Listener preferences for alternative dynamic-range-compressed audio configurations

Campbell, William, Paterson, Justin ORCID logoORCID: https://orcid.org/0000-0001-7822-319X and van der Linde, Ian (2017) Listener preferences for alternative dynamic-range-compressed audio configurations. Journal of the Audio Engineering Society, 65 (7/8). pp. 540-551. ISSN 1549-4950

10.17743/jaes.2017.0019

Listener preferences for alternative dynamic-range-compressed audio configurations

William Campbell, [1] *AES Member*, **Justin Paterson,** [2] **AND**
(william.campbell@anglia.ac.uk) (justin.paterson@uwl.ac.uk)

Ian van der Linde, [1]
(ian.vanderlinde@anglia.ac.uk)

[1] *Department of Computing and Technology, Anglia Ruskin University, Cambridge, UK*
[2] *London College of Music, University of West London, London, UK*

Some audio experts have proposed that using Dynamic Range Compression (DRC) to increase the loudness of music compromises audio quality. Conversely, in listening tests, researchers sometimes find that audio subjected to DRC is preferred over uncompressed audio. We test the hypothesis that it is DRC configuration, rather than the use of DRC per se, that determines listener preferences. In this study, 130 listeners completed 13 A/B preference trials using pairs of RMS loudness-equalized stimuli subjected to different DRC configurations: *viz*., two magnitudes (heavy, moderate) and two compression types (limiting, compression) applied at three different points in the mix chain (track, subgroup, and master buss, here termed full-sum), along with an uncompressed control stimulus. Our results suggest that listeners prefer audio in which moderate compression has been applied to fewer signals simultaneously and dislike heavy limiting, particularly when applied to the full-sum, presumably because heavy DRC (and particularly limiting) applied to pre-mixed signals produces disagreeable distortion or because tracks whose amplitude characteristics would not have reached the DRC threshold alone may be deleteriously affected (e.g., attenuated) as a consequence of amplitude peaks in other tracks with which they are grouped.

0 INTRODUCTION

Dynamic range compressors are ‘waveshapers’ that violate the homogeneity and additivity requirements of a linear system and produce intentional or unintentional harmonic and/or inharmonic frequency components not present in the source signal [1], [2]. Moore *et al.* [3] describe the perceptual effects of linear distortion in terms of changes in timbre, tonality or “coloration”, and the perceptual effects of nonlinear distortion in terms of “harshness”, “roughness” or “noisiness”.

Two overlapping domains in which Dynamic Range Compression (DRC) research is conducted are speech intelligibility in hearing aid design (see [4]–[8] for more information) and controlling the inherent loudness of musical signals [9], potentially at the expense of audio fidelity [10]–[19]. The algorithms and optimal settings in each domain differ. Our research focuses on the use of DRC for controlling the loudness of musical signals, rather than aesthetic effects like gain pumping [20], but also exploits current relevant hearing aid research.

Evidence in support of the ‘louder is better’ argument [12] for audio includes experimental results indicating that louder oration is more compelling [21], and that a relationship between loud music and elicited emotion exists [16], [22]. Schubert devised a computer-based experiment that tracked the perceived emotion conveyed by four romantic musical pieces using 67 volunteers that spanned a wide demographic range of age, gender, and musical expertise [22]. Listeners continuously reported the emotion that the music was intended to express (rather than the emotion the listener may have felt) by moving a mouse cursor in a two dimensional X-Y grid, wherein the X-dimension represented degrees of emotional valence (sad to happy) and the Y-dimension represented degrees of arousal (from sleepy to aroused). The four pieces of music were chosen for their predicted fit to the quadrants of perceived emotion and the variability of their measurable musical features:

loudness, tempo, melodic contour, texture, and spectral centroid (related to perceived timbral sharpness). The study found a significant positive correlation between loudness and both valence and arousal response dimensions, although it is also likely that loudness must be modulated to educe corresponding changes in emotional affect. Wendl and Lee correlated perceived loudness to perceived audio quality, finding a significant positive correlation for rock and jazz music, but no correlation for electronic music [13], suggesting that loudness impacts quality assessment ratings, but also that this relationship may be genre-specific

It is thought that, when used injudiciously, DRC can negatively affect audio quality [16], [21]. However, the question of whether listeners prefer objectively higher quality audio is not straightforward, and may be influenced by both long-term familiarity (which one might refer to as expertise) and short-term training. In a classic study, Kirk [23] examined the audio fidelity preferences of 210 college students over a 6–7 week period in the mid 1950s. The students' preferences were measured using A/B/A tests, in which the frequency spectra of five diverse high-fidelity phonograph recordings (string quartet, symphony orchestra, organ popular music and male speech) were altered using four band-pass filters of differing widths (180–3,000 Hz, 120–5,000 Hz, 90–9,000 Hz, 30–15,000 Hz), producing six comparisons per stimulus. Surprisingly, Kirk found that, participants *least* preferred the unrestricted (objectively higher quality) recordings (30–15,000 Hz), instead preferring a narrower frequency spectrum.

In a follow-up experiment, the same participants were divided into two groups, with one group invited to thirteen 40-minute listening sessions using unrestricted stimuli similar to those from the original experiment. The original listening test was then repeated, and it was found that this group had altered their preferences, and began to select higher fidelity recordings significantly more frequently. The second follow-up group listened to band-limited recordings (180–3000 Hz) and surprisingly, then expressed a significant preference for that configuration when the original test was repeated.

Together, these results suggest that untrained listeners who prefer lower-fidelity audio may do so because this matches their prior listening experiences (this study was conducted in the era of low-quality AM-radio), and that listening preferences can be manipulated in either direction (i.e., either towards objectively higher fidelity, or objectively lower fidelity recordings) by habituation (which one might refer to as training, although the preference change educed may be unconscious and not the result of guidance or feedback from the experimenter). It may be that DRC effects that negatively impact *objective* audio quality may be *subjectively* preferred by some listeners, perhaps due to long-term exposure to this process, and raises the possibility that only trained/expert listeners will find DRC-induced distortion disagreeable.

Olson [24], seemingly in contradiction to Kirk [23], found that untrained listeners *did* prefer an unrestricted frequency range, but this study related to live musical performances, as opposed to recorded music [23], which left the question of fidelity preferences in untrained listeners rather inconsistent, and seemingly contingent upon the environment in which the music was heard.

More recently, in Olive [25], [26] (additional participants added in latter article), the listening preferences of 18 school children and 40 college students were compared to those of a commercial speaker manufacturer's trained listening panel. The college students comprised pre-trained (expert) and untrained listeners. In experiment 1, four recordings (clapping, female vocal with guitar, female vocal with strings, and female vocal with orchestra) were played at both 128 kbit/s MP3 and CD-quality. Participants completed 12 counterbalanced A/B trials. In each trial, participants heard the MP3 and CD-quality reproductions of each recording four times. Olive found a significant preference for CD-quality over MP3 overall ($p < .05$). Around 80% of students with the most listening expertise preferred CD-quality audio, whilst only around 65% of non-expert students expressed this preference. These data also seem to support Kirk's findings [23], suggesting that listening preferences can be manipulated with habituation/training, since those without listening expertise expressed a lower rate of preference for objectively higher quality recordings. However, there is also likely to be a relationship between expertise and general appreciation for audio fidelity, which may have mediated these results (e.g., it may be that untrained listeners could not hear a noticeable difference, or were simply less interested in music and therefore less attentive during the study).

