Critical Examination of the Article *Impulse Noise Injury Prediction Based on the Cochlear Energy* by Zagadou, Chan, Ho, and Shelly

by G Richard Price, Joel T Kalb, and Charles R Jokel
NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer’s or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.
Critical Examination of the Article Impulse Noise Injury Prediction based on the Cochlear Energy by Zagadou, Chan, Ho, and Shelly

by G Richard Price and Joel T Kalb
Human Research and Engineering Directorate, ARL

Charles R Jokel
US Army Public Health Center, APG (Provisional)

Approved for public release; distribution is unlimited.
## Critical Examination of the Article Impulse Noise Injury Prediction Based on the Cochlear Energy by Zagadou, Chan, Ho, and Shelly

In an article claiming to produce an improved damage risk criterion for human exposure to intense impulsive sounds, Zagadou, Chan, Ho, and Shelly critiqued the Auditory Hazard Assessment Algorithm for Humans (AHAAH) model used by the US Department of Defense (DOD) in assessing noise hazard and also proposed an alternative. Their approach required that they 1) improve the fit of AHAAH to the transfer functions, 2) reproduce the stimulating conditions in the Albuquerque Studies (AS), record pressure histories at the eardrum position of a manikin, and 3) produce a better predictive algorithm. Unfortunately, they succeeded in none of their aims. They did not in fact use the DOD program but rather some unknown variant. The model they did create had a demonstrably worse fit to the transfer functions. For a variety of reasons, they did not reproduce the stimulating conditions in the AS. The acoustic data they published are rife with serious artifacts and cannot be considered a replication of the AS data. Finally, we have shown that the interpretation of the data and the analytical procedures simply ignored relevant data on the ear’s response to intense sounds. They consistently made assumptions, creating “data” that contained biases in the direction of energy explaining everything—a problem of logical circularity.
Contents

List of Figures

1. Introduction 1

2. The Non-Problem of the “Non-Monotonic” Response 2

3. Problems with the Model’s Variables? 3
   3.1 Basic Understanding 4
   3.2 Middle Ear Muscle Latency Inaccuracy 4
   3.3 Overlooking Insights from Animal Tests 5

4. The Non-Improvement Improvement 6

5. The Stapes Displacement Problem 6

6. Nonuse of the AHAH Model 6

7. Non-Improvement of AHAH-PK 8

8. Non-Replication of the AS Pressure Histories 10
   8.1 Measurement Issues 10
      8.1.1 Microphone Use 10
      8.1.2 Apparatus 11
      8.1.3 ATF Issues 11
      8.1.4 Distorted Pressure Histories 12

9. Philosophical Issues 15
   9.1 Experimental Technique 15
   9.2 Assumed and Extrapolated Data 16
   9.3 The Physiology of Loss 17
   9.4 Middle Ear Muscle Activity 18

10. Conclusions 19
List of Figures

Fig. 1  Stapes velocity transfer functions for what Zagadou et al. claimed to be the original AHAAH model (AHAAH ZCHS) and the same calculation with the actual AHAAH model (AHAAH PK) ........................................ 7

Fig. 2  Comparison of stapes velocity transfer functions for Zagadou et al.’s AHAAH-New, the original AHAAH (AHAAH PK), the human data of Aibara et al. and Kringlebotn and Gundersen ........................................ 9

Fig. 3  Eardrum pressure for 1 m, Level 1, from the ATF (blue curve) and the eardrum pressure calculated using the original AHAAH model (red curve) and the AS tragus data ........................................ 14

Fig. 4  Eardrum pressure for 1 m, Level 6, from the manikin (blue curve) and the eardrum pressure calculated using the original AHAAH model (red curve) and the AS tragus data ........................................ 15
1. Introduction

The ability to predict the hazard from very intense sounds typical of gunfire remains an active interest for many, including the armed forces, civilian sport shooters, law enforcement, industry, and society as a whole, who may be exposed to the intense noise of an automotive airbag deployment. The Auditory Hazard Assessment Algorithm for Humans (AHAAH) model is serving as the basis for such evaluations (SAE 2013; MIL-STD-1474E 2015) and is being considered as a basis for an American National Standards Institute (ANSI) standard for impulse noise exposure (ANSI S3.62 2017), the first of its kind. Consequently, when Zagadou, Chan, Ho, and Shelly (2016) proposed both major revisions to the current AHAAH model and immediate use of the revised model as a basis for hazard analysis, it was clearly appropriate to evaluate their proposals.

In reviewing their effort, we found that we had serious reservations at many places both with regard to the experimental work they did with their understanding and use of the AHAAH model, their interpretation of the work they report and the alternate model they propose. Our remarks here focus on the most critical issues, things that are essentially objective and affect the ultimate acceptance or rejection of their positions.

Zagadou et al. (2016) introduced their work by expressing concern for what they contended was the AHAAH model’s irrational performance in scaling the hazard within the Army’s Albuquerque Studies (AS). Namely, they contended that it behaved “non-monotonically” (i.e., higher peak pressures under some conditions involved lower predicted hazards). In their article, they proposed first to begin with an update of the variables in the model representing the ear’s physiology with the promise of improving the model’s fit to the appropriate transfer functions. They expected that these changes would correct what they argued was the irrational scaling for the AS stimuli. Second, they proposed to examine the improved model’s fit to the human data on threshold shift (TS) from intense impulses. The major portion of this part of their program was to reproduce exposure stimuli from the AS and expose an acoustic test fixture (ATF) at the subject locations. The ATF would be fitted with the various hearing protection devices (HPDs) used in the AS and pressure histories collected at the eardrum location of the ATF. Given their expressed intent to reproduce the AS stimuli, along with use of the HPDs from the AS, they felt they would be justified in interpreting the TS data from the subjects (Ss) in the AS studies using the acoustic data from the eardrum position of their ATF. No additional human exposures were conducted, so it was critical that the stimulating conditions of the AS be reproduced as closely as possible. Third, our comments address the proposals made by Zagadou et al. to adopt a newly created
energy-based model of hearing loss. Such proposal has the effect of losing most of
the validating data that support the existing AHAAH model. The approach they do
recommend is barely supported by very few pieces of data selected with a highly
questionable rationale and other premises that are clearly counter to the hearing loss
data appropriate to this high-intensity regime.

