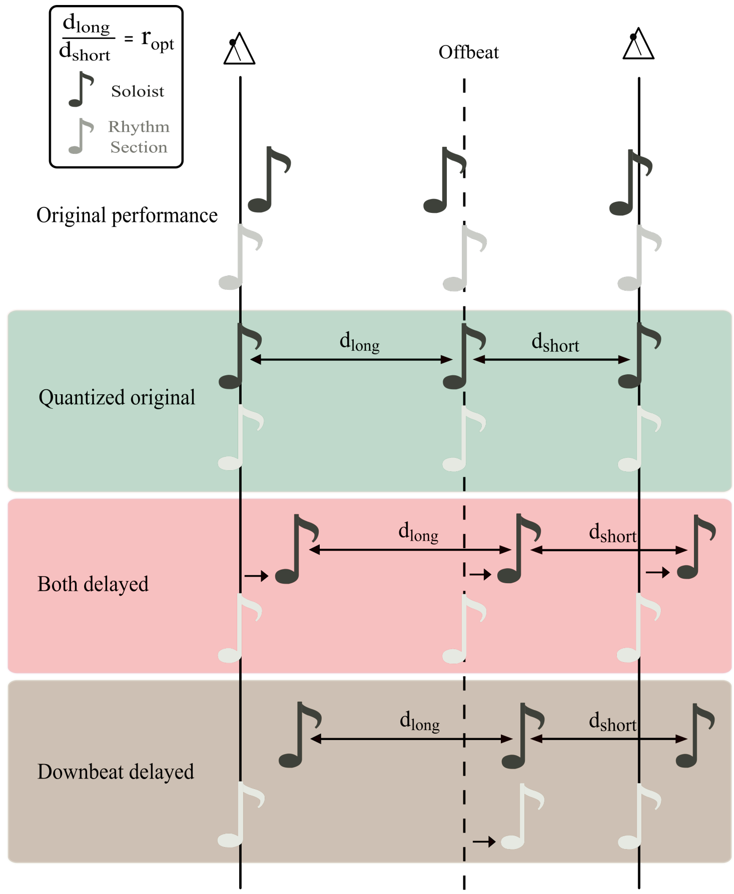


Figure 3: **Timing manipulations**. Schematic representation of the timing manipulations we used in the experiment to probe the effects of microtiming deviations on the swing feel. Importantly, all manipulations were done so as to keep the same swing ratio for the soloist (i.e. piano). Full lines represent exact quarter note positions (metronome beats). The dashed line shows the position of the offbeats corresponding to a chosen "optimal" swing-ratio, referred to as $r_{opt}$ in the upper-left frame. Black notes and grey notes denote timing positions of soloist and rhythm section, respectively, in the different manipulations. In the **quantized original** version (green background) underlying all further manipulations, the microtiming deviations of the soloist's original performance are suppressed and the notes are aligned with the grid. In the **both delayed** version (red background), all notes of the soloist are delayed by 85 ticks. Finally, in the **downbeat delayed** version (brown background), additionally the offbeats of the rhythm section are synchronized with the offbeats of the soloist. This procedure creates downbeat delays of 85 ticks for the soloist without changing the soloist's swing ratio, but increases the swing ratio of the rhythm section.

deviations, which are present in every human musical performance, do not enhance but rather tend to impede the swing feel. In the present work we now focus on studying the effect of *systematic microtiming deviations*.

Moreover, the analysis presented in the preceding section did not show whether musicians are also delaying their offbeats. The WJD only reports downbeats of drums as a reference, but does not give access to their offbeats. With our experimental approach, however, we are able to clarify the role of offbeat timing by studying how different versions with and without offbeat delays affect the swing feel.