In experiment 2 of Olive [25], [26], two CD-quality commercial popular music recordings, and four different commercial loudspeaker systems were tested, alternating playback from each in a random order until a listener preference was submitted. A preference for the most accurate, frequency-neutral (implicitly higher quality) loudspeaker system was found over loudspeakers with a measured acoustic performance that deviated from an ideal frequency response (implicitly lower quality). Although examining quite different things (audio source quality *vs*. speaker fidelity), these results are in contrast to Kirk [23], in which untrained listeners were found to prefer poorer fidelity music reproductions consistent with their prior listening experiences, since participants' prior listening experience in Olive [25], [26] was most likely to have been with poorer quality commercial speakers built for personal use (i.e., not professional reference speakers).

Reiss [27] conducted a systematic review of 18 studies published after 1980 that evaluated listeners' ability to discriminate high *vs*. standard resolution audio (defined in terms of bit depth and sample rate). It was reported that the ability of untrained listeners to discriminate between resolution was surprisingly poor, but did improve significantly after training, supporting the notion that an ability to discriminate between audio quality settings can be learned.

Similarly to the present study, Croghan, Arehart and Kates [11] examined the impact of DRC on perceived audio quality, along with loudness. Two uncompressed 13 s recordings from rock and classical genres served as source stimuli. DRC was applied to the final (summed) signal for each recording, yielding six compression thresholds (uncompressed, –8, –12, –16, –20, and –24 dBFS). Two versions of each stimulus were created, one in which loudness was not equalized between stimuli (UNEQ), and another in which loudness was equalized

(LEQ). Twenty three participants rated stimuli played in randomized pairs, within genre, on the metrics of preference, loudness, pleasantness, and dynamic range. Six hours of testing, spread over multiple visits, were completed. In both the UNEQ and LEQ conditions, the effect of compression was found to have a significant effect, for both rock and classical music stimuli, on loudness, dynamic range, pleasantness, and preference ratings (all $p \leq .01$). More specifically, these results indicated a genre-independent preference for light DRC over no DRC in the UNEQ condition. The dynamic range ratings for stimuli subject to DRC were markedly lower, showing that the principal outcome of DRC was noticeable. Heavy DRC was also found to decrease the likelihood of these stimuli being selected as preferred in both UNEQ and LEQ conditions, and to yield a commensurate reduction in pleasantness ratings.

Kates and Arehart [28] analyzed signals using the Hearing Aid Speech Quality Index (HASQI) system, modified to fit the quality ratings for musical signals reported in [29]. They measured the degree to which signal quality was affected by DRC by comparing the amplitude envelope and frequency spectra of processed and unprocessed signals, yielding a number between 0 (very low fidelity) and 1 (perfect fidelity). A reduction in fidelity with increased DRC was reported; however, no statistical analysis relating HASQI measurements to listener preference was provided, limiting the degree to which we may generalize from these findings.

Hjortkjær and Walther-Hansen [30] conducted listening tests that compared the original releases of several popular music recordings with subsequently remastered versions, on the premise that the peak-to-average ratio was smaller on the remasters (implying greater use of DRC). Test stimuli comprised fifteen 15 s clips, each from a different track, along with their remastered counterparts. All stimuli were CD-quality, chosen based on the anecdotal alleged inferior sound quality of the remastered versions. Loudness-equalized clips were played to 22 university music students, all naïve to the purpose of the study. The clips were heard in A/B combinations (master *vs*. remaster), and the participants were asked to indicate their preference. Each A/B pair was presented to subjects twice in the session, with the order of presentation randomized between trials. Subjects were permitted to replay the pairs as many times as necessary to finalize their decision. A two-sided sign test was used to determine whether the proportion of listeners that preferred each version deviated from chance. No significant preference for either version was found ($p = .12$, with 54.4% preferring the original version). However, since the popular musical stimuli used may have been known to participants *a priori*, the interpretation of putative DRC effects may have been complicated by preferences for familiar reproductions of these stimuli.

Furthermore, Hjortkjær and Walther-Hansen acknowledge that there were likely to have been other unquantified production differences between each pair of stimuli, such as equalization and spacialization. The authors report that "…subjects changed their preference with the presentation of the same stimulus on 46% of the trials". This was interpreted as an inability of participants to accurately differentiate between stimuli, but it may also relate to insufficient statistical power (the study was based on a relatively small sample size, and the non-parametric sign test [31] is known to be underpowered relative to its alternatives), or other limitations in experimental design (see above).

Studies examining listener preferences for music subject to DRC have typically applied this process to the summed signal [4], [11], [18]. However, the best point in the signal chain to apply DRC is also the subject of some debate. Ronan *et al.* [32] asked ten award-winning mix engineers to complete a 21-item questionnaire about track subgrouping choices, using questions formulated using sound engineering literature [33]–[35]. Responses were analysed using thematic analysis [36]. All subjects expressed a preference for the application of DRC to drum and vocal subgroups, assembled to maintain good gain structure (rather than purely for organization), and stated that pre-DRC subgrouping decisions were made that were contingent upon musical genre. Similarly, in both Ronan et al. [32] and Pestana and Reiss [18], subgroup or full-sum compression were reported to be preferred by sound engineers/audio mixing students, despite that one might expect DRC applied to pre-mixed signal groups to be more deleterious to overall quality as a consequence of the amplitude peaks in some signals in a group activating the compressor, and thereby affecting other signals with which they are grouped that would not have activated the compressor otherwise.

Wendl and Lee examined the relationship between DRC, loudness and perceived quality, but examined limiting-type DRC only and focussed upon the relationship between signal crest factor and quality between genres [13]. Giannoulis et al describe an experiment comparing the DRC settings chosen by 16 mix engineers (9 experts and 7 amateurs) to those of an automated system; however, no listening tests were conducted to evaluate the perceived quality of the alternatives produced [9]. Ma et al did measure listener preferences for different DRC ratio and threshold settings, again to compare the choices of 15 audio engineers to an automated mixing system, but did not examiner the impact of DRC point or type [37].

Using sinusoids, Toulson et al. [1] found that applying equivalent amounts of DRC to simple *vs*. summed signals affected the severity and type of distortion introduced. Specifically, reduced nonlinear intermodulation distortion was found when DRC was applied to signals prior to summation, leading the authors to propose that the application of DRC to pre-summed signals may be detrimental to sound fidelity. However, the question of whether findings based upon sinusoidal signals translate to changes in listener preferences for real musical stimuli is unclear.

