

Valves versus transistors

The results of a comparison among three different amplifiers

by James Moir, F.I.E.E. James Moir and Associates

What follows is the result of a series of listening tests commissioned by the Acoustical Manufacturing Company ("Quad"). The intention of the tests was to investigate claims that valve amplifiers sound better than transistor amplifiers.

IN recent years a cult has arisen in which the members worship valve amplifiers, claiming that amplifiers employing transistors are incapable of achieving the same high standard of sound quality that is obtainable from amplifiers using valves. The reasons for the claimed superiority of valved designs are never set out in detail, nor is there any attempt to prove the claims, but instead they appear to be based on the bottomless argument that absolute contradiction of their claims is not possible, so by a process of accelerated inference a remote possibility becomes converted to an absolute certainty. It is to be expected that a valve amplifier costing, say £1000, will have a better performance than a transistorised model costing £100, but it appears to the writer that the claims go beyond this and that it is being suggested that all present designs of transistorised amplifiers include some ingredient 'X' that, being beyond any possible measurement, automatically ensures that it is impossible to duplicate the performance of a valve amplifier with any unit employing transistors.

The discussion that follows describes one attempt at the professional level to discover whether there is in fact any basic difference in the performance of a group of valve and transistor amplifiers, all of them recognised as being at the top of their class at the time they were in production. Ingredient 'X' being, by definition, impossible to measure, any attempt at assessing the performance of the amplifiers by objective techniques would have been unconvincing and was therefore discarded, leaving listening tests as the only alternative likely to be acceptable to members of the cult. However it was decided that if listening tests did reveal any significant difference in the sound quality, then the subjective judgement would be followed by a determined attempt to segregate the cause of the observed quality differences by objective means. Listening test techniques that are completely free from any criticism are not

easy to arrange. There is a IEC Publication No. 543 covering the ground in a rather vague manner and a more recent IEC document 29B/WG5 providing a more detailed discussion of the subject. This is more specific in its suggestions and will presumably appear as a British Standard in due course. Where this document was applicable to the tests described, it was followed as far as possible.

In designing a listening test it appears reasonable to try to ensure that all the other elements in the reproducing system are at least an order better in performance than the element being submitted to a subjective judgement, although this is manifestly difficult to ensure when the components being judged are amplifiers of the highest class. If it is assumed that the non-linear distortions are at least a rough guide to the sound quality that can be obtained, then it is impossible to ensure that the recordings and loudspeakers that must be employed in any subjective assessment have a performance that is at all comparable to that of the best current amplifiers. Broadly speaking the situation is as follows.

There are power amplifiers on the market having distortions that are at least 90dB below their rated power output. The best current loudspeaker designs have distortions around 40dB down, professional tape recordings are perhaps 30 to 40 dB down, while the best current disc recordings have distortions that are no better than 25 dB below maximum output. Amplifiers have the additional advantage of having a large amount of headroom allowing the amplifier to be worked well below its overload point without compromising the system signal-to-noise ratio. In consequence the working distortions are much lower than are indicated by a quotation of the distortion content at full power output.

It is not the purpose of the present contribution to discuss all the distortions that exist in a reproducer system, but, assessed on the basis of the amplitude dependent components, it is reasonable to suggest that the best amplifiers have distortion products that are at least 40 to 50dB lower than the other elements in a hi-fi sound system. Thus it is a major weakness of a subjective assessment that the programme material and the equipment that must

be used for the evaluation has a performance that is far far worse than the amplifiers being evaluated.

The amplifiers employed for the listening tests were all the products of one manufacturer, Acoustical Manufacturing Co. Ltd, (Quad in other words), who commissioned the tests. It seems likely in any case that most people would agree that their amplifiers have been in the top class for very many years, right back to the time when valved types were the only models available. In addition, using the products of one manufacturer seems essential if comparison with other manufacturers' products and design skills is to be avoided. Quad II amplifiers were the valve model used and the performance was compared with that of the type 303, their first transistor design, and with their model 405, the present current dumping transistor design.