We address, in turn, the problem of the “non-monotonic response”, the proposed
improvements to the model, the replication of the acoustic data from the AS, and
the revised model and its relationship to assessment of hearing loss.

2. The Non-problem of the “Non-monotonic” Response

Zagadou et al.’s contention that the AHAAH model behaved irrationally was
baseless and, in fact, the model’s predictions have been noted, found to be internally
consistent, and, importantly, in line with hearing loss data (Price 2007, 2010).
Specifically, for the 1- and 3-m exposures in the AS, the model did make the
unusual prediction that the most-hazardous impulses would be in the middle of the
intensity range (Level 4) and that the impulses with the higher peak pressures and
energies would in fact be progressively less hazardous. It was perhaps a surprising
prediction, but we believe that one of the functions of a model is to challenge
preconceptions and produce insight into processes. In any event, the ultimate arbiter
in these cases is data. If there were a problem with the model, a bad prediction
would be evidence that something needed to be reformulated and tested, and a
successful outcome would produce an improved model. However, Zagadou et al.
skipped the evidence phase of this process (the model’s predictions in these cases
had not been shown to be faulty), and they simply presumed the predictions of
lower hazard for those particular impulses as evidence of obvious errors within the
model.

The heart of the problem with Zagadou et al.’s analysis, we believe, is essentially
its internal logic. They began with the premise that there could be no such thing as
what they referred to as a “non-monotonic” prediction (i.e., a higher peak pressure
with essentially the same A-duration must always produce an increase in hazard).
Then much of their effort was guided by making a new version of the AHAAH
model that would conform to that premise. We note that conformity with their
premise was not data-driven—hearing loss data supporting it—or required by
established scientific principle. The test of a premise, from the theoretician’s
viewpoint, is whether or not it leads to accurate predictions. As we will see, the
premise that the peak pressure level is controlling is simplistic and results in
predictive errors.
The paradoxical reduction in hazard with increased peak pressure has not only been noted, discussed, and explained in talks and in print over a number of years (Price 2007), but it is also consistent with other impulses that produced similar outcomes (Nixon and Sommer 1973; Price 2003). For all the waveforms we have tested with the AHAAH model, if a waveform increases in sound pressure level with no change in shape, the predicted hazard growth is indeed monotonic. The critical point in the case of the AS is that the waveforms changed shape with AS Level, which explains the apparent non-monotonicity. At the higher AS Levels, the shape of the waveform was such that the low-frequency elements forced the stapes into clipping, which meant that fewer damaging oscillations were transmitted to the cochlea, and the prediction was that the impulse with the higher peak pressure was in fact safer for the ear. The model’s displays clearly demonstrate this phenomenon. We have no explanation as to how Zagadou et al. overlooked these publications and the logic supporting the contention, given that they cite an article that explained the very problem they purported to address (Price 2007).

Significantly, the data provide at least partial support for the prediction of reduced effect with increasing AS Level for the 1-m impulses. In the AS dataset, not all conditions in the stimulation matrix were in fact run, which means that the model’s prediction of decreasing hazard could not be fully tested. However, some of the data did apply. As predicted by the model, a lower-level impulse (1 m, Level 5) is in fact more hazardous than the next higher level (about 45% of the Ss showed loss at Level 5, decreasing to about 35% at Level 6) (Price 2007). More data would have been desirable, but these are consistent with the model’s insight. These data and the studies cited earlier all are consistent with the premise that higher peak pressure levels (even with similar spectra) are not necessarily associated with higher hazard. We conclude that the evidence so far is that the problem of non-monotonicity Zagadou et al. raised is in fact a non-problem. Yet, it was the underlying assumption that drove the proposed changes to the model and a basis for the claims by Zagadou et al. that they had improved the model.

3. Problems with the Model’s Variables?

The first major project Zagadou et al. set out to accomplish was to improve the model’s performance by updating the physiological variables so that they better matched the human ear’s characteristics and thereby improved the model’s performance in calculating various transfer functions. The goal was worthwhile; however, valid tests require that the model being tested is understood and accurately used. Unfortunately, there is abundant evidence in their article that Zagadou et al. have neither understood the original model nor have they actually used it in their tests. In the following, we examine the specific questions.

Approved for public release; distribution is unlimited.
3.1 Basic Understanding

Consider first the question of their basic understanding of the model. In one instance they maintain that “the parameters of the current AHAAH model, originally developed for the cat, were assigned to the human model by assuming that cats and humans have the same ear properties”. Their assertion is simply not true. As an electroacoustic model of the ear, conformal with the ear’s physiology, the values in the original cat model were indeed appropriate to the cat ear. That model was developed to produce good fits to the transfer functions available for that ear (Kalb and Price 1987; Price and Kalb 1999). Beginning in the late 1980s, the full software codes were shared with Dr Chan, one of the article’s authors, through the workings of a North Atlantic Treaty Organization Research Study Group. The values of the parameters in the human model have always been appropriate to the human ear and produce good fits to the transfer functions for that ear. Because of the common physiology of mammalian ears, the cat and human models have a similar structure, but the specific values in the computational models are appropriate to each species. For some values in the model (about 20%), the 2 ears were essentially the same (resting strain of the annular ligament, breaking strain of the annular ligament, damping/compliance ratio, etc.). In spite of the fact that the variables in the model are listed specifically in a coefficients file, Zagadou et al. seem to have not noticed that about 80% of the variables differ between the species, a point essential to understanding the essence of the model.