With such conflict in extant results, and limited consideration of the impact of specific DRC settings on listener preferences for musical stimuli, there is scope for further work to improve understanding of the relationship between listener preferences and DRC. In this study, untrained listeners preferentially rated musical stimuli subject to different DRC configurations to ascertain the impact, independently and in combination, of compression magnitude (moderate or heavy), type (compression or limiting), and the point in the mix chain at which DRC was applied (track, subgroup or full-sum).

1 METHOD

1.1 Apparatus

Audio stimuli were played on AKG K550 closed-back/over-ear dynamic reference headphones connected to an Apple MacBook Pro running MATLAB with the PsychToolbox/VideoToolbox extensions [38]–[40].

Audio stimuli were created using Avid Pro Tools software that incorporates the peak-sensing Compressor/Limiter III. Stimulus normalization was performed with SpectaFoo metering software configured to meet the AES Standard RMS reference.

1.2 Participants

130 participants were recruited using opportunity sampling, of which 74 were male and 56 were female (M_{age} = 22.35, SD_{age} = 5.57). All participants had normal hearing (self-reported). A subset of participants (34) were registered on a cognate (music/audio) undergraduate course, indicating above-average experience with audio, but had no specific training in DRC or audio quality assessment practices. Participants were not trained, except that their task in the experiment was carefully explained (see Section 1.4), in order that instinctive preference data could be collected, mitigating against 'demand characteristics' [41] in which participant responses are biased by perceived experimenter expectations. All participants were naïve to the purpose of the experiment, and were unpaid.

1.3 Stimuli

Stimuli were prepared using a 10 s monophonic 16-bit/44.1 kHz recording of a five-piece pop band playing together (as opposed to being overdubbed). The excerpt would not have been known to participants *a priori*, eliminating the potential for any preferences towards a familiar rendition to influence responses (*cf.* [30]). The excerpt consisted of eight tracks (bass drum, drum overhead, bass guitar, electric guitar, electric piano, viola, vocal 1, vocal 2) without DRC, equalization, or any other post-production effects, to mitigate against the possibility of unwanted artefacts. The excerpt had considerable amplitude variation (Fig. 1), ensuring that, with appropriate settings, the compressor/limiter was periodically activated.

A Pro Tools mix was configured to allow DRC to be applied at one of three points in the signal path (Fig. 2): 1] 'Track' (T), having DRC on each track only; 2] 'Subgroup' (S), having DRC on each of three subgroups (consisting of a drum subgroup, vocal subgroup, and an other instruments subgroup) only; 3] 'Full-sum' (F), having DRC on the sum of the three subgroups only.

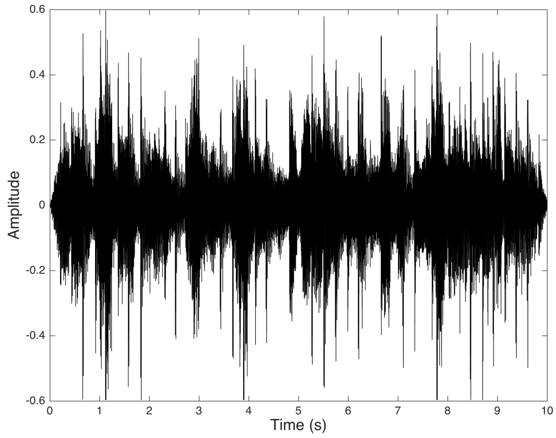

Fig. 1 Time-domain plot of source stimulus (full-sum).

The Pro Tools Compressor/Limiter III was installed on each track, subgroup, and full-sum simultaneously, in bypass mode as required, ensuring that the signal path was identical for each stimulus generated. The appropriate DRC component was activated as required to produce each of the three compression configurations described above, and deactivated otherwise.

Five combinations of DRC magnitude and type were applied in each of the mix configurations, consisting of Moderate Compression (MC), Heavy Compression (HC), Moderate Limiting (ML) and Heavy Limiting (HL), and No Compression (NC). Differentiation between compression and limiting was by compression ratio, with compression having a ratio of 1.8:1 and limiting having a ratio of 10:1, replicating the settings used by Stone *et al.* [42]. Differentiation between moderate and heavy compression/limiting was determined by the threshold above which the signal was compressed, with MC defined by 8% of the signal samples exceeding the absolute threshold, and HC by 25% of the signal samples exceeding it, again replicating the compression magnitude settings in Stone et al. [42]. An algorithm (implemented in MATLAB) was used to reverse engineer the required thresholds for MC and HC by iteratively determining the amplitude thresholds for which either 8% or 25% of the samples exceeded that threshold.

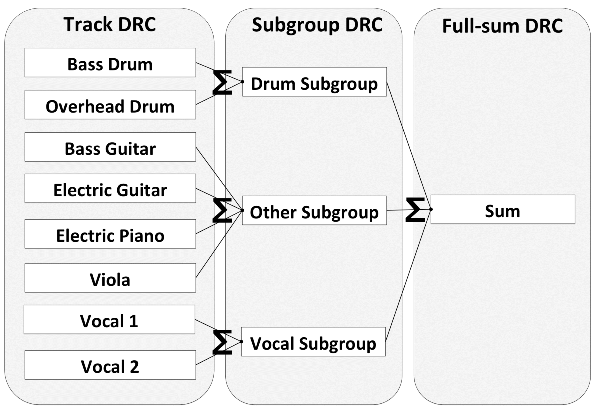

Fig. 2 Signal path for grouped stimulus production.

Next, thresholds were multiplied by $20\log_{10}(|amplitude|)$ to convert to standard logarithmic voltage units. This yielded thresholds of

-9.9dbFS (8% signal above threshold) for MC and ML, and -14.8dbFS (25% signal above threshold) for HC and HL. The resultant mean absolute amplitudes of all stimuli were within a tolerance of 1%. Stone et al. [42] used DRC settings of 0.15 ms attack and 5 ms release for spoken voice signals. Here, attack was set to 0.5 ms (with release held at 5 ms), to better suit musical stimuli.

A total of 13 stimuli were produced (Table 1): Moderate Compression Track (MCT), Heavy Compression Track (MCT), Moderate Limiting Track (MLT), Heavy Limiting Track (MLT), Moderate Compression Subgroup (MCS), Heavy Compression Subgroup (HCS), Moderate Limiting Subgroup (MLS), Heavy Limiting Subgroup (HLS), Moderate Compression Full-sum (MCF), Heavy Compression Full-sum (MCF), Moderate Limiting Full-sum (MLF), Heavy Limiting Full-sum (HLF), and No Compression (NC).

During stimulus production, RMS-normalization to -15dbFS was applied at each stage prior to DRC to ensure that signal loudness was maintained (i.e., at track, subgroup, full-sum stages). Additionally, each final stimulus defined in Table 2 (including NC) was subject to one final RMS-normalization to -15dbFS. RMS-loudness, rather than the Loudness Unit (LU) metric specificed in ITU-R BS-1770 [43], was used due to the ongoing debate concerning the suitability of LU for musical signals [14], [44], the wide availability of RMS loudness in music production software, and its consequent widespread use by professional sound engineers. However, the LUFS of our processed stimuli had a mean of -17.68 LUFS with SD = 0.09, showing very little variance between stimuli and good agreement (in terms of dispersal) between LU and RMS loudness for our stimulus set.