We prepared audio extracts presenting different kinds of systematic microtiming deviations in jazz piano performances ("soloist") with respect to a quantized rhythm section ("rhythm section"). The

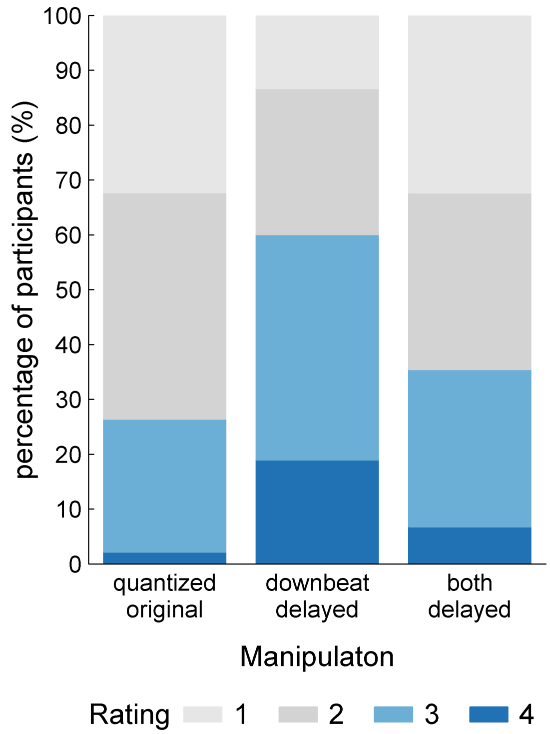


Figure 4: **Distribution of swing ratings given by professional and semiprofessional jazz musicians to different manipulated versions.** The three stacked histograms display the proportions of different possible ratings from 4 ("very much") to 1 ("not at all") averaged over three pieces. The *downbeat delayed* manipulation in the center elicits a much larger portion of high ratings (3 and 4 in blue colors) than the two other manipulations.

manipulations we carried out on real performances are explained in detail in the Methods section. and sketched in Fig. 3. We based all manipulations on a *quantized original* version, which suppresses all fluctuations (i.e. microtiming deviations) by aligning the notes to a grid. We needed to take such a step for the sake of providing well-controlled distinguishable conditions. We think that this is justified as a minor intervention, as we previously showed that random microtiming deviations do not play a positive role for the swing feel [20]. We hypothesized that (i) a *both delayed* manipulation, where all notes of the soloist are uniformly delayed with respect to the rhythm section, and/or (ii) a *downbeat delayed* manipulation, where the soloist notes are delayed apart from the offbeats (which are synchronized with the rhythm section) might have a positive effect on the swing feel.

In order to measure the swing feel using an operational definition, we presented the manipulated audio extracts of three different pieces ("The smudge", "Texas blues", "Jordu") to professional and semiprofessional jazz musicians in an online experiment. Separately for each piece participants compared all three manipulations with each other and rated the perceived swing feel and the perceived groove on different scales from 1 ("not at all") to 4 ("very much"). The results show that professional and semiprofessional jazz musicians gave the highest swing ratings to versions with delayed downbeats and synchronized offbeats (i.e. the *downbeat delayed* version). This is apparent in the average distribution of swing ratings across three pieces shown in Fig. 4 as well as in Supplementary Fig. 9. One can see that the *downbeat delayed* version obtained a large proportion of high ratings (3 and 4, blue colors) while the *quantized original* or *both delayed* versions received considerably smaller fractions of high ratings. The results on the groove ratings show a similar pattern with considerably smaller effect sizes of the manipulations (see Supplementary Information Sect. 2.4). It is also interesting to note that professional

