Comparison of the CAM2 and NAL-NL2 Hearing Aid Fitting Methods

Brian C. J. Moore¹ and Aleksander Sęk¹,²

Objective: To compare preference judgments for sounds processed via a simulated five-channel compression hearing aid with gains and compression ratios selected according to two recently introduced fitting methods, CAMEQ2-HF (hereafter named CAM2) and NAL-NL2.

Design: There were 15 participants with sloping sensorineural hearing loss. They had mild losses, typical of people who might be candidates for wide-bandwidth hearing aids. Within a given trial, the same segment of sound was presented twice in succession to one ear, once with CAM2 settings and once with NAL-NL2 settings, in random order. The participant had to indicate which one was preferred and by how much. Judgments of overall sound quality were obtained for female and male speech in quiet and for four types of music (classical, jazz, a man singing, and percussion). Judgments of speech clarity were obtained for female and male speech in speech-shaped noise, female speech in a male-talker background, and male speech in a female-talker background. Factors investigated included compression speed (slow or fast) and input sound level (50, 65, or 80 dB SPL).

Results: The pattern of the results was reasonably consistent across participants, but the magnitude of the effects was small. For judgments of overall sound quality, nine participants preferred CAM2 relative to NAL-NL2, and the remainder showed no clear preference. There was a significant overall preference for CAM2. The preference for CAM2 over NAL-NL2 in overall sound quality was present for all types of stimuli, both compression speeds, and all three levels. For judgments of the clarity of speech in noise, five participants preferred CAM2 over NAL-NL2, one showed the opposite preference, and the remainder showed no clear preference. There was a significant overall preference for CAM2. The preference for CAM2 over NAL-NL2 in the context of clarity of speech in noise was present for all types of stimuli, both compression speeds, and all three levels. For judgments of the clarity of speech in a background talker, CAM2 was significantly preferred overall relative to NAL-NL2, but the effect was very small.

Conclusions: For participants with mild sloping hearing loss, a simulated hearing aid unilaterally fitted using CAM2 was preferred over the same aid fitted using NAL-NL2 for overall sound quality and the clarity of speech in noise. Preferences differed only very slightly for the clarity of speech in a background talker. Further work is needed to establish whether similar preferences would be found in everyday life.

INTRODUCTION

Most modern hearing aids incorporate multichannel amplitude compression to compensate for the loudness recruitment and reduced dynamic range associated with cochlear hearing loss (Dillon 2001; Moore 2007). The initial fitting of such hearing aids is usually based on the audiogram of the client, either using a published fitting formula, such as DSL V5 (Scollie et al. 2005), NAL-NL1 (Dillon 1999; Byrne et al. 2001), or CAMEQ (Moore et al. 1999; Moore 2005), or using a manufacturer’s proprietary fitting method. Additional “fine tuning” is usually required to suit the preferences of the individual (Moore et al. 2005), but it is generally acknowledged that a good initial fitting is important (Mueller et al. 1992).

Recently, two new fitting methods have been introduced, the NAL-NL2 method from the National Acoustics Laboratories in Australia (Keidser et al. 2011) and the CAMEQ2-HF method (hereafter called CAM2; Moore et al. 2010b). NAL-NL2 was designed with the goals of making speech intelligible and overall loudness comfortable; these are basically the same goals as for NAL-NL1. To achieve this, a speech intelligibility model, based on a modified version of the Speech Intelligibility Index (ANSI 1997), was combined with a loudness model (Moore & Glasberg 2004) in an adaptive computer-controlled optimization process. The Speech Intelligibility Index model was modified by introduction of an “effective audibility factor,” which takes into account the finding that, as the hearing loss gets more severe, less information is extracted from the speech signal even when the speech signal is above the detection threshold (Ching et al. 1998, 2001; Turner & Cummings 1999; Johnson & Dillon 2011; Keidser et al. 2011). In addition, empirical data on the loudness and quality of sounds amplified according to the NAL-NL1 prescription were taken into account (Smeds et al. 2006; Keidser et al. 2007, 2008). NAL-NL2 recommends less mid-frequency gain and more low- and high-frequency gain than NAL-NL1 does (Johnson & Dillon 2011). NAL-NL2 recommends gain for frequencies up to 8 kHz, whereas the limit for NAL-NL1 is 6 kHz.