Table 1 DRC configuration and settings of 13 experimental stimuli.

ACRONYM	STIMULUS	DRC CONFIGURATION			DRC SETTINGS			
		Magnitude	Type	Point	Attack (ms)	Release (ms)	Ratio	Threshold
MCT	**1**	Moderate	Compression	Track	0.5	5	1.8:1	−9.9dBFS (8%)
HCT	**2**	Heavy			0.5	5	1.8:1	−14.8dBFS (25%)
MLT	**3**	Moderate	Limiting		0.5	5	10:1	−9.9dBFS (8%)
HLT	**4**	Heavy			0.5	5	10:1	−14.8dBFS (25%)
MCS	**5**	Moderate	Compression	Subgroup	0.5	5	1.8:1	−9.9dBFS (8%)
HCS	**6**	Heavy			0.5	5	1.8:1	−14.8dBFS (25%)
MLS	**7**	Moderate	Limiting		0.5	5	10:1	−9.9dBFS (8%)
HLS	**8**	Heavy			0.5	5	10:1	−14.8dBFS (25%)
MCF	**9**	Moderate	Compression	Full-sum	0.5	5	1.8:1	−9.9dBFS (8%)
HCF	**10**	Heavy			0.5	5	1.8:1	−14.8dBFS (25%)
MLF	**11**	Moderate	Limiting		0.5	5	10:1	−9.9dBFS (8%)
HLF	**12**	Heavy			0.5	5	10:1	−14.8dBFS (25%)
NC	**13**	No Compression			N/A	N/A	N/A	N/A

1.4 Procedure

Participants were seated in a secluded, low-noise, low-distraction listening space. Age, gender and audio expertise were recorded.

Each participant compared one 10 s stimulus (their *comparator*), selected using the procedure described below, to the 12 remaining 10 s stimuli, and compared the comparator to itself (a control). A sequential, two-interval forced-choice (2IFC) procedure was used [45], wherein participants were required to choose which audio excerpt they preferred (first or second) by depressing key 1 or 2 on a computer keyboard. A 0.5 s interstimulus interval was used, consistent with ITU-R BS.562-3 [46]. The order that the stimuli were presented to participants was randomized in two ways. First, the order in which the 13 stimuli were presented; second, the order in which the participants' assigned stimulus was presented for comparison to the second stimulus (i.e., first or second). This randomization procedure was intended to minimize any learning/fatigue/temporal biases caused by stimulus presentation order. Stimuli were intentionally short to enable all stimulus variants to be heard by all participants (all participants completed the experiment in under 5 minutes), and are within the ITU-R BS.562-3 recommended maximum stimulus duration of 15-20 s [46]. Each of the 13 2IFC trials began when the participant pressed a key, enabling rest breaks to be taken as desired.

Participant number was used to determine which comparator stimulus was allocated to each participant. For example, participant 1 compared stimulus 1 to each of the 13 stimuli; participant 13 compared stimulus 13 to each of the 13 stimuli. For subsequent participants, the comparator (c) was the modulus of the number of stimuli (n) by participant number (p), i.e. $c = n - p\left\lfloor\frac{n}{p}\right\rfloor$. Since 130 participants performed 13 A/B comparisons, each of the 13 stimuli served as a comparator for 10 participants.

1.5 Design

The non-parametric Pearson two-tailed chi-square goodness-of-fit test [47], shown in Eq. 1, was used to determine if observed (O) and expected (E) listener-preference counts differed significantly. Summing the addend over each of the i stimuli in each group of n stimuli, it determines if selection rate differed more than one would expect by chance (i.e., at a statistically significant level, with α = .05). Expected counts are one half of the number of times that each stimulus was presented (i.e., a per-trial chance selection probability = .5, representing the behavior of a random responder).

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \qquad \text{(Eq. 1)}$$

The dependent variable was therefore the frequency with which each stimulus was selected. Independent variables were DRC configuration (with 13 levels), which may be subdivided into magnitude (with 2 levels), point (with 3 levels), and type (with 2 levels), or combinations thereof, along with no compression.

In a closed system such as this, it is the case that if some stimuli are selected less frequently than chance, other *must* be selected more frequently than chance to balance the count. It is therefore difficult to be certain which was the cause and which was the casualty (i.e., did participants particularly dislike some DRC configurations, or particularly like others?), a general limitation of count-based analyses. However, grouped analyses enable us to probe the results further to make reasoned deductions. Therefore, additional combinatorial analyses were conducted, also justified by the postulation of Maddams [48], that the interaction of DRC magnitude and type, more so than these effects in isolation, influences perceived audio quality. In these analyses, the independent variable was again stimulus selection frequency, but was contingent upon magnitude and type in combination (pooling over point) yielding 4 levels (MC, HC, ML, HL), type and point in combination (pooling over magnitude) yielding 6 levels (CT, LT, CS, LS, CF, LF), and magnitude and point in combination (pooling over type) yielding 6 levels (MT, HT, MS, HS, MF, HF).

Where chi-square analyses were significant overall, a standard post-hoc pairwise comparison procedure was run. This established whether the observed counts of each stimuli/stimulus group differed from the expected count for that stimulus/stimulus group and the observed and expected counts of all remaining stimuli in that particular analysis, giving one degree of freedom per comparison. Bonferroni correction was used to correct for multiple comparisons, where appropriate.

2 RESULTS

Chi-square results are summarized in Table 2. Raw listener preference count data is available for download for further analysis [49].

Table 2 Results of Chi-square analyses (~ indicates marginal significance, * indicates p ≤ .05, ** indicates p ≤ .01, and *** indicates p ≤ .001.

Independent Variable	Pooling Over	*df*	*N*	χ^2	*p*	
configuration	none	12	1690	64.28	<.001	***
magnitude	point and type	2	1690	3.80	.06	~
type	magnitude and point	2	1690	13.66	.01	**
point	magnitude and type	3	1690	17.37	<.001	***
magnitude and type	point	4	1690	23.97	<.001	***
type and point	magnitude	6	1690	35.21	<.001	***
magnitude and point	type	6	1690	38.46	<.001	***

The frequency with which the 13 DRC configurations defined in Table 1 were selected as preferred was found to differ significantly from chance [χ^2(12, N = 1690) = 64.28, p < .001], Fig. 3. Post-hoc pairwise comparisons indicate that HCT and HCS were selected more frequently than expected (p < .01 and p = .02 respectively), whereas HLS and HLF were selected significantly less frequently than expected (p = .02 and p < .001 respectively). MCT, MLT, HLT, MCS, MLS, MCF, HCF, MLF and NC were not selected at a frequency that different significantly from chance (p > .05). Using Bonferroni correction for four comparisons (i.e., each stimuli giving a provisionally significant result *vs*. all remaining stimuli), the threshold for declaring significance is reduced to p < .0125, meaning that only HCT and HLF remain significant, with HCS and HLS becoming marginal. However, Bonferroni correction is known to be particularly conservative; subsequent grouped analyses allow effects to be examined using fewer pairwise comparisons.