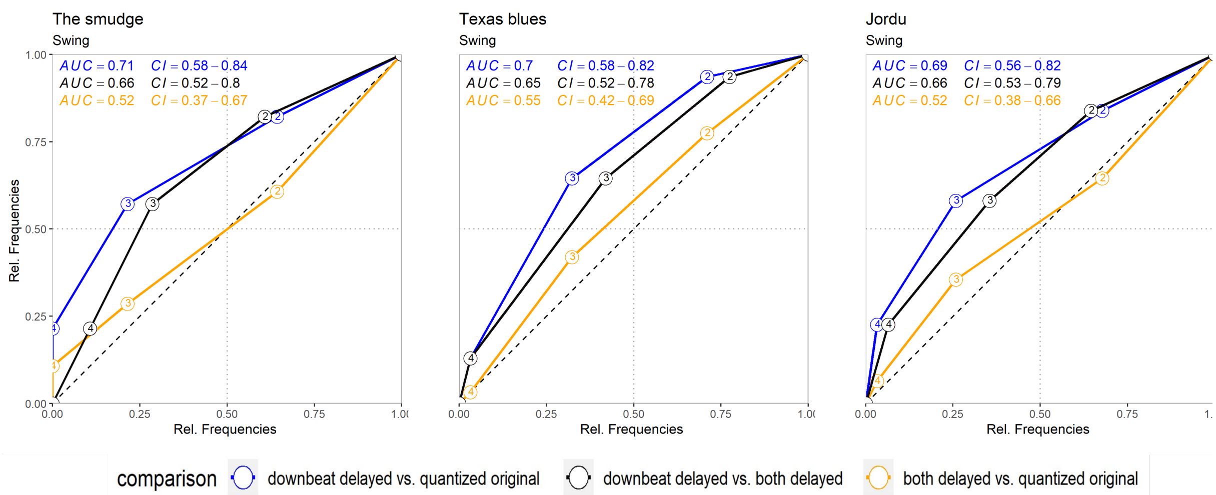


Figure 5: **Receiver Operating Characteristic (ROC) curves for the swing ratings of three pieces.** These curves compare cumulative proportions of the ratings 4 to 1 for two conditions mapped along the horizontal and the vertical axis, i.e. two of the stacked histograms of Fig. 4 are plotted against each other along an axis each. A deviation from the diagonal to either side indicates higher swing ratings for one of the conditions and shows that listeners discriminate between the versions. The area under the curve (AUC) quantifies the deviation from the diagonal ($AUC = 0.5$ means no discrimination). Comparison of the *downbeat delayed* with the *quantized original* manipulation (blue curves) and with the *both delayed* manipulation (black curves) shows a significant preference for the *downbeat delayed* versions. Statistical significance is confirmed for these two curves by the AUC confidence intervals (CI), which do not contain $AUC = 0.5$.

| Condition | OR | CI | $p$ |
|---|---|---|---|
| **DD** vs **QO** (condition 1) | 7.48 | [3.19 ; 17.54] | $< .001$*** |
| **BD** vs **QO** (condition 2) | 1.30 | [0.67 ; 2.54] | .440 |
| **semi-pro** vs **pro** (muscat) | 3.94 | [1.26 ; 12.31] | .019* |
| **condition 1** × **muscat** | 7.00 | [1.32 ; 37.11] | .022* |
| **condition 2** × **muscat** | 2.33 | [0.62 ; 8.82] | .212 |

Table 1: **Results of ordinal logistic regression for swing ratings.** OR denotes odds ratios for the comparison of different manipulations, musician categories, and their interaction, CI are 95% confidence intervals for the OR. The $p$ values are marked by *** for $p < .001$ and by * for $p < .05$. Odds ratio $OR = 7.48$ for DD vs. QO means that it is about 7.5 times more likely that jazz musicians perceive the *downbeat delayed* manipulation (DD) as more swinging than the *quantized original* (QO). BD refers to *both delayed*, pro and semipro refer to jazz musician categories.

musicians gave overall lower ratings than semiprofessionals, which is noticeable in particular for the highest rating in the downbeat delayed version (6.5% vs 31.4% for professionals and semiprofessionals, respectively; see Supplementary Fig. 9). We made a similar observation in our earlier study [20]. This finding probably reflects the higher standards and expectations of professional musicians. An ordinal logistic regression of the swing ratings upon manipulation, musician category, and their interaction statistically confirmed the results described above (cf. Table 1). The *downbeat delayed* versions received significantly higher swing ratings than *quantized original* versions not having any delays ($p < .001$). No significant difference was observed comparing the swing ratings of the *both delayed* versions to those of the *quantized original* versions ($p = .440$). Professional jazz musicians gave overall lower ratings than semiprofessional musicians ($p = .019$). In addition, the effect of the *downbeat delayed* versions (vs. *quantized original*) was larger for semiprofessionals than for professionals ($p = .022$). The odds ratios as well as their associated confidence intervals for the different conditions are summarized in Table 1. The odds ratio of the *downbeat delayed* versions compared to the *quantized original* versions was 7.48. In other words, delaying the soloist's downbeats while synchronizing the offbeats makes it more than *seven times* more likely that jazz musicians perceive the recording as more swinging than the *quantized original*. To further validate this effect we performed three additional checks to analyze the statistical power and to test for potential effects of outliers and sample size (see Supplementary Information, Sect. 2.2.1). They yield a very high statistical power together with a high robustness of the effects.