In the spirit of full disclosure, we note that we did incorporate the implicit assumption that the hair cell in the cat and the hair cell in the human are essentially similar structures and would respond to mechanical stress similarly. The hair cells’ microscopic anatomy is highly similar, and we know of no data that suggest that they behave differently. The model calculates basilar membrane displacement (which is the forcing function for hair cell displacement). Thus, when the transfer functions for the transmission of sound from air to the inner ear are right for each species, the assumption of similarity of hair cell properties allows the data from exposures to the cat ear to be interpreted with only a modest extrapolation to the human case.

3.2 Middle Ear Muscle Latency Inaccuracy

In a second instance, they state that the maximum gain in the middle ear muscle response is “reached with a 9 ms delay”. That is simply not the case. It is true that there is a latent period of 9 ms before the middle ear muscle response begins (a short latency appropriate at the very high intensity of these impulses). However, the middle ear muscle response then grows exponentially with a time constant of
11.7 ms (both the latency and time constant are variables in the AHAAH model’s coefficients file). This time constant was developed based on the experiments of Dallos (1964) on the dynamics of the middle ear muscle reflex. Time to a full contraction is about 50 ms from the beginning of the contraction. It is difficult to see how Zagadou et al. could have missed these values because in addition to being in the coefficients file, they are also displayed graphically on every analysis done with the AHAAH model.

### 3.3 Overlooking Insights from Animal Tests

There is another sense in which Zagadou et al. appear to not have understood the scientific strength of the AHAAH modeling approach. The conceptual structure of the approach is very important because it enabled real tests of the model with live ears and appropriate stimuli critical to its development and validation. In contrast, Zagadou et al., in the end, without discussing the intra-cochlear elements of the AHAAH model or providing evidence for shortcomings, recommend the use of intra-cochlear energy as the metric for hazard. Further, their position was based on very few pieces of TS data: 2 rifle fire exposures and the 28 ears that showed a TS of 15 dB or greater in the AS (Price 2003). Understandably, there are essentially no human data in which real damage has been observed in experiments. But the history of experimental exposure of human ears is much greater than the studies Zagadou et al. cite (Price 2007).

Fortunately for science, however, in developing the cat model, we were able to subject cat ears to real stresses for a variety of impulses that did produce varying degrees of loss in more than 200 ears (Price and Wansack 1989; Pierson et al. 1995; Price and Kalb 1999; Price 2003). That is the experimental basis that has produced the AHAAH model’s loss formula that connects calculated displacements to damage. For 12 different exposures to 12 groups of 10 pairs of cat ears, the correlation between the calculated value and the mean immediate TS for each group was $r = 0.94$ (Price 2003). Furthermore, the correspondence between the change in the auditory sensitivity was matched by hair cell losses (Price 2006). The very high correlation between the model’s predictions and actual losses indicated that the majority of the variance in the hearing loss data was being explained by the model, which in turn argued that it should be created in a human form and tested. The modeling approach ultimately recommended by Zagadou et al. simply ignores this work. It appears that Zagadou et al. simply missed this powerful method of connecting to the cat model and the considerable database associated with it.
4. The Non-improvement Improvement

The authors also maintain that they have improved the model by getting the circumference of the human head correct (their Table 2). This assertion is truly baffling. The AHAAH model for the human ear has neither a head circumference nor a radius in its variables, so we are at a loss to understand what changes they actually made to whatever model they were working with. Within the AHAAH model, the acoustic effects of the head are calculated from the variables associated with the diffraction field around the head. It is one thing to overlook something in an account but quite another to positively assert it, especially when it is contrary to fact.

5. The Stapes Displacement Problem

Zagadou et al. claimed to have done a thorough analysis of the variables associated with the AHAAH model and improved its fit to the newest human data—a worthy exercise. The 7 changes to the middle ear elements they propose are found in their Table 2.

One of the elements they “corrected” was the area of the stapes. Two comments apply. In changing from the centimeter-gram-second to the millimeter-kilogram-second system of dimensions, they wound up with a stapes 100 times too large. We believe that the correct physical area of the stapes footplate for the human ear is 3.2 mm² (in the old system). The value we actually used in AHAAH was 2.1 mm², a deliberate choice to reflect what we believe is the effective area for a human stapes operating at very high amplitudes. The problem in the case of the human is that the annular ligament is not uniform in its width or thickness, with the often-observed result that the stapes’ motion is not piston-like but rocking (von Bekesy 1960; Gyo et al. 1987; Heiland et al. 1999; Hato et al. 2003; Lauxmann et al. 2014). The rocking motion is likely to produce a radiation pattern that approaches that of a dipole, which is less effective than a simple piston. Therefore, in the interests of simplicity, sometimes held to be a scientific virtue, we opted to use a smaller area for the stapes’ footplate for calculations at these very high levels. Zagadou et al. seemed to have missed the problem of non-piston-like movement of the stapes at high pressure levels.

6. Non-use of the AHAAH Model

In the interests of scientific curiosity, we made their suggested changes to our copy of the original AHAAH model (AHAAH-PK) and ran it (correcting their area of the stapes to 3.2 mm²) hoping to be able to compare their version of the AHAAH
model with the original. Unfortunately for science, that was not to be, because we discovered that, apparently, they did not actually use the original AHAHH model.

Their Fig. 13 portrays the calculation of the middle ear pressure gain for what they maintain is the original AHAHH model along with their “New AHAHH”. Our Fig. 1 reproduces the data from their figure representing their version of the original model (labeled AHAHH ZCHS) along with the same calculation using the original model (AHAHH PK). The first important observation is that AHAHH ZCHS does not come close to matching AHAHH PK (the 2 models should have been identical). We conclude that what they referred to as the AHAHH model in their article had been changed in some unidentified ways and was not in fact the original model. This is a critical problem for the purposes of their work, because comments regarding the original AHAHH cannot be taken seriously. We are at a loss to explain this problem with Zagadou et al.’s use of the model. It is possible that transcription errors could have occurred or other programming errors could have entered in, but comparison of the model outputs should have quickly identified such problems. The AHAHH model has for many years been freely available for download at the US Army Research Laboratory website. A simple comparison of the output with that model with their version of it on virtually any waveform could have alerted Zagadou et al. that there was a serious problem with their version of the AHAHH model.