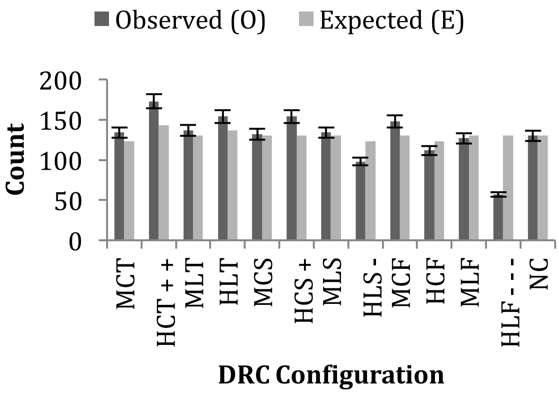

Fig. 3 Listener preferences (observed [dark gray] *vs*. expected [light gray]) for all DRC configurations. --- indicates observed < expected at p < .001, - indicates observed < expected at p < .05, + indicates observed > expected at p < .05, and ++ indicates observed > expected at p < .01. Error bars show ±5% error.

The frequency with which different DRC magnitude settings (moderate and heavy) were selected as preferred was not found to differ significantly from chance [χ^2(2, N = 1690) = 3.80, p = .06], i.e., no significant difference in DRC magnitude preference was observed when data were pooled across other independent variables (type and point). However, this result was marginal (see Section 3 for discussion), so no post-hoc pairwise comparisons were performed.

The frequency with which different DRC type settings (compression and limiting) were selected as preferred differs significantly from chance [χ^2(2, N = 1690) = 13.66, p = .01], i.e., some DRC type settings were selected significantly more or less frequently than expected when data were pooled across other independent variables (magnitude and point), Fig. 4. Post-hoc pairwise comparisons show that Compression (C) was selected significantly more frequently than expected (p < .001), whereas Limiting (L) was selected significantly less frequently than expected (p < .001). No compression (NC) was not selected at a frequency that differed significantly from chance. These findings survive Bonferroni correction for two comparisons.

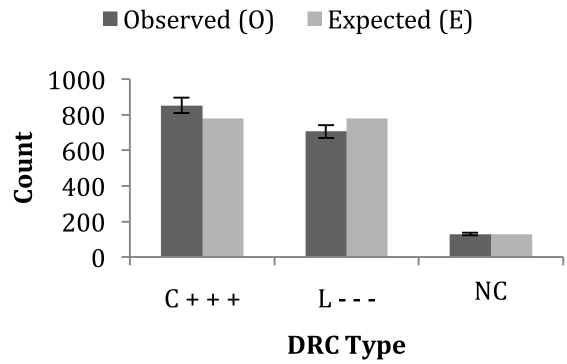

Fig. 4 Listener preferences (observed [dark gray] *vs*. expected [light gray]) for DRC type settings. (C: compression; L: limiting; NC: no compression), --- indicates observed < expected at p < .001, +++ indicates observed > expected at p < .001. Error bars show ±5% error.

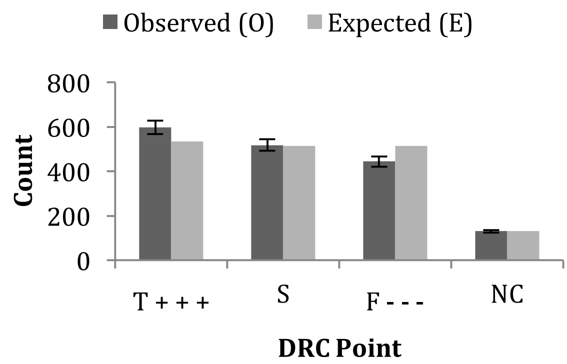

Fig. 5 Listener preferences (observed [dark gray] *vs*. expected [light gray]) for DRC point settings. (T: track; S: subgroup; F: full-sum, NC: no compression), --- indicates observed < expected a p < .001, +++ indicates observed > expected at p < .001. Error bars show ±5% error.

The frequency with which different DRC point settings (track, subgroup and full-sum) were selected as preferred differs significantly from chance [χ^2 (3, N = 1690) = 17.37, p < .001], i.e., some DRC point settings were selected significantly more or less often than expected when data were pooled across other independent variables (magnitude and type), Fig. 5. Post-hoc pairwise comparisons show that Track (T) was selected significantly more frequently than expected (p < .001), whereas Full-sum (F) was selected significantly less frequently than expected (p < .001). Subgroup (S) and No Compression (NC) were not selected at a frequency that differed significantly from chance. These findings survive Bonferroni correction for two comparisons.

The frequency with which different DRC magnitude and type combinations were selected as preferred was calculated, pooling over the other independent variable (point), producing five settings: Moderate Compression (MC), Moderate Limiting (ML), Heavy Compression

(HC), Heavy Limiting (HL), and No Compression (NC). Some settings were found to be selected at a frequency that differed significantly from chance [χ^2(4, N = 1690) = 23.97, p < .001]. Post-hoc pairwise comparisons show that HC was selected more frequently than expected (p = .02), whereas HL was selected less frequently than expected (p < .001). MC, ML and NC were not selected at a frequency that differed significantly from chance. These findings survive Bonferroni correction for two comparisons.

The frequency with which different DRC type and point combinations were selected as preferred was calculated, pooling over the other independent variable (magnitude), producing seven settings: Compression Track (CT), Compression Subgroup (CS), Compression Full-sum (CF), Limiting Track (LT), Limiting Subgroup (LS), Limiting Full-sum (LF), and No Compression (NC). Some settings were found to be selected at a frequency that differed significantly from chance [χ^2(6, N = 1690) = 35.21, p < .001]. Post-hoc pairwise comparisons show that CT was selected more frequently than expected (p < .01), whereas LF was selected less frequently than expected (p < .001). CF, CS, LT, LS and NC were not selected at a frequency that differed significantly from chance. These findings survive Bonferroni correction for two comparisons.

The frequency with which different DRC magnitude and point combinations were selected as preferred was calculated, pooling over the other independent variable (type), producing seven settings: Moderate Track (MT), Moderate Subgroup (MS), Moderate Full-sum (MF), Heavy Track (HT), Heavy Subgroup (HS), Heavy Full-sum (HF), and No Compression (NC). Some settings were selected at a rate that differed significantly from chance [χ^2(6, N = 1690) = 38.46, p < .001]. Post-hoc pairwise comparisons show that HT was selected more frequently than expected (p < .01), whereas HF was selected less frequently than expected (p < .001). HS, MF, MT, MS and NC were not selected at a frequency that differed significantly from chance. The findings survive Bonferroni correction for two comparisons.