Choosing the programme material for a listening comparison is a very difficult problem when the products being judged are 'state-of-the-art' amplifiers. About forty programme samples on 15 i.p.s. tape were available from four of the best-known studios in the country, with some additional material from several other sources. All were original recordings or first generation copies of original recordings made on machines of the highest professional standard. These samples had been provided as the best examples of current recording practice in the particular studio, but these samples were further distilled by careful listening comparisons until we were left with four selections that were considered to be outstanding in respect of frequency response, low distortion and acoustic clarity. The examples of programme finally used consisted of a concert orchestra, a light orchestral section, a group of male singers and finally a 'pop' group, all thought to be broadly representative of the type of music played at home by the average enthusiast.

The tapes were replayed on a Studer A80 recorder, the signal output being applied directly to the three amplifiers through resistive potentiometers to achieve the same output voltage from each of the power amplifiers. Pre-amplifiers were not necessary and were not used. A double-beam 'scope was installed to monitor the output signal from the amplifier to ensure that over-

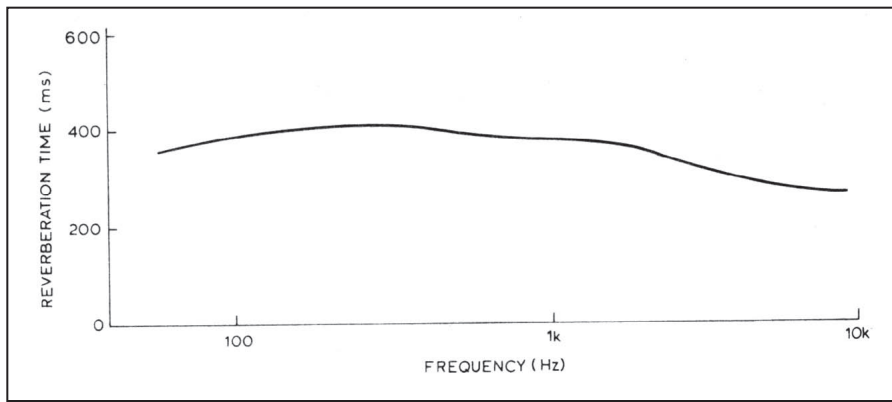


Fig. 1. Listening room reverberation time response.

loading did not appear even on instantaneous peaks of very short duration.

Yamaha Type NS 1000 loudspeakers were employed, the choice being that of one of the cult members as a condition of his participation in the tests,

The cult members that were invited to take part in the tests accepted but subsequently withdrew from the listening group, but by that time considerable effort had been devoted to determining the effect of the speaker impedance on the frequency response of each of the amplifiers and by then there was inadequate time available to investigate the performance of any substitute speaker system.

The output from the amplifiers being compared was switched to the loudspeakers by relays with gold-plated contacts to avoid any suggestion that contact-resistance or rectifying-action at the contacts was in any way responsible for the findings. These relays were operated through a switching system that allowed a randomised selection of any pair of amplifiers to be connected to the loudspeakers. At the same time the switching system operated a series of lamps that indicated the number of the particular test to the listening panel. Separate A and B lamps were employed to indicate which of the two amplifiers being compared was connected to the loudspeaker, although the panel had no means of knowing the types of amplifier in use in any particular comparison; all the technical equipment was operated in an adjacent room. In a large number of the comparisons the same amplifier was used in both the

'A' and 'B' positions.

The listening panel were all well known and experienced listeners. They were seated in two rows at a distance of approximately 4.0 metres from the two loudspeakers, but they were free to interchange seating positions as often as they wished. The test was conducted in a typically-furnished lounge having the measured reverberation time/frequency relation shown in Fig. 1 and an ambient noise level around 22dBA in the absence of the panel, rising to 28 dBA at the quietest moments when the panel in a form suitable for statistical training run before judging commenced, the loudness level was adjusted to that thought reasonable by the panel, the level being continuously monitored by the double-beam crt across the speaker line to ensure that this level was maintained through the series of tests,

It is probably impossible to assemble a reproducing system that is absolutely beyond all criticism but the system used had a 'state-of-the-art' performance that was far beyond the facilities of any ordinary enthusiast.