Fig. 1  Stapes velocity transfer functions for what Zagadou et al. claimed to be the original AHAHH model (AHAHH ZCHS) and the same calculation with the actual AHAHH model (AHAHH PK)
7. Non-improvement of AHAHAH-PK

A critical focus of the Zagadou et al. work, however, was to improve the fit of the model to “the latest human data”. If we continue to focus on the middle ear pressure gain (their Fig. 13), we discover that they managed to reduce the goodness of the fit between the model and the data rather than improve it. In our Fig. 2 we have plotted the middle ear gain for the Aibara et al. (2001) data on 12 cadaver ears that they were trying to match. For comparison, we also plotted the calculated responses of the “new” AHAHAH (their improved version of the model), the original AHAHAH PK, and the work of Kringlebotn and Gundersen (1985) (data from 68 cadaver ears and used in the development of the original AHAHAH). Visual inspection of the curves shows that at frequencies below 1 kHz the AHAHAH-New curve sits well above the Aibara et al. data as well as the Kringlebotn and Gundersen data and shows peak sensitivity at a lower frequency than all the other curves. As a way of quantifying the fit, we calculated the correlation coefficients between the curves. The correlation between Aibara et al. and AHAHAH PK is 0.91, while the correlation between Aibara et al. and AHAHAH-New is 0.79—clearly not an improvement. If we compare the models’ predictions with Kringlebotn and Gundersen, the correlation with AHAHAH PK is still higher (0.93 vs. 0.86 for AHAHAH-New). These data support the contention that in spite of their expressed intentions, Zagadou et al. did not in fact improve the fit of the model to at least these transfer function data. The original model (AHAHAH PK) fits the Aibara et al. and the Kringlebotn and Gundersen datasets much better than Zagadou et al.’s “improved” version.
Fig. 2  Comparison of stapes velocity transfer functions for Zagadou et al.’s AHAHAH-New, the original AHAHAH (AHAHAH PK), the human data of Aibara et al. (2001), and Kringlebotn and Gundersen (1985)

In passing, we note that the data from Aibara et al. might not be the ultimate data to match for 2 reasons. First, as we noted earlier, there is general agreement that the human annular ligament is not symmetrical, with the result that the motion of the stapes is not that of a simple piston but rather includes rocking elements in 2 planes, especially at high amplitudes. The test fixture used by Aibara et al. appears to have held the head of the stapes so as to limit it to piston-like movement in one plane, producing a motion not truly characteristic of the human stapes. Second, the measurements of compliance were made using very-low-frequency driving (near 1 Hz). Measurements of annular ligament compliance at very low frequencies approaching steady state (Gan et al. 2011) and subsequent analysis using temperature–frequency superposition (Zhang and Gan 2014), show stiffening of the annular ligament with increasing frequency. It seems to us that the performance of the ligament needs to be evaluated at realistic auditory frequencies, at least for the purpose of predicting conductivity in the normal auditory range.

In the end we conclude that whereas Zagadou et al. set out to improve the fit of the AHAHAH model to the conduction data, however laudable their intent, they simply failed to produce a better fit than the actual original model.
8. Non-replication of the AS Pressure Histories

A major area of critical concern with the work of Zagadou et al. was their apparent failure to reproduce the waveforms from the AS—a task central to their experimental design. Zagadou et al. invested a great deal of effort attempting to recreate the impulse-producing apparatus used in the AS. This involved C4 explosive charges detonated in particular structures that were 1, 3, or 5 m from the Ss’ ears. Their experimental design required them to recreate the AS pressure histories, measuring them in the free-field and at the eardrum position of an ATF fitted with the same type(s) of HPD used in the AS (as well as a helmet and goggles, which the Ss also wore). If the ATF were an adequate substitute for the human Ss, they argued that the waveforms could legitimately be analyzed using AHAH’s algorithms permitting data to be entered at the eardrum position. And if those waveform data matched the AS data, they had proposed to use the TS data from the AS to validate and/or adjust the model. This approach did have the effect of not requiring the testing of additional human Ss, but it does require scrupulous attention to the details of the physical measurement. It is obviously critical that the pressure histories match those of the AS if we are to allow them to be meaningful predictors of the hearing loss data from the AS. Following an analysis of the limited data available in their report, we believe that they failed to reproduce the AS in numerous ways, so much so that we conclude that they simply did not in fact reproduce the AS stimulating conditions as required. We offer several examples of this problem in the following.

8.1 Measurement Issues

8.1.1 Microphone Use

First we address the nonstandard use of pencil gauges in making the free-field measures. The pencil gauge is designed to be pointed directly at the impulse source so that the incident wave passes the sensor’s surface at grazing incidence. In Zagadou et al.’s Figs. 2 and 3, which picture the gauges in use, we note that they are pointed horizontally in the direction of the impulse source. In the case of the 5-m condition, the actual impulse source was considerably elevated above the horizontal, which means that the pencil gauges were susceptible to off-axis effects. In the original AS, measurements were made with pancake gauges oriented vertically, which minimized the off-axis influences.
8.1.2 Apparatus

Second, we note that in spite of Zagadou et al.’s assertions that the stimulating apparatus was faithfully reproduced, the apparatus for the 1- and 3-m conditions was significantly different from that used in the AS exposures, which would alter the true exposure waveforms (see their Fig 2.). In the original AS apparatus, the Ss were seated next to a vertically oriented, mortar-like tube on a raised platform made of expanded metal, which was essentially acoustically transparent. The muzzle of the “mortar” was about 1.5 m above the platform, which was in turn a little less than 2 m above a hard surface. Thus, the ground reflection in the AS data had a path length of about 7 m (probably somewhat more, because the effective center of the stimulating impulse is above the muzzle by some amount). However, in the setup pictured in the Zagadou et al. article, there appears to be a hard, deck-like reflecting surface at the level of the platform rather than the expanded metal used in the original apparatus. Thus, the ground reflection was closer by about 4 m, which to a first approximation would mean that the reflected waveform would appear to be about 10–12 ms earlier in the waveform than it did in the AS. In deciding whether or not they had reproduced the AS impulses, Zagadou et al. used metrics that examined only the initial peak in the waveform.