Results are summarized in Fig. 6, in which significant and marginally significant stimuli/stimulus groups that were/were not preferred are shown.

	Preferred	**Not Preferred**
significant at p ≤ .05	C, CT, HCT, T, HT, HC	L, HL, HLF, F, LF, HF
marginal	HCS	HLS

Fig. 6 Summarized listener preference results following post-hoc pairwise comparisons.

3 DISCUSSION

Our results indicate that the following stimulus groups were selected less often then expected by chance: L, HL, HLF, F, LF, and HF. A relationship between these stimuli is clearly apparent: many involve heavy magnitude limiting applied to the full-sum. This is as one might expect, given that it is this combination of settings that would be most likely to introduce distortion to the signal. One stimulus groups selected less often than chance that did not survive Bonferroni correction (HLS) had similar characteristics, entailing heavy magnitude limiting, but here applied to subgroups.

Our results also indicate that the following stimulus groups were selected more often than expected by chance: C, CT, HCT, T, HT and HC. One stimulus did not survive Bonferroni correction (HCS). Again, these preferences are closely related, many entailing DRC applied to fewer signals (track or subgroup), and the use of compression rather than limiting. These findings are in agreement with some earlier studies in which compression was found to be preferred over no compression [37], [48], although Ma et al. [37] do stipulate that listeners may prefer no compression, contingent upon genre. These findings are consistent the hypothesis that DRC applied to fewer signals simultaneously (i.e., to tracks, rather than subgroups or the full-sum) produces the most agreeable results for listeners. Listener preferences for compression over limiting may be linked to the perceptual impact of nonlinear distortion: the higher DRC ratios used in limiting may increase the degree to which nonlinear frequency components are generated in the original signal, potentially causing it to sound increasingly "harsh" [3], and reduce "pleasantness" [11]. A preference for light DRC over no DRC in the UNEQ condition of [11] was also reported. Considering that the function of DRC is to reduce variation in loudness, perhaps heavy DRC, in addition to introducing unpleasant distortion, also diminishes the expressiveness of music [50].

In Hjortkjær & Walther-Hansen [30], no evidence of listener preference for "less compressed music" was found, in contrast to our finding that compression was preferred over limiting (and no compression). It may be that moderate DRC is preferred over no DRC because the sonic characteristics imparted by DRC were pleasing to listeners for our musical stimuli. Potentially, this is a learned preference resulting from the widespread use of DRC in popular music production, making the sound of 'no DRC' high-fidelity audio sound less familiar and consequenly less agreeable.

Listener preferences for DRC applied to individual tracks rather than the full-sum may be because DRC (whether limiting or compression) reduces audio fidelity to a lesser degree when fewer signals interact simultaneously by restricting the introduction of the aforementioned sum and difference components [2]. The suggestion of Pestana & Reiss [18] that DRC is justified for the compensation of "erratic loudness ranges" is reasonable, although our results do appear to be at odds with their recommendation that DRC be used on the "overall mix" (referred to in this study as full-sum) as best practice.

In this study, DRC attack and release parameters were set to function in a generalized way, rather than being specifically optimized for each stimulus configuration, which may have influenced listener preferences. However, it would be difficult to obtain quantitative/comparable results had subjective adjustment of these parameters been undertaken, due to the variability of the individual signal envelopes. Furthermore, the organization of instruments into subgroup was based on instrument type (e.g., drums, vocals, and other instruments). There are numerous ways in which the subgroups could have been organized which may have influenced DRC behavior, and therefore listener preferences. However, our subgroups were broadly consistent with professional sound engineering preferences (for drum and vocal subgroups), reported in [32], [51].

This experiment only tested listeners' preferences for different DRC configurations using one popular music excerpt (unknown to participants beforehand), restricting the degree to which we can generalize our findings to other genres, or indeed to other stimuli within the popular music genre. However, using a single unknown musical excerpt (with different DRC configurations) as a source stimulus enabled participants to make a detailed audio quality comparison that would have been compromised had more audio stimuli (e.g., from multiple genres), or familiar stimuli, been used. Furthermore, the experiment required considerable concentration from participants, and the introduction of multiple musical clips would have rendered it too long to expect that concentration could be sustained (as would full counterbalancing, wherein every stimulus was compared to every other stimulus in both sequential orders by all participants). Informal feedback from participants suggested that the duration of the test protocol already may have been about as long as they were willing to tolerate. However, since the order that stimuli were used was partially counterbalanced (such that each participant was allocated one comparator stimulus), fatigue/boredom effects are not expected to have unduly influenced the pattern of results reported.

It is possible that the synchronicity of instruments may cause DRC to be invoked as a result of an ensemble of signals interacting constructively, rather than as a result of a single loud signal, which may lead to different pattern of results within genre, depending upon the timing and phase relationships between signals. Further work is required to verify whether our results generalize to different music genres (i.e., to a wider range and combinations of timbres), durations, instrumentations, and musical structures, and to more exhaustively examine the impact of DRC ratio, attack, and release parameters on listener preferences.

4 CONCLUSIONS

By manipulating the point in the mix chain at which DRC was applied, this study supports the hypothesis that listeners prefer music with DRC applied to fewer signals simultaneously (i.e., to tracks prior to grouping/summation), which is expected to have reduced distortions associated with the application of DRC to pre-mixed signals [2], [3], [11], [13], [16], [29], [30], [42], [52]. Our findings also suggest that listeners prefer compression over limiting, and the use of moderate DRC over none. Our results are compatible with those of Croghan, Arehart and Kates [11], who found that heavy DRC applied to maximize loudness reduced listener preference relative to where moderate DRC was used. In current industry practice, the application of compression to subgroups is commonplace; furthermore, limiting is often applied at the end of the signal chain, not just for overload protection, but also to increase loudness. Conversely, our findings suggest that listeners prefer music to which DRC is applied early in the signal chain, and where compression (rather than limiting) is used.

5 REFERENCES

[1] E. R. Toulson, J. L. Paterson, and W. Campbell, "Evaluating Harmonic and Intermodulation Distortion of Mixed Signals Processed with Dynamic Range Compression," in *KES Transactions on Innovation in Music*, 1st ed., Future Technology Press, Shoreham-by-Sea, UK., (2014).

[2] C. Roads, "A Tutorial on Non-Linear Distortion or Waveshaping Synthesis," *Comput. Music J.*, vol. 3, no. 2, pp. 29-34, (1979 Jun.). https://doi.org/10.2307/3680281

[3] B. C. Moore, C.-T. Tan, N. Zacharov, and V.-V. Mattila, "Measuring and Predicting the Perceived Quality of Music and Speech Subjected to Combined Linear and Nonlinear Distortion," *J. Audio Eng. Soc.*, vol. 52, no. 12, pp. 1228–1244, (2004 Dec.).