CAM2 is conceptually similar to the CAMEQ method (Moore et al. 1999; Moore 2005), which was developed using a loudness model (Moore & Glasberg 1997). The loudness model makes use of two key concepts. The first is the excitation pattern, which can be thought of as representing the distribution of excitation along the basilar membrane evoked by a given sound (Fletcher 1940; Zwicker 1970; Moore & Glasberg 1983; Glasberg & Moore 1990). Within the loudness model, the frequency scale is transformed to an ERBₙ-number scale, which is a scale derived from the equivalent rectangular bandwidth of the auditory filter, ERBₙ, for listeners with normal hearing (Glashberg & Moore 1990; Moore 2012). A one-unit step on the ERBₙ-number scale corresponds to a frequency change of 1 ERBₙ at that center frequency. The unit of the ERBₙ-number scale is the Cam. For example, for a center frequency of 1000 Hz the value of ERBₙ is 130 Hz, so a change from 935 to 1065 Hz represents a step of 1 Cam on the ERBₙ-number scale. The second key concept is specific loudness, which is a kind of loudness density. It is defined as the loudness that would be evoked by the excitation within a 1-Cam–wide range on the basilar membrane, if it were possible to present that excitation alone (with- out any excitation at adjacent regions on the basilar membrane). A plot of specific loudness as a function of ERBₙ number is...
called a specific loudness pattern. In the loudness model, the overall loudness is calculated by summing the specific loudness over the whole range of ERB<sub>n</sub> numbers, which is equivalent to calculating the total area under the specific loudness pattern.

The CAM2 procedure is based on the goal of placing as much of the speech spectrum as possible above absolute threshold for a given overall loudness (Moore & Glasberg 1998). This goal is achieved by amplifying speech so that, on average, the specific loudness is the same for all frequencies within the range 500 to 4000 Hz; this is the range that is most important for speech intelligibility (ANSI 1997). The CAMEQ procedure is also based on the goal of giving approximately the same overall loudness as normal for speech over a wide range of sound levels.

The CAM2 procedure was developed partly to allow the fitting of wide-bandwidth hearing aids that some manufacturers are producing. The main differences between CAMEQ and CAM2 are: (1) CAM2 gives recommended gains for center frequencies up to 10 kHz, whereas the upper limit for CAMEQ is 6 kHz; (2) CAMEQ is based on the assumption that the hearing aid user faces the person they wish to hear, and uses a free-field-to-eardrum transfer function for frontal incidence. CAM2 is based on the assumption that the user may wish to hear sounds from many directions, and uses a diffuse-field-to-eardrum transfer function; (3) CAM2 is based on an improved loudness model for impaired hearing (Moore & Glasberg 2004); (4) CAM2 is based on recent wide band measurements of the average spectrum of speech (Moore et al. 2008). For frequencies above 6 kHz, the gains prescribed by CAM2 are intended to lead to a specific loudness that is similar to or slightly less than obtained in a normal ear without amplification. The gains recommended by CAM2 are close to, but typically 1 to 3 dB lower, than those recommended by CAMEQ for frequencies from 1 to 4 kHz.

A field trial with hearing aids fitted using CAM2 (for frequencies ≤7.5 kHz) showed that CAM2 generally led to satisfactory loudness in everyday life, except for some complaints about the loudness of transient sounds (perhaps related to the slow compression used in that study) (Moore & Füllgrabe 2010). A laboratory study (Moore et al. 2011) suggested that some inexperienced hearing aid users prefer slightly less high-frequency gain than prescribed by the first version of CAM2 (Moore et al. 2010b), and that finding is now incorporated in the CAM2 software; slightly reduced gains are prescribed for inexperienced users. We are not aware of any published studies evaluating the NAL-NL2 fitting method. Despite this, the method is already being adopted by hearing aid manufacturers and dispensers.

For a hearing aid fitting method to be considered satisfactory, several requirements should be met, including: (1) It should lead to good clarity of speech, especially when background sounds are present; (2) It should lead to satisfactory sound quality for a variety of sounds including speech and music; (3) It should lead to satisfactory loudness and quality of sounds for a wide range of input sound levels. The work described in this article addresses the relative extent to which CAM2 and NAL-NL2 achieve these goals by answering the following questions:

(1) What are the relative preferences for sound quality and speech clarity of hearing aid fittings based on CAM2 and on NAL-NL2?

(2) Do relative preferences for CAM2 and NAL-NL2 depend on the input sound level?

(3) Do relative preferences for sound quality depend on the type of signal (speech or music)?

(4) Do relative preferences for speech clarity depend on the type of background sound (noise or a competing talker)?

(5) Do relative preferences depend on the compression speed of the hearing aid?

These questions were addressed using a method of paired comparisons in a laboratory study using a simulated hearing aid. The participants selected had mild hearing losses, typical of people who might be candidates for wide-bandwidth hearing aids (Moore et al. 2008). The simulated hearing aid was individually programmed using the CAM2 and NAL-NL2 methods for each hearing-impaired ear that was tested. The method of paired comparisons provides an efficient and sensitive way of comparing a large number of stimuli and settings in a way that would not be possible in a field trial with wearable aids. Also, the use of a simulated hearing aid makes it possible to achieve target insertion gains very accurately, whereas with wearable hearing aids compromises are often necessary.