The AHAAH model indicates that the ground reflection is a nontrivial issue, even though the reflected wave is much less intense than the initial waveform. The problem is that at these high pressure levels (170+ dB), peak-clipping by the stapes produces a waveform entering the cochlea in which only a small percentage of the hazard is calculated as being due to the initial peak of the waveform, and a majority of the effect is assignable to the latter part of the waveform (which is still very intense) (Price 2012). So the reflected energy may be critical to understanding the hazard from a particular stimulating situation. It appears that in this regard, Zagadou et al. have not reproduced the stimulating conditions in the AS.

8.1.3 ATF Issues

Third, we observe that the data from the ATF they used were grossly distorted by some combination of unspecified factors. Zagadou et al. in effect noted this problem in their data, and as a result, they found it necessary to “extrapolate” lower-pressure-level data to achieve the higher-pressure-level data they needed! We comment more on this extraordinary procedure later, but there really is no substitute for good measurements. Their Fig. 6 displays the transfer functions measured with the ATF, and it is clear that their data, especially at the lower frequencies, did not come close to matching the normal transfer function for the human ear. We have no way of diagnosing the problem(s) for certain, but such problems can arise when the ATF design does not sufficiently isolate the “back path” to the microphone.
Such isolation is very hard to achieve, especially for the low frequencies. We also note that the distorted data they cite were produced by impulses from a shock tube operating at much lower pressures (155–160 dBP) than those in the AS dataset (178–195 dBP). The data they have shown appear to be full of measurement artifacts, which should preclude their use in a test such as this. Furthermore, it is reasonable to suppose that the measurement artifacts must have gotten worse as the pressures rose to the Levels actually used in the AS. Our experience has shown that valid pressure histories are admittedly difficult to make in this pressure region, especially under conditions where heat from the exploding sources can itself cause the microphone output to distort. Distortions of this sort can even be seen in some of the original AS data.

It also occurs to us that the problem of simulating the stimulating conditions in the AS was perhaps even impossibly complex. Not only did the Ss wear an HPD, but they also wore combat helmets and protective goggles. The goggles and helmet were part of the immediate acoustic environment and their interaction with the ATF may have produced unexpected effects, given that skin on the scalp and hair are not normally part of an ATF. Also, at the very high pressures used in these studies (up to 195 dB, or more than 100 kPa), it was possible to get physical displacement of objects in the sound field.

8.1.4 Distorted Pressure Histories

The foregoing problem with measurements becomes more apparent when we examine the pressure histories Zagadou et al. actually used in their calculation and the ones they provide in the article. We read the pressure history data in their Fig. 9, which represented pressures at the eardrum position for the 5-m data and the defeated muff. Fortunately, samples of the waveforms from the original AS have been made available to the research community, and we were able to compare them with the Zagadou et al. data. A careful examination reveals a host of serious problems:

- The growth of the peak pressure level (PPL) of the impulse does not match the AS data. Zagadou et al. show essentially no growth of PPL at the ear drum between Level 1 and Level 4 (PPL in the free-field in the AS rose 8.3 dB between those levels, and the corresponding PPL under the muff at the Ss’ tragus rose 7.2 dB over that same interval). The lack of growth in the data of Zagadou et al. makes no physical sense and runs counter to their claim that they matched the growth of PPL.

- When faced with artifacts in their pressure history data, Zagadou et al. did a most extraordinary thing—they “extrapolated data”. Their extrapolation,
however well-intentioned, was based on linear thinking in a world of nonlinear processes, not the least of which was the muff themselves (an issue the authors neither noted nor commented upon). Consider the nonlinearity introduced by the muff. If we use the waveforms from the original AS, we find that if we compare the A-weighted energy in the impulse outside the muff with the A-weighted energy at the tragus, we get an estimate of the attenuation provided by the muff (for the record, peak pressure measures behave similarly). The effects were nontrivial. In the case of the 1-m data, the attenuation of the muff improved 11 dB, while the outside energies rose by 18 dB. Under the muff, then, the total rise of A-weighted energy was only 7 dB! In contrast, for the intact muff in the AS, the attenuation became less by 5.4 dB, which means that the energy rose by 23 dB under the muff for a nominal 18-dB rise outside it (Price 2007). Assumptions concerning the pressure histories, under these circumstances, can only be viewed as very risky, especially if the exact shape of the pressure history matters. In any event, Zagadou et al. made linear assumptions and, not surprisingly, their results reflect their assumptions. The PPLs for their Levels 5–7 (at the eardrum) rose 5.6, 3.1, and 2.5 dB, respectively (11.2 dB total). For the same data range, the free-field pressure (FFP) data from the AS showed rises of 2.6, 3.3 and 2.0 dB, respectively (7.9 dB total), and the PPLs at the tragus in the AS rose 1.5, 4.5, and 1.6 dB, respectively (7.6 dB total), with pressure growth at the tragus paralleling the free-field. Arguably, the data from Zagadou et al. showed a rise about 3.5 dB higher than the measured responses. Such data can hardly be used with confidence.