[4] J. M. Kates, "Principles of Digital Dynamic-Range Compression," *Trends Amplif.*, vol. 9, no. 2, pp. 45–76, (2005 Jun.). https://doi.org/10.1177/108471380500900202

[5] C.T. Tan and B. C. Moore, "Comparison of Two Forms of Fast-Acting Compression Using Physical and Subjective Measures," in *Proc 18th Int. Con. Acoust., Kyoto JP.*, (2004).

[6] R. A. van Buuren, J. M. Festen, and T. Houtgast, "Compression and Expansion of the Temporal Envelope: Evaluation of Speech Intelligibility and Sound Quality," *J. Acoust. Soc. Am.*, vol. 105, no. 5, pp. 2903–2913, (1999 May). https://doi.org/10.1121/1.426943

[7] M. A. Stone and B. C. J. Moore, "Effect of the Speed of a Single-Channel Dynamic Range Compressor on Intelligibility in a Competing Speech Task," *J. Acoust. Soc. Am.*, vol. 114, no. 2, pp. 1023-1034, (2003 Aug.). https://doi.org/10.1121/1.1592160

[8] A. C. Neuman, M. H. Bakke, C. Mackersie, S. Hellman, and H. Levitt, "The Effect of Compression Ratio and Release Time on the Categorical Rating of Sound Quality," *J. Acoust. Soc. Am.*, vol. 103, no. 5, pp. 2273–2281, (1998 May). https://doi.org/10.1121/1.422745

[9] D. Giannoulis, M. Massberg, and J. D. Reiss, "Digital Dynamic Range Compressor Design—A Tutorial and Analysis," *J. Audio Eng. Soc.*, vol. 60, no. 6, pp. 399–408, (2012 Jul.).

[10] B. Aarseth, "DAB Norway - Implementation of Loudness Normalization," *EBU Technical Review*, Q1, (2012 Mar.).

[11] N. B. H. Croghan, K. H. Arehart, and J. M. Kates, "Quality and Loudness Judgments for Music Subjected to Compression Limiting," *J. Acoust. Soc. Am.*, vol. 132, no. 2, pp. 1177–1188, (2012 Aug.). https://doi.org/10.1121/1.4730881

[12] C. Essling, J. Koenen, and C. Peukert, "Loudness War: How Record Labels Compete for Consumer Attention," *SSRN Electron. J.*, (2014). https://doi.org/10.2139/ssrn.2444708

[13] M. Wendl and H. Lee, "The Effect of Dynamic Range Compression on Loudness and Quality Perception in Relation to Crest Factor," in *Audio Engineering Society Convention 136*, *Berlin, GR.,* (2014 Apr.).

[14] E. Deruty and D. Tardieu, "About Dynamic Processing in Mainstream Music," *J. Audio Eng. Soc.*, vol. 62, no. 1/2, pp. 42–55, (2014 Feb.). https://doi.org/10.17743/jaes.2014.0001

[15] M. Kirchberger and F. A. Russo, "Dynamic Range Across Music Genres and the Perception of Dynamic Compression in Hearing-Impaired Listeners," *Trends Hear.*, vol. 20, no. 2, (2016 Feb.). https://doi.org/10.1177/2331216516630549

[16] E. Vickers, "The Loudness War: Background, Speculation, and Recommendations," in *Audio Engineering Society Convention 129*, *San Francisco, US.,* (2010 Nov.).

[17] D. Viney, "The Obsession With Compression," *Master of Music Technol. Thesis Lond. Coll. Music, Thames Val. Univ.*, (2008 Dec.).

[18] P. Pestana and J. Reiss, "Intelligent Audio Production Strategies Informed by Best Practices," in *Audio Engineering Society 53rd International Conference on Semantic Audio, London, UK.,* (2014 Jan.)

[19] S. Gorlow and J. D. Reiss, "Model-Based Inversion of Dynamic Range Compression," *IEEE Trans. Audio Speech Lang. Process.*, vol. 21, no. 7, pp. 1434–1444, (2013 Jul.). https://doi.org/10.1109/tasl.2013.2253099

[20] R. T. Solberg, "'Waiting for the Bass to Drop': Correlations Between Intense Emotional Experiences and Production Techniques in Build-Up and Drop Sections of Electronic Dance Music," *Dancecult*, vol. 6, no. 1, pp. 61–82, (2014 Jun.). https://doi.org/10.12801/1947-5403.2014.06.01.04

[21] M. Ronan, R. Sazdov, and N. Ward, "Factors Influencing Listener Preference for Dynamic Range Compression," in *Audio Engineering Society Convention 137, Los Angeles, CA., US.,* (2014 Oct.).

[22] E. Schubert, "Modeling Perceived Emotion With Continuous Musical Features," *Music Percept.*, vol. 21, no. 4, pp. 561–585, (2004 Jun.). https://doi.org/10.1525/mp.2004.21.4.561

[23] R. E. Kirk, "Learning, a Major Factor Influencing Preferences for High-Fidelity Reproducing Systems," *J. Acoust. Soc. Am.*, vol. 28, no. 6, pp. 1113–1116, (1956 Nov.). https://doi.org/10.1121/1.1908573

[24] H. F. Olson, "Frequency Range Preference for Speech and Music," *J. Acoust. Soc. Am.*, vol. 19, no. 4, pp. 549–555, (1947 Jul.). https://doi.org/10.1121/1.1916520

[25] S. Olive, "Some New Evidence That Teenagers May Prefer Accurate Sound Reproduction," in *Audio Engineering Society Convention 131*, *New York, U.S.,* (2011 Oct.).

[26] S. Olive, "Some New Evidence That Teenagers and College Students May Prefer Accurate Sound Reproduction," in *Audio Engineering Society Convention 132*, *Budapest, HU.,* (2012 Apr.).

[27] J. Reiss, "A Meta-Analysis of High Resolution Audio Perceptual Evaluation," *J. Audio Eng. Soc.*, vol. 64, no. 6, pp. 364–379, (2016 Jun.). https://doi.org/10.17743/jaes.2016.0015

[28] J. Kates and K. Arehart, "The Hearing-Aid Speech Quality Index (HASQI) Version 2," *J. Audio Eng. Soc.*, vol. 62, no. 3, pp. 99–117, (2014 Mar.). https://doi.org/10.17743/jaes.2014.0006

[29] K. H. Arehart, J. M. Kates, and M. C. Anderson, "Effects of Noise, Nonlinear Processing, and Linear Filtering on Perceived Music Quality," *Int. J. Audiol.*, vol. 50, no. 3, pp. 177–190, (2011 Mar.). https://doi.org/10.3109/14992027.2010.539273

[30] J. Hjortkjær and M. Walther-Hansen, "Perceptual Effects of Dynamic Range Compression in Popular Music Recordings," *J. Audio Eng. Soc.*, vol. 62, no. 1/2, pp. 37–41, (2014 Feb.). https://doi.org/10.17743/jaes.2014.0003

[31] W. J. Dixon and A. M. Mood, "The Statistical Sign Test," *J. Am. Stat. Assoc.*, vol. 41, no. 236, pp. 557-566, (1946 Dec.). https://doi.org/10.2307/2280577

[32] D. Ronan, B. De Man, H. Gunes, and J. D. Reiss, "The Impact of Subgrouping Practices on the Perception of Multitrack Music Mixes," in *Audio Engineering Society Convention 139, New York, US.,* (2015 Oct.).