The last question described above has received only scant attention in the past. It is known that fast-acting compression can lead to an increase in the loudness of sounds for normal-hearing listeners for a fixed root-mean-square level (Moore et al. 2003), but it is not known how, and by how much, compression speed affects the loudness of sounds for hearing-impaired people when the compression is tailored to suit their hearing loss. The NAL-NL2 procedure incorporates adjustments to the gains based on compression speed, but these adjustments have not been experimentally validated. Also, the adjustments are generally very small (<1 dB) and unlikely to have any noticeable perceptual effect.

**METHODS AND MATERIALS**

**The Simulated Hearing Aid**

As in our previous studies with simulated hearing aids (Füllgrabe et al. 2010; Moore et al. 2010a, 2011), we used an aid with five compression channels. This reflects a compromise between having sufficient flexibility to implement the frequency-dependent gains and compression ratios required to match the prescription while avoiding the reduction of spectral and temporal contrast that can occur when many channels are used (Plump 1988; Bor et al. 2008; Stone & Moore 2008).

The simulated hearing aid was the same as described by Moore et al. (2010a), and was implemented using Matlab. The aid simulator included a digital filter for overall shaping of the frequency response; the insertion gain (IG) for a speech-spectrum noise (SSN) with a level of 65 dB SPL could be set for center frequencies of 0.25, 0.5, 1, 2, 3, 4, 6, 8, and 10 kHz. The IGs were set according to the relevant prescription method, either CAM2 or NAL-NL2. The version of the CAM2 software was 1.0.26.0. The version of the NAL-NL2 software was v2.0 (dll v2.15) (the software was kindly provided by Harvey Dillon and Scott Brewer of NAL). Because NAL-NL2 did not recommend an IG at 10 kHz, the IG at 10 kHz for NAL-NL2 was set equal to that prescribed at 8 kHz. The SSN had the spectral shape described by Moore et al. (2008). The prescribed IGs would be similar for a noise with the spectral shape used in the derivation of NAL-NL2. The shaping of frequency response
in the simulated aid was done using a single linear-phase FIR filter, applied before compression. The four higher compression channels were each 1-octave wide and were centered on frequencies of 1, 2, 4, and 8 kHz. The lowest channel included all frequencies up to 0.71 kHz. FIR filters were used to create the channel signals, and the delay introduced by these filters was removed. The frequency response of each filter overlapped with that of the neighboring filters at the –6 dB point, and overlapped with that of the next-but-one neighboring filters at the –65 dB point or lower. The compression thresholds were set to 49, 41, 40, 34, and 28 dB SPL for channels 1 to 5, respectively. These values were chosen to be representative of those chosen by default in the fitting software of some hearing aid manufacturers (neither CAM2 nor NAL-NL2 gives recommended compression thresholds). The compression ratio could be set independently for each channel; the values were set according to the relevant prescription method, either CAM2 or NAL-NL2.

The attack and release times (ANSI 2003) could be set to any desired value for each channel. To simulate fast compression, the attack/release times were set to 10/100 msec for all channels. To simulate slow compression, the attack/release times were set to 50/3000 msec for all channels. For CAM2, the prescribed gains do not depend on compression speed. For NAL-NL2 there are adjustments based on compression speed, but these are very small. For both compression speeds, we used the gains prescribed by the NAL-NL2 software when the option “Dual/Intermediate/Adaptive” was selected. The gains prescribed in this way differed by less than 1 dB from the gains prescribed when the options “Fast” or “Slow” were selected.

All processing was performed offline, using at least 24-bit precision. Stimuli were generated via an Echo Indigo 24-bit sound card, using a sampling rate of 44.1 kHz, and were presented via Sennheiser HDA200 headphones. These headphones are often used to measure audiometric thresholds for frequencies above 8 kHz. The response of the headphones was measured with a KEMAR (Burkhard & Sachs 1975), using the “large” pinnae. Digital filtering was used to “correct” the response of the headphones so that it corresponded closely to the diffuse-field response of the ear as specified in ANSI (2007). This ensured that the prescribed IGs were implemented accurately. The system as a whole had a frequency response extending up to 10 kHz.

It should be noted that both the prescribed IGs and the compression ratios could be set very accurately in the software. Also, the output of the headphones for frequencies up to 10 kHz is rather consistent across individual ears (Han & Poulsen 1998). Thus, the prescriptions were probably implemented more accurately here than would be the case in a clinical setting with real hearing aids.