To elucidate the discriminability between the different manipulations, we determined *receiver operating characteristic* curves (ROC) for each piece in Fig. 5. These ROC curves compare the cumulative proportions of the four ratings for two conditions mapped along the horizontal and the vertical axis, i.e. two of the stacked histograms of Fig. 4 are plotted against each other along an axis each. A deviation from the diagonal to either side indicates higher swing ratings for one of the conditions and shows that listeners discriminate between the versions and perceive one of them as more swinging. The area under the curve (AUC) quantifies the deviation from the diagonal ($AUC = 0.5$ means no discrimination) and is an effect size that can be tested for significance. The effect is statistically significant, if 0.5 is not within the AUC confidence interval (CI). Comparing the *downbeat delayed* to the *quantized original* manipulations (blue curves in Fig. 5) shows a clear preference for the *downbeat delayed* versions with significant AUC values for all three pieces: $AUC_{\text{The smudge}} = 0.71 \pm 0.13$, $AUC_{\text{Texas blues}} = 0.70 \pm 0.12$ and $AUC_{\text{Jordu}} = 0.69 \pm 0.13$. Comparing the *downbeat delayed* and *both delayed* manipulations (black curves in Fig. 5) also shows a preference for the *downbeat delayed* versions with significant AUC values. By contrast, the yellow curves and their AUC values display no significant difference between the *both delayed* and *quantized original* versions. Taken together, these findings imply that delaying the soloist's downbeats while synchronizing offbeats has a significant positive impact on the swing feel, whereas uniformly delaying all soloist notes does not.

Note that the experiment also included a fourth piece ("Serenade to a Cuckoo"), which is discussed in detail in the Supplementary Information, Sect. 3. The swing ratio is an important parameter affecting the swing feel and it is critical to optimally adjust its value. An ill-chosen value can lead to inconclusive results, since our *downbeat delayed* manipulation increases the swing ratio of the rhythm section (see Fig. 3 and its caption). In our experiment we originally chose a non-optimal swing ratio for "Serenade to a Cuckoo", which resulted in a very large swing ratio of the rhythm section of 2.91 in the *downbeat delayed* manipulation. We therefore conducted a second experiment on this fourth piece testing the influence of different swing ratios. The results of the second experiment showed that for an optimally chosen swing ratio there was a significant preference for the *downbeat delayed* condition (Supplementary Information, Sect. 3).

## Discussion

The research presented in this article aimed at identifying systematic microtiming deviations in recorded jazz solos and clarifying their possible role for the swing feel in jazz. Our observational study analyzing more than 400 recordings showed that downbeat delays, although piece and player dependent, are used by many jazz soloists and follow a clear tempo dependent trend with increasing delays for decreasing tempo.

To find out whether these downbeat delays are relevant for the swing feel, we conducted an experimental study. In lack of a generally accepted definition of swing, clarifying their role was only possible based on an operational definition by measuring the swing feel through ratings of musical experts. This approach required introducing a number of simplifications. In particular, we used a quantized original version as a well-defined starting point for manipulating the recordings. Another simplification was to consider a solo instrument, a piano, playing on top of a quantized rhythm section. Moreover, we focused on pieces with many downbeat–offbeat pairs, which are prominent in jazz music, in order to study the role of their microtiming. As soloists sometimes vary their playing style within a piece or even within a solo, it was necessary and worthwhile to make such simplifications, in order to reveal general trends.

Our experimental study yielded the clear and significant result that soloists delaying their downbeats while synchronizing their offbeats with the rhythm section do enhance the swing feel. This is the first time a positive impact of certain systematic microtiming deviations on the swing feel was shown. While the authors of early observational studies [11, 13] investigating downbeat timing could only speculate that observed downbeat delays play a role in the swing feel, our study provides direct evidence of their positive effect.

Our findings are of interest to various fields, from the physics of social interactions and human behavior to psychoacoustics and the perception of musical rhythms. They also have implications for music education and music production. Many modern digital audio workstations offer options for "swingifying" computer-generated music. So far these features are of limited value, as they mainly serve to introduce a suitable swing ratio. Adding downbeat delays according to our findings would help improve these features for digital music production. The question might come up, whether downbeat delays are specific for swinging jazz, or whether they show up more generally. We therefore also carried out timing analyses of latin music of various origins, e.g. The Latin Pianist by PG-Music [29]. We found that downbeat delays, where they occur, are very small, sometimes negative, and are mostly below the threshold of timing accuracy.