[33] A. Case, *Mix Smart*. 1st ed., Focal Press, Oxford, UK., (2011).

[34] B. Owsinski, *The Mixing Engineer's Handbook*. 2nd ed., Course Technology Inc., Vallejo, CA., US., (2013).

[35] R. Izhaki, *Mixing Audio: Concepts, Practices and Tools*. 2nd ed., Focal Press, Oxford, UK., (2013).
[36] V. Braun and V. Clarke, "Using Thematic Analysis in Psychology," *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, (2006 Jan.). https://doi.org/10.1191/1478088706qp063oa
[37] Z. Ma, B. De Man, P. Pestana, D. Black, and J. Reiss, "Intelligent Multitrack Dynamic Range Compression," *J. Audio Eng. Soc.*, pp. 412–426, (2015 Jun.). https://doi.org/10.17743/jaes.2015.0053
[38] D. H. Brainard, "The Psychophysics Toolbox," *Spat. Vis.*, vol. 10, no. 4, pp. 433–436, (1997 Jan.). https://doi.org/10.1163/156856897x00357
[39] D. G. Pelli, "The VideoToolbox Software for Visual Psychophysics: Transforming Numbers Into Movies," *Spat. Vis.*, vol. 10, no. 4, pp. 437–442, (1997 Jan.). https://doi.org/10.1163/156856897x00366
[40] M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, and C. Broussard, "What's New in Psychtoolbox-3," *Perception*, vol. 36, no. 14, p. 1, (2007 Aug.).
[41] A. L. Nichols and J. K. Maner, "The Good-Subject Effect: Investigating Participant Demand Characteristics," *J. Gen. Psychol.*, vol. 135, no. 2, pp. 151–166, (2008 Apr.). https://doi.org/10.3200/genp.135.2.151-166
[42] M. A. Stone, B. C. J. Moore, C. Füllgrabe, and A. C. Hinton, "Multichannel Fast-Acting Dynamic Range Compression Hinders Performance By Young Normal-Hearing Listeners in a Two Two-Talker Separation Task," *J. Audio Eng. Soc.*, vol. 57, no. 7/8, pp. 532–546, (2009 Aug.).
[43] ITU, "ITU Recomendation BS.1770, Algorithms to Measure Audio Programme Loudness and True-Peak Audio Level," *International Telecommunications Union*, Geneva, (2006).
[44] P. D. Pestana, J. D. Reiss, and A. Barbosa, "Loudness Measurement of Multitrack Audio Content Using Modifications of ITU-R BS. 1770," in *Audio Engineering Society Convention 134, Rome, IT.,* (2013 May).
[45] N. A. Macmillan and D. C. Creelman, *Detection Theory: A User's Guide*. Lawrence Erlbaum, Mahwah, NJ., US., (2004).
[46] ITU, "ITU Recomendation BS. 562-3, Subjective Assessment of Sound Quality." *International Telecommunication Union*, Geneva, (1990).
[47] K. Pearson, "On The Criterion That a Given System of Deviations From the Probable in the Case of a Correlated System of Variables is Such That it Can Be Reasonably Supposed to Have Arisen From Random Sampling," *Philos. Mag. Ser. 5*, vol. 50, no. 302, pp. 157–175, (1900 Jul.). https://doi.org/10.1080/14786440009463897
[48] J. A. Maddams, S. Finn, and J. D. Reiss, "An Autonomous Method for Multi-Track Dynamic Range Compression," in *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12), York, UK.,* (2012 Sep.).
[49] Open Science Framework Data: https://osf.io/e8m5p/?view_only=1b4405f58d4f4934b941cc901ea65000.
[50] P. N. Juslin and P. Laukka, "Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?," *Psychol. Bull.*, vol. 129, no. 5, pp. 770–814, (2003). https://doi.org/10.1037/0033-2909.129.5.770
[51] D. Ronan, H. Gunes, and J. D. Reiss, "Analysis of the Subgrouping Practices of Professional Mix Engineers," *Audio Engineering Society Convention 142, (2017, May)*.
[52] U. Zölzer, "DAFX: Digital Audio Effects," 2nd ed., John Wiley & Sons, Hoboken, NJ., US., (2011).

THE AUTHORS


William Campbell


Justin Paterson


Ian van der Linde

William Campbell is Lecturer on the JAMES accredited Audio Music Technology degree in the Department of Computing & Technology, Anglia Ruskin University, Cambridge, UK, a Sound Engineer and Producer for Film, Music and Games. He holds a BSc (Hons) in Audio Music Technology, and is presently studying for a PhD. Before entering academia, he was Aircraft Structural Maintenance Engineer at Lockheed Martin and Raytheon Aerospace, and Production Supervisor for the US Air Force. He is a Fellow of the UK Higher Education Academy, Member of the Music Producers Guild, and Member of the Audio Engineering Society.

●

Justin Paterson studied Electronic Engineering at the University of Dundee, UK before being awarded an MSc in Microelectronic System Design from Brunel University, UK in 1989. He also graduated from the Musician's Institute London in 1991, and in 2015 was conferred as Doctor of Music at the University of West London, UK. Throughout the 1990s, Justin was a professional drummer based in London, but latterly moved towards music production, in which he has specialized ever since. Still an active professional practitioner, he joined London College of Music | University of West London in 2004, where he leads the MA Advanced Music Technology. He is currently Associate Professor of Music Technology. Justin is the author of *The Drum Programming Handbook*. His research has an international profile ranging from journal articles to book chapters. In 2015, he developed a novel app-based format of music release featuring interactive playback, together with Professor Rob Toulson of the University of Westminster – a project that at present has a patent pending based on intelligent audio crossfades. Recent research has been around interactive 3-D audio for virtual reality (VR), collaborating with record label *Ninja Tune* and VR company *MelodyVR*. Justin is also co-chair of the Innovation in Music conference series, and a co-editor of its associated journal.

●

Ian van der Linde is Reader in the Department of Computing & Technology at Anglia Ruskin University, Cambridge, UK and a research scientist at the Vision & Eye Research Institute (VERU) in the Postgraduate Medical Institute, Cambridge. He holds a BSc (Hons) in Software Engineering, an MSc (Lond) in Cognitive & Clinical Neuroscience, and a PhD in Human/Computer Vision. He is a Chartered Professional Fellow of the British Computer Society (BCS), Fellow of the UK Higher Education Academy, Chartered Engineer of the UK Engineering Council, and Chartered Scientist of the UK Science Council. Formerly visiting sholar at the Laboratory for Image & Video Engineering (LIVE), University of Texas at Austin, he has authored over 65 journal and conference outputs, specialising in experimental psychology relating to visual perception, memory and attention.