Each item in the musical programme was presented to the panel as two 30s repeats, separated by an interval of one or two seconds during which the amplifiers were switched. This was followed by an interval of about 15-20s before the second piece of music was presented in the same general format.

Obtaining the opinion of a listening

panel in a form suitable for statistical analysis requires some careful consideration, for it is not as simple as might appear at first thought. When three identical amplifiers are compared, then if a sufficiently large number of opinions are taken, each amplifier will get one-third of the votes in much the same way as an unbiased penny will come up heads on 50% of the throws, but only if there are a large number of attempts. Thus a large number of independent quality judgements are required if the result is to be even moderately conclusive. If a small number of judgements are made, any one of the three amplifiers is likely to find favour by sheer chance, in much the same way as the neutral penny tossed three times will confirm that it is weight-biased because heads will come up twice as often as tails.

To judge the amplifier performance, each of the four pieces of music was played twenty-four times to a panel of six judges, their opinion on the performance of each pair of amplifiers being given after hearing each of the four pieces on music on each of two amplifiers. After each of the four pieces of music, the panel members were asked to record their opinion on that particular pair of amplifiers in the form:-

1. I prefer A.
2. I prefer B.
3. I have no preference.

If a preference was expressed the panel members were asked to indicate their reasons for that preference. It was thought just possible that an expressed preference might be connected in some way with the particular seating position, so each panel member was also asked to mark his position on a small seating plan on the score sheet.

To avoid listening fatigue there were gaps in the comparison process after twelve judgements had been made, with longer gaps after twenty-four judgements. Lunch was taken in the interval between the first and second groups of twelve judgements, a whole day being devoted to the comparisons. Every possible effort was made to ensure that the test conditions were as free from criticism as could be achieved, but it would be too optimistic to believe that the arrangements were beyond all criticism.

Table 1.

PAIRED COMPARISON TEST RESULTS

| Comparison | Quad 11/405 | | | Quad 11/303 | | | Quad 303/405 | | | Same Amplifier | |
|---------------|-------------|--------|-------|-------------|--------|-------|--------------|--------|-------|----------------|------------|
| | Prefer | Prefer | No. | Prefer | Prefer | No. | Prefer | Prefer | No. | Preference | No |
| | 11 | 405 | Pref. | 11 | 303 | Pref. | 303 | 405 | Pref. | Preference | Preference |
| Listener a | 5 | 4 | 15 | 7 | 6 | 11 | 5 | 6 | 13 | 11 | 13 |
| Listener b | 2 | 2 | 20 | 1 | 3 | 20 | 4 | 3 | 17 | 3 | 21 |
| Listener c | 3 | 6 | 15 | 5 | 3 | 16 | 4 | 1 | 19 | 7 | 17 |
| Listener d | 4 | 9 | 11 | 2 | 4 | 18 | 7 | 4 | 13 | 8 | 16 |
| Listener e | 2 | 3 | 19 | 2 | 2 | 20 | 3 | 1 | 20 | 3 | 21 |
| Listener f | 8 | 7 | 9 | 8 | 10 | 6 | 7 | 4 | 13 | 14 | 10 |
| Group results | 24 | 31 | 89 | 25 | 28 | 91 | 30 | 19 | 95 | 46 | 98 |

When statistically analysed using the 50% Probability Test none of these results indicates either on a group basis, or an individual basis, that there are any audible differences among the performance of the three amplifier.

A summary of the listening panel's scoring is given in Table 1. Each panel member had 24 opinions to record on each paired comparison and the Table indicates his views,

The data collected is sufficient to allow two of the many standard statistical tests to be applied to determine how far the result obtained is likely to be due to sheer chance (luck) rather than to any real difference in the performance of the three amplifiers. There are several statistical tests that can be used for this purpose, but two that appeared particularly applicable have been applied. The 50% probability test applied to a paired comparison of samples thought to be identical reveals how far the consensus opinion is due to sheer chance and how far it is due to a real difference between the amplifier being compared.