- Next we note with surprise that the absolute PPLs in the Zagadou et al. data (for pressure at the eardrum under a protector) are higher than the corresponding FFPs in the AS—by 1.3–8.7 dB! Further, if one looks at the PPLs at the tragus for the AS studies, the PPLs of the Zagadou et al. data (nominally at the ear drum) were 12–22 dB higher than the AS data at the tragus. It is obvious through simple inspection that something is drastically wrong with the data at the eardrum of the ATF. Others using the same ATF did not find such an aberration in PPLs (Murphy et al. 2015). And when we have exercised the AHAAH model to predict pressure histories at the eardrum, we find that, not surprisingly, there is little difference predicted between pressures at the tragus and at the eardrum. We will cover more of this later.

- Furthermore, the wave shape of the original AS data at the tragus had a slow rise and smooth, rounded form, and an extended duration, as one would
expect to find in a pressure history measured under a protector. In contrast, the Zagadou et al. data look essentially like free-field waveforms with significant high-frequency content (unlikely under a muff). When we tried using AHAAH to predict pressure histories at the eardrum, given pressures at the tragus, the calculated shape changed little. We can conceive of no physical way that such a pressure level and shape change could have occurred in a proper measurement.

- The problems of this sort are not confined to the 5-m data. Zagadou et al. reported no waveform data for the 3-m condition, but their Fig. 8a contains overlaid waveforms for the 1-m condition. We have digitized their waveforms from the printed figure, with the obvious limitations on accuracy. That need not concern us given that the issues that concern us are not subtle. In our Fig. 3, we show their 1-m Level 1 waveform at the eardrum with our calculated waveform (using pressures at the tragus from the AS) for the same condition. In this case, the lowest pressure level for this condition, we find that the agreement between the 2 pressure histories is quite good.

![Fig. 3 Eardrum pressure for 1 m, Level 1, from the ATF (blue curve) and the eardrum pressure calculated using the original AHAAH model (red curve) and the AS tragus data](image)

In our Fig. 4, we compare our calculated 1-m Level 6 waveform with theirs, and the difference is substantial. The PPL of their waveform is 120 kPa, about 25 kPa higher than the free-field pressure and more than 100 kPa (18 dB) higher than the calculated pressure. Furthermore, the A-duration of Zagadou et al.’s impulse was less than 1 ms—the same as that of the FFP—while the calculated pressure had a rounded leading edge and an A-duration of a little over 2 ms (matching the A-duration at the tragus in the AS). Clearly some consequential measurement issues developed as the pressure rose. These data contain serious anomalies.
The foregoing points make the case that Zagadou et al.’s measurement system was defective, and that they did not in fact reproduce the AS data. These shortcomings invalidate any subsequent analysis using these data in assessing the sensitivity of the human ear to damage from intense impulse noise (as Zagadou et al. have done).

Furthermore, it is difficult to conceive how this critical set of issues could have escaped their attention and brings into serious question their concluding recommendation that the “use of the ATF method to collect eardrum data as model input will correctly account for the HPD and orientation effects”. The use of an ATF has much potential in the assessment of HPD effects as well as the complexities of exposures that do not occur in the free-field. However, there are still many technical issues that need to be addressed—especially in the case of very intense stimuli—before a standard ATF can be agreed upon.

9. Philosophical Issues

The logic underlying the design of these experiments seems problematic in that it pervaded and confused the scientific effort, leading to what we believe are invalid outcomes. One set of issues involves the experimental technique and the other the presumptions in their analysis. We cover both in turn.

9.1 Experimental Technique

The study set out to critically examine the AHAAH model and improve its fit to the human hearing loss data. That is a worthwhile focus. Their experimental approach included reproducing the stimuli used in the AS—possibly a useful effort, depending upon the use to which such data might be put. Certainly, there have been
questions regarding the true variability in the AS style exposures (Price 2010). Also, it would be good to have waveforms measured under the nonlinear earplugs tested in those studies (not currently available) and to measure the response of additional protectors that might be tested as well. This research might have produced interesting and useful data. Though they said they made such measures, they did not report them.

However, Zagadou et al. made the G.R.A.S. ATF an integral part of their measurement setup. It would be fair to say that there is no consensus in the technical community that the perfect ATF has been developed and that differences between ATFs constitute a significant element of variability (Murphy et al. 2015). One of the difficulties in using such devices at these high pressures is in the nonlinearity of microphones used in them. In an “unprotected” exposure, the PPLs can be very high indeed. There are gauges that can meet the requirement for the high peak pressure, but they tend to be less sensitive and have a poor signal-to-noise ratio at the lower pressures encountered with less-intense stimuli and protected exposures. The AHAAH model, furthermore, has made the measurement world more difficult by requiring the measurement of very-low-frequency energy (down to about 1 Hz). That energy does not itself produce hearing loss, but it does affect the conductive path by causing the stapes suspension to clip, which modulates the flow of higher-frequency energy that does cause intra-cochlear damage. A real series of technical challenges needs to be worked out before measurements can be made confidently. There is also the question regarding the back-path conduction (equivalent to “bone conduction”) of the ATF. Controlling the microphone’s exposure to these influences is essential to making valid measurements. As we have seen, the data that Zagadou et al. report in their Fig. 6 suggest that this was a real problem in their experiments. In nonlinear systems one cannot simply assume away the problem of a measurement error. In our estimate, this is another fatal flaw in the execution of the study.