Other papers have investigated the role of microtiming deviations for *groove*, defined as the impetus to move along with the rhythm. Groove can exist in music without swing, but it is generally accepted that swing requires groove. In our study, we also asked about the perceived groove feel (see Supplementary Information Sect. 2.4). It appears that swing ratings are influenced more strongly by the *downbeat delays* than groove ratings, hinting at a partial dissociation of these concepts. Validating this dissociation, however, will require further experimental studies.

Considering future research, there is plenty of room for lifting the restrictions we imposed in our study. For instance, we did not study ensemble microtiming such as in big band performances, but we believe that similar mechanisms are at work as in the case of soloists. Furthermore, jazz soloists also typically use other rhythmic values, e.g. 8th-note triplets among others. Future work should aim at elucidating trends of microtiming deviations for these other rhythmic values.

It is important to note, that the downbeat delays of the order of 30 ms studied here are not related to the well-known laid-back style that is occasionally applied by jazz musicians. These small downbeat delays were not perceivable as such by professional jazz musicians in the recordings. The much larger delays in laid-back playing – on the other hand – are easily perceivable and are also applied to offbeats. In their comments to the online study, some professional jazz musicians reported that they could perceive a pleasant friction between soloist and rhythm section, but were amazed that they could not determine its nature. They apparently could ”feel it”, but they just couldn’t ”explain it”.

# Methods

## Timing analysis of jazz solos

To clarify the question, whether jazz musicians apply systematic MTD, we analyzed 456 jazz solos from the Weimar Jazz Database [21]. The transcriptions in this database were obtained using the sonic visualizer software, a software enabling precise spectral visualization of audio data [30]. In the database, each note of a piece is stored as a collection of entries describing its properties. Features like pitch, onset (in ms) and quarter note subdivision are available, as well as the position (in ms) of the measure to which it belongs. This detailed representation permits computation of mean quantities such as average downbeat and offbeat position as a function of tempo. To determine the root mean squared error of the note positions in the database, we performed two successive transcriptions of John Coltrane’s ”Giant steps” with the help of sonic visualizer and compared the obtained values. We found the root mean squared error of the note positions to be around $20ms$. This reflects the human nature of the transcriptions and should be taken into account when considering the variance in our results. We converted our results to *ticks*, the unit of time in MIDI format, as ticks time units are required to manipulate MIDI files. The conversion is done via:

$$t_{ms} = 1000 \times \frac{t_{ticks}}{tpq} \times \frac{60}{tempo} \tag{1}$$

Where *tpq* is a MIDI variable representing the number of *ticks per quarter notes* in the piece, its default value is 960. We also computed the mean *swing ratio* as a function of tempo for the pieces of the database. The swing ratio is a quantity used to measure the asymmetry of downbeat-offbeat pairs in jazz. For a triplet of three adjacent downbeat, offbeat and downbeat, it can be defined as:

$$r = \frac{p_{ob} - p_{db_1}}{p_{db_2} - p_{ob}} \tag{2}$$

with $p_{ob}$ the position of the offbeat, $p_{db_1}$ and $p_{db_2}$ the position of the first and second downbeat, resp. The mean swing ratio of a given piece is then the average of equation (2) for all such triplets in the piece.

## Timing manipulations

The code to perform the manipulations on the MIDI recordings was written in Julia, using the Music-Manipulations.jl package [31]. The mp3 audio examples were generated using the Reaper software with plugins by Native Instruments (*”The Gentleman”* piano, *”50s drummer”* drumset and the acoustic bass from the standard library). All recordings used in the present study were first quantized to a grid (see Fig. 3) whose swing ratio was adjusted to a value guided by the average swing ratios observed by Friberg and Sundström [13, 27] and our own analyses of the Weimar Jazz Database, assuming that this provides optimized swing ratios. The chosen swing ratios of the three pieces were close to the optimal values $r_{opt}$ of Fig. 2 and are listed in Table 2. In our timing manipulations, we took care not to modify the swing ratio of the piano track across versions. This ensured that the participants’ ratings reflect the interaction between the timing of soloist and rhythm section and not a difference in the soloist's swing ratios. Triplets were not manipulated as they rarely occurred in the recordings we used and their dependence on tempo as well as their relation to surrounding $8th$-notes would be harder to quantify and would distract from our main objective. After quantizing a given recording, we performed two further manipulations, as sketched in Fig. 3 in the *both delayed* and *downbeat delayed* boxes. In the *downbeat delayed* version, soloist downbeats were delayed while the offbeats of soloist and rhythm section were synchronized. The delay value was fixed to 85 ticks as this allowed us to have a common delay value for all pieces (see *Timing analysis of jazz solos* in the Methods section), taking into account the tempo range of our recordings and the variance of our observations (see Fig. 1b). In order to keep the swing ratio of the soloist constant, we first delayed the whole soloist track and then synchronized all rhythm offbeats with the delayed soloist offbeats. In the *both delayed* manipulation, the whole soloist track was delayed with respect to the rhythm section. This version allowed us to test, whether the mere presence of a delay is relevant for the swing feel, or whether the offbeat synchronicity is also crucial.