As a second test of the validity of the listening panel's opinion the Chi-square test was applied to their scoring. Both tests confirm that the residual preferences expressed by the panel were no more than would be achieved by sheer chance (guesswork is the crude term). The analysis is not reproduced in detail but one simple and easy-to-understand result is worth quoting.

There were a large number of 'no preference' votes, sufficient to allow them to be separated into a 'no preference' group when a single amplifier was used in both the A and B positions and a second group of 'no preference' votes when two different amplifiers were being compared. The percentage of 'no preference' votes when one amplifier was being compared with itself was 68%, while the number of 'no preference' votes when two different amplifiers were being compared was 64%. The panel judgements amplifiers indicated that they did not prefer anyone amplifier to either of the other types and that there are no consistent audible differences between any of the amplifiers being compared.

It is worth commenting that during a trial run some days before the test described, a different expert panel and different programme material were used, but the result did not differ in any significant respect from those obtained in the 'official' test. There was no indication of a consistent preference for any one type of amplifier either by any individual member of the panel, or by the panel as a group, or by the combined result of two separate group tests.

The test was primarily aimed at discovering whether there were any real differences in the sound quality that could be achieved from valve and transistor amplifiers, but there were other incidental differences between the designs that reflect the developments in technology that have taken place since the appearance of the valve design around 1960. The valve design necessarily employed an iron-cored output transformer, whereas the 303 includes a series capacitor of 2000 μ F, while the 405

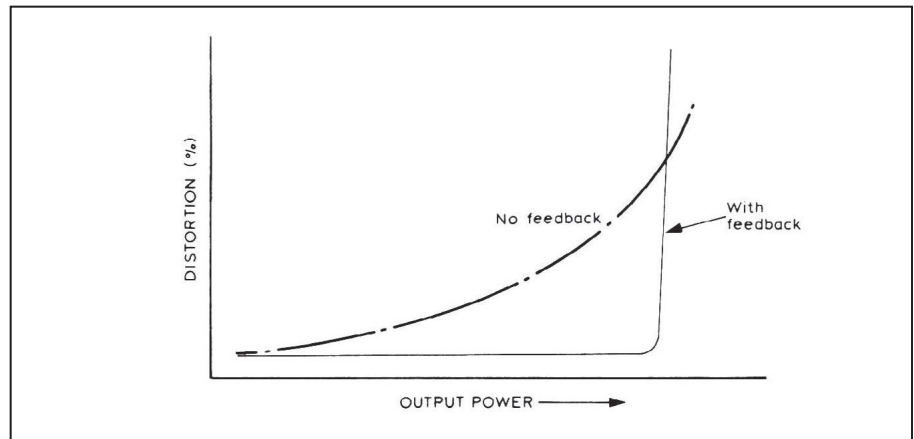


Fig. 2. Output power against percentage distortion characteristics for amplifiers with and without feedback.

amplifier has the loudspeaker directly-coupled to the output transistor. Separate power supplies are used in the Quad II but a common power supply is used for both channels in the 303. Protection circuits of different designs are used in both the 303 and 405. The 405 includes a circuit that provides a sharp cut-off below a frequency of 20Hz whereas the Quad II falls away more slowly below 20Hz.

The comparison also includes amplifier designs completed before t.i.d. (transient intermodulation distortion) became a misunderstood explanation for every subjectively-assessed difference in amplifier performance. All of these differences have at some time been claimed to be responsible for large difference in sound quality, but it will be appreciated that the comparison tests quoted show that in the hands of a skilled designer none of these factors appears to be of real importance.

It seems reasonable to conclude that if a dozen expert listeners working for a day with the best available equipment and the best obtainable programme material cannot find any significant difference between the amplifier types, then it is highly unlikely that such differences as are inherent in designs that span more than twenty years will be of any significance to any user.

Finally it is worth considering how far the findings are reasonable from an engineering point of view. Insofar as the understood distortions in any of the three amplifiers are far lower than in any part of the record-replay system, the findings are logical and are what might be expected. The residual amplifier distortions are likely to have been masked by the much greater distortions in the recording/replay elements, even though these were representative of the best current professional practice.