9.2 Assumed and Extrapolated Data

The measurement problem is compounded by the authors’ willingness to assume data rather than to strictly adhere to empirical evidence. For example, they comment that in interpreting the hearing loss data from the AS, they analyzed them “in accordance to the original approved statistical protocol”. That means that they assumed that energy is the critical variable for an ear’s susceptibility and proceeded to create data consistent with that presumption. In the original AS, if an ear failed at a particular energy level in the exposure matrix, for statistical purposes it was presumed to fail at all higher levels as well. This meant that more than half of the failures in their statistical analysis had been assumed rather than observed. For the
purposes of a simple statistical analysis, in the words of Hamlet, “tis a consummation devoutly to be wished”. Unfortunately for this approach, the presumption of failure at higher levels was in fact not met in the AS data. A detailed analysis of the data has shown that there were 28 failures (greater than 15 dB of TS). Of those, 25 went on in the protocol to finally fail at higher levels, sometimes much higher levels. Only 3 ears met the presumption in the experimental design and failed when re-exposed to the same energy (Price 2010). Clearly the presumption of failure at all higher levels based on energy at one level is not supported by the AS data. The problem has been noted in the literature (Price 2007); yet in their analysis, Zagadou et al. embodied the presumption that energy in the free-field was the controlling variable (which was clearly not the case).

Likewise, in our discussion of their creation of data for the 5-m condition (their Fig. 9), we have seen that the outcome is clearly inappropriate (e.g., apart from the inappropriately high levels, the expected growth of PPL with intensity was only present in the data they created). The circularity of their argument makes it unsurprising that they were able to find a match between their model and energy. Such procedures are anathema to scientific rigor.

9.3 The Physiology of Loss

There is another critical issue overshadowing Zagadou et al.’s model and the analytic approach they recommend, namely the issue of what has been termed a “critical level” (CL) for the ear. A lot of data, from multiple species including guinea pig, monkey, cat, and man, for various measures including histology, electrophysiological potentials, rate of growth of loss, recovery times, and hearing thresholds, all agree that there is a level/zone in which the loss mechanisms change from metabolic exhaustion to mechanical stress (Price 1981). Simply put, the ear goes from being tired out to being torn up. Hence, it is reasonable to expect to find that a metric appropriate to the lower level will not apply at the higher level. Experience with exposure to factory noises has shown that one could well expect ears to recover from a 35- or 40-dB TS every working day for many years with a permanent loss gradually growing over a working lifetime. However, experiments with animal ears have shown that the same level of TS in response to very intense sounds, such as gunfire, would produce about 20 dB of permanent TS from one exposure.

Consider the rate of loss once the CL has been crossed. Data from the human ear, the chinchilla ear, and the cat ear agree that for a given exposure, the loss measured immediately after exposure grows very rapidly with increases in level: about 6 or 7 dB for every decibel increase in PPL of the exposure (Price 1981; Hamernik
et al. 1987; Price and Wansack 1989). Thus, once the threshold for mechanically based loss has been reached for some exposure, a 10- to 15-dB increase in PPL is the difference between no loss and total loss with no hair cells left in that part of the cochlea—from one exposure. Clearly the processes are different on either side of the CL, and those differences can be ignored only at great peril.

Zagadou et al. neither note nor account for the issues associated with a CL. For instance, in their model for loss, the function that relates hazard to the level and number of impulses is what they refer to “the traditional $10\log_{10}(N)$ rule”, where $N$ represents the number of shots. One of the strong observations for a CL is that the rate at which loss accumulates goes from linear in log-time (like the “traditional rule” Zagadou et al. commend) to a function that is linear in time (for multiple measures and different types of ears [see Price 1981]). Price and Kalb (1991) have argued that this change is consistent with the mechanism of mechanical fatigue and have embodied it in the cochlear portion of the AHAAH model. Zagadou et al. essentially reject this element without noting or commenting on it.

We also believe that their approach would produce only a very bad fit to the hearing loss data should they try to plot a full range of them. In their Figs. 16 and 17, they plot a logistic regression analysis of the very small set of hearing loss data they have developed. We note that what has been plotted is essentially a population susceptibility curve portraying the percentage of the population that would reach a threshold of damage on the ordinate given an increasing exposure on the abscissa. If we read from their plots, the interval between a 5% incidence to a 95% incidence is a range of about 65 dB. The data from the human ear relating to a CL (Price 1981) as well as data from the cat and chinchilla ears (Hamernik et al. 1987) show that once you have reached the zone of the CL, the 5%–95% interval in a population is much closer to 20 dB than the 65 dB plotted by Zagadou et al. The 20-dB range is, in the AHAAH model, held to be indicative of a susceptibility normally distributed with a 6-dB standard deviation. As of this writing, that analysis fits the data. We do not doubt that ranges like Zagadou et al. plot may exist where the actual exposure conditions are not well defined (angle of exposure is not well controlled, middle ear muscle activity is not accounted for, etc.), but in the present case where the conditions are clearly specified, their assumed range of 65 dB appears to be much too large.

9.4 Middle Ear Muscle Activity

Finally, in the area of what we believe are bad assumptions, Zagadou et al. have essentially ducked the question of middle ear muscle activity. Like Scarlett O’Hara, they have chosen to “think about that tomorrow”, and have nonetheless gone
forward with a specific, highly questionable viewpoint. It is true that, like many other aspects of hearing loss to intense sounds, more data with respect to human middle ear muscle effects would be desirable. Yet recent data have shown that 94% of the human Ss tested have a middle ear muscle reflex (Flamme 2015) and a number of other studies have shown them to be conditionable in animals and man (Djupesland 1964, 1965; Brasher et al. 1969; Yonovitz 1976). Beyond that, the cognitive capacity of the human adds to the probability of a conditionable response. For instance, it has been shown that Ss’ middle ear muscles contracted as they contemplated handling a toy that was thought to be noisy (Marshall et al. 1975). The evidence suggests that it is reasonable to expect the middle ear muscles can contract in advance of the auditory stimulus. In fact, it would be difficult to make the argument that we should not expect a contraction in advance of the stimulus! Interestingly, we note that the original AHAH model (for the cat ear) was developed with data in which experimental designs were arranged to specifically avoid middle ear muscle involvement (animals were anesthetized at the time of exposure). To obtain a realistic response for a normal population, the muscle effects were then added once the pure cochlear loss mechanisms were quantified.