In consequence, we had three versions of each piece: the quantized original, a version with a delayed soloist, and a version where downbeats were delayed but offbeats were synchronized. In Fig. 3, these versions are called *quantized original*, *both delayed*, and *downbeat delayed*.

## Online experiments

Jazz musicians recruited through musical conservatories, universities, big bands, and personal contacts were asked to participate in an anonymous online study designed with the EFS Survey (Unipark, 2019). The software did not allow collecting any data or meta-data from participants thus guaranteeing anonymity. Participants were free to end the study at any time. Since the study aimed at professional jazz musicians, we included data from 19 semiprofessional and 18 professional musicians, who took sufficient time to rate the recordings (at least 5 minutes per piece with 3 versions). The majority of respondents were male ($n = 30$), 5 were females and 2 were participants without gender information. Mean age was 38.59 ($SD = 16.10$). Participants provided information about their current daily practice and the number of concerts played within the last year (see Supplementary Information). Non-musicians were also allowed to participate, but were notified that their results would not be taken into account. In the experiment, participants were presented with three versions of four different pieces. The three versions resulted from the manipulations described in subsection . The pieces consisted of live MIDI recordings of two professional jazz pianists to which a standard swing drum track was added, which consisted of $8th$-notes on the ride cymbal and hi-hat hits on the 2 and 4 beat of each measure. A bass-line was also included, playing quarter notes to outline the harmonic background. The pieces "Serenade to a cuckoo" and "Jordu" were recorded in our accoustics lab on a Kawai ES7 keyboard, the other two "The smudge" and "Texas blues" were performed by Miles Black on the PG music "*Oscar Peterson multimedia CD*". These pieces were chosen because of their large number of $8th$-notes, maximizing the effects we seek to study. As noted above, one of the pieces of the experiment (*"Serenade to a Cuckoo"*) originally had an ill-chosen swing ratio, which led to inconclusive results due to a very large swing ratio of the rhythm section of 2.91 in the *downbeat delayed* manipulation. We therefore did not include *"Serenade to a Cuckoo"* in the analysis and ran an additional study for this piece to clarify the influence of different swing ratios (see Supplementary Information, Sect. 3 for details and findings).

All Participants were asked to use headphones to better perceive the subtle differences between versions and to minimize external noise. Before starting the experiment, participants were reminded the notions of *groove* and *swing* and were given an audio example for each in order to get sensitized to the differences of these concepts. The swing feel requires groove, but there can be groove without swing.

In the main part of the study, the different pieces were presented in random order. For each piece, they were given all three manipulated versions on one page to allow them to switch back and forth between versions and listen to them as often as they liked. For each version participants had to rate the perceived groove and the swing feel on a 4-point scale ranging from "not at all" to "very much" (see Supplementary Information for results on groove ratings). When the same swing ratings were given to different versions,

| Recording | BPM | $\bar{d}$ | $\bar{r}$ | quantized r | delayed r |
|---|---|---|---|---|---|
| Jordu | 150 | -1 | 1.73 | 1.7 | 2.54 |
| Texas blues | 168 | 43 | 1.86 | 1.7 | 2.54 |
| The smudge | 175 | 42 | 1.88 | 1.65 | 2.46 |

Table 2: **Characteristic parameters of the recordings**. Tempo is given in beats per minute (BPM). $\bar{d}$ is the average position of downbeats in the original recordings in ticks and $\bar{r}$ is the mean swing ratio of the original recordings. Quantized r represents an optimized swing ratio with which we quantized soloist and rhythm section before the manipulations (see Fig. 3). Delayed r is the swing ratio of the rhythm section after the *downbeat delayed* manipulation.

participants were asked in addition, whether they could perceive a difference between the versions.