The absence of preference for anyone type of amplifier is exactly what an engineer might expect, so it is reasonable to ask why other published listening tests and reviews appear to show differences so large that one expert was able to speak of the type 405

amplifier as producing '100 watts of squawking sound'. This aspect is worthy of some comment.

It has been stressed that the amplifiers were all operated within their power ratings and this may be one possible key to the difference between findings of this listening panel and of others that have been published. An amplifier which employs a relatively small amount of feedback overloads in an entirely different way to an amplifier having a large amount of negative feedback, a result indicated by Fig. 2. Negative feedback can reduce amplifier distortions by a large factor, but only below the point at which the distortion without feedback is less than a few per cent. Above this power the overall distortion is greatly increased by the application of negative feedback.

In consequence an amplifier design employing a small amount of feedback will approach its rated distortion limit rather slowly, the overall distortion increasing gradually as the input signal is increased. In contrast an amplifier employing a large amount of feedback will exhibit much lower amounts of distortion at output powers below the rated value, but the distortion will increase very rapidly above this 'overload' value.

The difference is well illustrated by the distortion/power output curves of Fig. 2. In practice it is almost impossible to specify with any real accuracy the distortion content of an amplifier employing large amounts of negative feedback at output power levels near, or above, the overload point. Beyond this point the distortion increases so rapidly that small changes ($\pm 1\%$ etc.) in mains supply voltage or signal input voltage may increase or decrease the measured distortion by four or five times without there being any significant increase in the power output.

Above its overload point, every amplifier type exhibits its own particular overload characteristics depending on the cause of the distortion. Slew-rate limiting, transient intermodulation distortion, dynamic intermodulation distortions, amplitude compression, mains frequency modulation and the well-understood distortions due to curvature in the overall transfer characteristics will all introduce their own

characteristic acoustic effects.

However it is unreasonable to operate any amplifier above its claimed power output. If the amplifier has to be overloaded to achieve an adequately loud signal then a more efficient loudspeaker or a more powerful amplifier should be substituted.

Differences in the acoustic performance of an amplifier can also be induced by applying test signals that are outside the designed frequency band of the amplifier. In a misguided attempt to assess the transient performance of an amplifier short square wave dc pulses or short pulses of sinusoidal tone are often applied to an amplifier, but it is easy to show that such short pulses contain components up to a frequency in the region of 80 to 100kHz. For example a pulse 10µs long will have a first zero in its amplitude response at 100kHz and the amplitude of components at 80kHz will be only a few dB lower than the maximum. In most amplifiers this is well outside the designed frequency range and overloading is easily produced by signals that would be far from overloading the amplifier if they were of the same amplitude but inside the designed frequency pass band.

Finally some comment about the pitfalls that are possible when connecting up the components in any hi-fi system. Fig. 3 shows the same units having their earthing connections set out in different ways that are superficially identical. Yet in practice there may be large differences in the performance of the system in respect of sound quality depending on just where the earthing wire is connected to the circuit earth bus or the chassis. An assembly in which two or more amplifiers are connected to the same earthing system and the same loudspeaker will almost certainly not achieve their catalogue performance unless some considerable expertise is employed in assembling and testing the completed system.

The writer took no part in reaching the judgements discussed, but it is his firm personal opinion that no 'X' ingredient, nor any 'black magic' is necessary to explain the results. As Mr. Peter Walker has commented, "If an engineer finds that a couple of measurements appear to contradict Ohm's Law he does not immediately rush into print with his findings, he looks again at the measuring technique employed."

This comment should be taken to heart by all those reviewers who are so eager to print comment that flatly contradicts all reason. They might, with advantage, look at the experience of a Canadian journal.¹ Their reviewing panel listened to half a dozen amplifiers, all in the top class and found large differences in sound quality. Doubting the findings, they set out to investigate the reasons for this. After having eliminated all the little problems that they were able to unearth, a repeat of the listening test revealed that all the quality differences had vanished. It is a