In any event, the AHAH model has been validated with the active middle ear muscle effects as an option, and the acceptable levels within it reflect the appropriate presence (or absence) of the response (Price 2007). Actually, the basic question is whether or not a conditioned response can be demonstrated. The unconditioned response (as in a string of continuing impulses, or even an exposure in the presence of a moderately loud background sound) is not really debated. It seems to us rather perverse, in the presence of so much positive information, to simply conduct an analysis assuming that there was no muscle response. Due scientific diligence would surely have commended that in the presence of doubt, both analyses be performed and the differences examined. Zagadou et al., again, assumed an outcome and did not qualify or limit the application of their proposal.

### 10. Conclusions

The work reported in the Zagadou et al. article was nominally aimed at producing an improved damage risk criterion for human exposure to intense impulsive sounds. We take no pleasure in concluding that the scientific rigor in this work was woefully deficient, and as a result this work failed to deal with the problems it set out to address.

Zagadou et al. began with the premise that the “non-monotonic” response of the AHAH model (presently used as a hazard assessment tool) was evidence of a serious problem. We have seen that it has long been known that there was no true
non-monotonicity and furthermore that the available data did in fact fit the model’s prediction.

Zagadou et al., in fact, did their work on an unknown version of the AHAAH model that differs in unspecified ways from the official version. Any comments regarding problems with or improvements to AHAAH are therefore moot.

Zagadou et al. claimed to improve the model’s fit to the transfer functions for the human ear. We have seen from a transfer function that they published that they did not in fact use the original model. Yet it should have been obvious to them that this error existed. Furthermore, we showed that the changes they made, for at least the data to which we had access, resulted in a poorer fit to one of the critical transfer functions for the human ear. Their efforts at model improvement produced a demonstrably less good fit to the data.

Zagadou et al. proposed to reproduce the AS exposure stimuli and measure the acoustic response at the eardrum location of an ATF. We have seen that the data they have published are rife with serious artifacts and cannot in any sense be considered to have been a replication of the AS data.

Finally, we have shown that in the interpretation of the data and the analytical procedures they used, Zagadou et al. simply ignored relevant data on the ear’s response to intense sounds and consistently made assumptions that favored a bias in the direction of energy explaining everything—a problem of logical circularity.
11. References


### List of Symbols, Abbreviations, and Acronyms

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHAHAH</td>
<td>Auditory Hazard Assessment Algorithm for Humans</td>
</tr>
<tr>
<td>AHAHAH ZCHS</td>
<td>article AHAHAH model</td>
</tr>
<tr>
<td>AHAHAH PK</td>
<td>original AHAHAH model</td>
</tr>
<tr>
<td>ANSI</td>
<td>American National Standards Institute</td>
</tr>
<tr>
<td>AS</td>
<td>Albuquerque Studies</td>
</tr>
<tr>
<td>ATF</td>
<td>acoustic test fixture</td>
</tr>
<tr>
<td>CL</td>
<td>critical level</td>
</tr>
<tr>
<td>FFP</td>
<td>free-field pressure</td>
</tr>
<tr>
<td>HPD</td>
<td>hearing protection device</td>
</tr>
<tr>
<td>PPL</td>
<td>peak pressure level</td>
</tr>
<tr>
<td>S</td>
<td>subject</td>
</tr>
<tr>
<td>TS</td>
<td>threshold shift</td>
</tr>
<tr>
<td>1</td>
<td>DEFENSE TECHNICAL INFORMATION CTR DTIC OCA</td>
</tr>
<tr>
<td>2</td>
<td>DIRECTOR US ARMY RESEARCH LAB RDRL CIO L IMAL HRA MAIL &amp; RECORDS MGMT</td>
</tr>
<tr>
<td>1</td>
<td>GOVT PRINTG OFC A MALHOTRA</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LAB – HRED RDRL HRB B T DAVIS BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LAB – HRED RDRL HRO COL H BUHL RDRL HRA I MARTINEZ RDRL HRR R SOTTILARE RDRL HRA C RODRIGUEZ RDRL HRA B J HART RDRL HRA A C METEVIER RDRL HRA D B PETTIT 12423 RESEARCH PARKWAY ORLANDO FL 32826</td>
</tr>
<tr>
<td>1</td>
<td>USA ARMY G1 DAPE HSI B KNAPP 300 ARMY PENTAGON RM 2C489 WASHINGTON DC 20310-0300</td>
</tr>
<tr>
<td>1</td>
<td>USAF 711 HPW 711 HPW/RH K GEISS 2698 G ST BLDG 190 WRIGHT PATTERSON AFB OH 45433-7604</td>
</tr>
<tr>
<td>1</td>
<td>USN ONR ONR CODE 34 P MASON 875 N RANDOLPH STREET SUITE 1425 ARLINGTON VA 22203-1995</td>
</tr>
<tr>
<td>1</td>
<td>USA HQDA ARI M SAMS 6000 6TH STREET BLDG 1464 FT BELVOIR VA 22060</td>
</tr>
<tr>
<td>1</td>
<td>USA NSRDEC RDNS D TAMILIO 10 GENERAL GREENE AVE NATICK MA 01760-2642</td>
</tr>
<tr>
<td>1</td>
<td>OSD OUSD ATL HPT&amp;B B PETRO 4800 MARK CENTER DRIVE SUITE 17E08 ALEXANDRIA VA 22350</td>
</tr>
<tr>
<td>12</td>
<td>DIR USARL RDRL HR L ALLENDER P FRANASZCZUK RDRL HRB J LOCKETT RDRL HRB A M LAFIANDRA RDRL HRB C J GRYNOVICKI RDRL HRB D D HEADLEY RDRL HRF K OIE RDRL HRF A A DECOSTANZA RDRL HRF B A EVANS RDRL HRF C R PRICE J GASTON RDRL HRF D A MARATHE</td>
</tr>
</tbody>
</table>

Approved for public release; distribution is unlimited.
INTENTIONALLY LEFT BLANK.