## Statistical Analyses

To analyze the participants' swing ratings and their dependence on the manipulations and the categories of musicians, we performed an ordinal logistic regression. It takes into account the ordinal nature of our 4-point rating scale, with ordered but unstructured thresholds for our four response categories [32] and was based on the statistical model

$$\log\left[\frac{\Pr(\,rating \le j\,)}{1-\Pr(\,rating \le j\,)}\right] = \alpha_j + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 \tag{3}$$

Variables $x_1$ and $x_2$ represent the manipulation and musician category, $x_1 x_2$ their interaction. The list of all parameters for the model can be found in the Supplementary Information. The *quantized original* version and semiprofessional jazz musicians were chosen as reference categories to be compared to the other versions and musician category. The resulting odds ratios indicate how much more likely it is that the respective version elicits higher swing ratings than the *quantized original* version. A value larger than 1 signifies a higher probability, a value lower than 1 a reduced probability. For detailed analyses of statistical power and robustness of the effects see Supplementary Information, Sect. 2.2.1.

To analyze whether and how participants discriminated between the different versions of a piece, we determined receiver operating characteristic curves (ROC) and computed areas under the curve (AUC). The AUC is a measure of discriminability and an effect size. In a ROC analysis, one version is assigned to the abscissa, the other to the ordinate. The ROC-curve reflects the cumulative frequencies of each rating category starting from 4 (very much) to 1 (not at all). A diagonal line results, if participants do not differentiate between versions or have no preference. The more the ROC-curve deviates from the diagonal to the top, the higher the ratings for the version assigned to the ordinate in comparison to the version assigned to the abscissa. For the discriminability of the versions to be significant, the ROC-curve must deviate substantially from the diagonal, that is, the confidence interval of the AUC must not include 0.5 (the area under the diagonal) [33].

Ethics statement. All experimental procedures adhere to the Ethical Principles of the American Psychological Association [34] and are in full accordance with the guidelines of the local ethics committee and federal regulations. All participants were fully informed about the aims and procedures of the study, and gave informed consent before participating in the survey. The study was conducted anonymously.

## References


[1] Treadwell B. Big Book Of Swing. Cambridge House; 1946.

[2] Rendel DM. The New Harvard Dictionary Of Music. Harvard University Press; 1986.

[3] For swing modifications of audio signals see e.g. Gouyon F, Fabig L, Bonada J. Rhythmic expressiveness transformations of audio recordings: swing modifications. Proc. Digital Audio Effects Workshop. 2003

[4] Keil C. Participatory discrepancies and the power of music. Cultural Anthropology. 1987;2(3):275–283.

[5] Keil CM. Motion and feeling through music. Journal of aesthetics and art criticism. 1966; p. 337–349.

[6] Prögler JA. Searching for swing: Participatory discrepancies in the jazz rhythm section. Ethnomusicology. 1995;39(1):21–54.

[7] Mermikides M. Changes over time: Theory and practice. University of Surrey.; 2010.

[8] Kilchenmann L, Senn O. Microtiming in swing and funk affects the body movement behavior of music expert listeners. Frontiers in psychology. 2015;6:1232.

[9] Matsushita S, Nomura S. The asymmetrical influence of timing asynchrony of bass guitar and drum sounds on groove. Music Perception: An Interdisciplinary Journal. 2016;34(2):123–131.

[10] Hofmann A, Wesolowski BC, Goebl W. The tight-interlocked rhythm section: Production and perception of synchronisation in jazz trio performance. Journal of new music research. 2017;46(4):329–341.

[11] Ellis MC. An analysis of "swing" subdivision and asynchronization in three jazz saxophonists. Perceptual and Motor Skills. 1991;73(3):707–713.

[12] Iyer V. Embodied mind, situated cognition, and expressive microtiming in African-American music. Music perception. 2002;19(3):387–414.