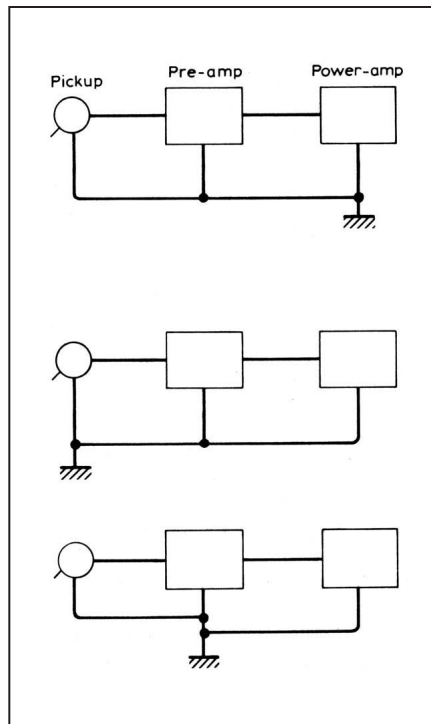


Fig. 3. Nominally-equivalent earthing arrangements.

mundane explanation perhaps, and one that is without journalistic appeal, but then the truth is often less sensational than pseudo-science or witchcraft.

Reference and further reading

1. Six Amplifiers, How Did They Compare? Ian G. Masters, Audio Scene, May 1977.
2. Audible Amplifier Distortion is not a Mystery, P. Baxandall, *Wireless World*, Nov. 1977.
3. Rational Amplifier Testing, P. Walker, Hi-Fi News, July 1977.
4. Can You Hear Any Difference? Adrian Hope, Hi-Fi News, June 1978.
5. www.keith-snook.info keith@snookeu

How accurate are audio reviews? From p. 67

new range of Wharfedale speakers that British hi fi customers had bought more of their XP speakers than any other brand in the past ten years, in spite of no press coverage or adverse reviews. The four large comparative reviews of bookshelf speakers that had appeared in the previous four months had grouped up to 16 loudspeakers, not one of them a Wharfedale, yet 20% of readers, he asserted, would be buying a Wharfedale product. At least five competitors had appeared in three out of four reviews.

"When the Wharfedale Denton or Linton products have been reviewed there has been virtually no consideration of the price the consumer will pay. A recent letter in the hi fi magazine had a customer asking advice on the purchase of a pair of Dentons, to which the magazine said there were a lot better products on the market and proceeded to list some which would cost the consumer between 30% and 100% more than the Denton, hardly fair advice."

"We have survived this approach for some years, but we are naturally concerned about the eventual impact of what we consider to be slightly unfair journalism. Particularly with the increasing competition from Japanese loudspeakers, who have the weight of advertising expenditure to overcome any type of press coverage."

Reviewers should take price into account, he said. Later he told *Wireless World* that he also questioned the usual method of reviewing, by comparing the tested equipment with a reference monitor. "It's a bit like comparing a Cortina to a Jensen: you are always going to find areas of performance that are not quite up to the standard of the Jensen, whereas a Cortina would normally be measured against standards expected of that category of car."

Electronics well-represented in Queen's awards

THE FOLLOWING firms are among those listed in the latest Queen's awards to industry: Brookdeal Electronics, for technical achievement in signal recovery; General Instrument Microelectronics, for exports; International Aeradio, for exports; Marconi Avionics, for exports; Racal Dana, for technological achievement in programmable synthesised signal generators.

Three audio companies were represented. Electrosonic won an export award for audio visual and lighting equipment. The Acoustical Manufacturing Company won a technology award for their QUAD current dumping amplifier, and Bowers & Wilkins won an award for exporting almost 90% of production.

Disaster avoided

Over 300 mobile radio users attended Pye's two-day symposium, (see other news items, this issue) including some from the Ministry of Defence, the Home Office, police forces, area health authorities, fire brigades and water authorities. The symposium was a triumph over adversity. Pye employees at first said they would picket the week's events as a protest against their £1 million cost. No sooner had they been persuaded not to do so than the giant marquees in which the symposium was to be held, set up on fields next to the Cam river, were flooded by three feet of water, this only days before the opening. The same employees won the eternal gratitude of Pye executives by working day and night to clean up the mess. They did, so effectively that delegates had to be shown photos of the inundated marquees to believe what had happened.