[13] Friberg A, Sundströöm A. Swing Ratios and Ensemble Timing in Jazz Performance: Evidence for a Common Rhythmic Pattern. Music Perception. 2002;19(3):333–349.

[14] Dittmar C, Pfleiderer M, Balke S, Müller M. A swingogram representation for tracking micro-rhythmic variation in jazz performances. Journal of New Music Research. 2018;47(2):97–113.

[15] Butterfield M. Participatory discrepancies and the perception of beats in jazz. Music Perception. 2010;27(3):157–176.

[16] Davies M, Madison G, Silva P, Gouyon F. The effect of microtiming deviations on the perception of groove in short rhythms. Music Perception: An Interdisciplinary Journal. 2013;30(5):497–510.

[17] Frühauf J, Kopiez R, Platz F. Music on the timing grid: The influence of microtiming on the perceived groove quality of a simple drum pattern performance. Musicae Scientiae. 2013;17(2):246–260.

[18] Madison G, Sioros G. What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. Frontiers in Psychology. 2014;5:894.

[19] Senn O, Kilchenmann L, Von Georgi R, Bullerjahn C. The effect of expert performance microtiming on listeners' experience of groove in swing or funk music. Frontiers in Psychology. 2016;7:1487.

[20] Datseris G, Ziereis A, Albrecht T, Hagmayer Y, Priesemann V, Geisel T. Microtiming deviations and swing feel in jazz. Scientific reports. 2019;9(1):1–10.

[21] Pfleiderer M, Frieler K, Abeßer J, Zaddach WG, Burkhart B. Inside the Jazzomat - New Perspectives for Jazz Research. Schott Campus; 2017.

[22] Polak R, London J, and Jacoby N. Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus study of Malian jembe drumming. Frontiers in Neuroscience. 2016;10:285.

[23] Polak R., Jacoby N., Fischinger T., Goldberg D., Holzapfel A., and London J. Rhythmic prototypes across cultures: a comparative study of tapping synchronization Music Perception: An Interdisciplinary Journal 2018;36:1–23.

[24] Jacoby N, et al. Universality and cross-cultural variation in mental representations of music revealed by global comparison of rhythm priors PsyArXiv 2021.

[25] Gouyon, F., Fabig, L. Bonada, J. Rhythmic expressiveness transformations of audio recordings: swing modifications. In Proc. Digital Audio Effects Workshop (DAFx) 2003.

[26] Honing, H., De Haas, W. B. Swing once more: Relating timing and tempo in expert jazz drumming. Music Perception 2008;25:471–476.

[27] Friberg, A., Sundström, A. Preferred swing ratio in jazz as a function of tempo. TMH-QPSR 1997;38:019–027.

[28] Dittmar, C., Pfleiderer, M., Müller, M. Automated estimation of ride cymbal swing ratios in jazz recordings. 2015; p. 271–277.

[29] The Latin pianist, Computer Software Program, PG Music inc. 1998.

[30] Cannam, C., Landone, C., Sandler, M. Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In Proceedings of the ACM Multimedia 2010 International Conference 2010. 2010; p. 1467–1468.

[31] Datseris, G., Hobson, J. MIDI.jl: Simple and intuitive handling of midi data. In The Journal of Open Source Software 2019;4:1166.

[32] McCullagh, P. Regression models for ordinal data. In Journal of the Royal Statistical Society: Series B (Methodological) 1980;42:109–127.

[33] DeLong, E. R., DeLong, D. M., Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988; p. 837–845.

[34] American Psychological Association. Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2017). http://www.apa.org/ethics/code/index.html

## Acknowledgements

We would like to thank Uwe Meile for providing part of the piano recordings we used in our experimental study. We are indebted to Tim Dudek for invaluable help, providing expert advice in many stages and throughout this research, as well as for mixing and mastering our audio samples in his studio. C.N. thanks George Datseris for sharing his expertise from a previous project. We also thank Ulrich Nicolai for advice and for his help in spreading our online study, and all participating musicians for their collaboration. Finally, we thank PG music for allowing us to publicly share the extracts used in our experiment.

## Author contributions statement

TG: conceived the study
TA, YH, TG: designed research
CN, EMS, TA, YH, TG: performed research
CN, EMS: analyzed data
CN, EMS, TA, YH, TG: wrote the paper.

## Competing interests

The authors declare no competing interests.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationDownbeatdelays.pdf