Stimuli
Comparisons of sound quality were obtained for four types of music signals and male and female speech. The music signals were: a 7.3 sec segment of a jazz trio (piano, bass, and drums from Oscar Peterson’s We Get Requests); a 5.6 sec segment of an orchestra (including brass instruments and cymbals) performing Bizet’s Carmen; a 3.5 sec segment extracted from track 27 of the compact disc produced by Bang & Olufsen called Music for Archimedes (CD B&O 101) consisting of a xylophone playing the Sabre Dance by Khatchaturian (anechoic recording); and a male singing (a countertenor accompanied by guitar and recorder). For all four music signals, the input level to the simulated hearing aid was 50, 65, or 80 dB SPL. The level was always the same for the two sounds within a pair. The long-term average spectra of the first three signals are shown in Figure 1 of Moore et al. (2011).

The male or female speech was presented either in quiet or in a background of either SSN (with signal-to-noise ratio [SNR] = 0 or −3 dB) or a single background talker with opposite gender to the target talker (with signal-to-background ratio [SBR] = 0 dB or −3 dB). The speech stimuli were digitally recorded segments of running speech (connected discourse) obtained from six male and six female talkers of British English. The recordings were a subset of those described by Moore et al. (2008). Examples of the spectra of the speech stimuli are given in that article. The recordings were edited using CoolEdit to remove pauses for breath and extraneous sounds, but natural sounding gaps of between 100 and 300 msec were left. One 4.8 sec segment of speech was selected for each talker. The input level of the target speech to the simulated hearing aid was 50, 65, or 80 dB SPL.

Paired-Comparison Procedure
All testing was conducted with the participant seated in a double-walled sound-attenuating booth. The participant could see a computer screen and had a mouse that was linked to the screen. The experimenter was seated outside the booth. The same segment of sound was presented twice in succession, once with the gains and compression ratios (CRs) prescribed by NAL-NL2 and once with the gains and CRs prescribed by CAM2. The possible orders were used equally often and the order was randomized across trials. Within a given pair of sounds (one trial), the only difference between the signals was in the fitting method; the input level and compression speed were always the same. For each pair of sounds, the participant was asked to indicate which of the two was preferred and by how much. Each pair was only presented once. Participants responded using a slider on the screen, which could be moved, using the mouse, along a continuum labelled “1 much better,” “1 moderately better,” “1 slightly better,” “equal,” “2 slightly better,” “2 moderately better,” and “2 much better.” The participant moved the slider to the appropriate point to indicate his/her preference. Choices were not restricted to the labelled points; any point along the slider could be chosen. In separate blocks of trials, participants were asked to base their judgments on: (1) overall sound quality; (2) (for speech only) clarity/intelligibility. For the judgments of overall sound quality, the target speech signals were presented in quiet because it is easier to judge the quality of speech when it is presented in quiet. For the judgments of speech clarity/intelligibility, the target speech was presented with either an SSN or competing-talker background.

For overall sound quality judgments, the design was as follows. Within each block of trials, six types of signals were presented (classical music, jazz, male singing, percussion, female speech, and male speech). The two compression speeds were also used within a single block, but the input level was kept constant. Within a block, the 24 pairs of sounds (6 Signal Types × 2 Compression Speeds × 2 Presentation Orders [CAM2 first or NAL-NL2 first]) were presented in random order. Two blocks of trials were used for each participant and each input level. The order of presentation of input levels across blocks was random.
For judgments of the clarity of speech in a background of SSN, a block contained 48 pairs of sounds, and participants were given a short break after 24 pairs had been presented. Twelve different talkers (6 female and 6 male) and two compression speeds were used within a block, but the input level and SNR (0 or –3 dB) were kept constant. Within a block, the 48 pairs of sounds (12 Talkers × 2 Compression Speeds × 2 Presentation Orders) were presented in random order. A single block of trials was used for each participant for each input level and each SNR. Within that block, CAM2 and NAL-NL2 were compared 24 times (12 Talkers × 2 Presentation Orders) for each compression speed. The order of presentation of input levels and SNRs across blocks was random.

For judgments of the clarity of speech in a background of a single talker, the gender of the target talker was fixed within a block. The background talker was always of the opposite gender to the target talker, and the background talker on a given trial was selected randomly from six possible talkers. Both compression speeds were used within a block, but the input level and SBR (0 or –3 dB) were kept constant. In a given block, the 24 pairs of sounds (6 Target Talkers × 2 Compression Speeds × 2 Presentation Orders) were presented in random order. Two blocks of trials were used for each participant for each input level, SBR, and gender of the target talker. The order of presentation of input levels, SBRs, and talker gender across blocks was random.

Preference scores for each participant and each condition were computed in the following way. Regardless of the order of presentation in a given trial (CAM2 first or NAL-NL2 first), if CAM2 was preferred the slider position was coded as a negative number and, if NAL-NL2 was preferred the slider position was coded as a positive number. For example, if the order on a given trial was NAL-NL2 first and CAM2 second, and the participant set the slider position midway between “2 slightly better” and “2 moderately better,” the score for that trial was assigned a value of −1.5. The overall score for a given fitting method and stimulus type (e.g., classical music) was obtained by averaging all the subscores obtained for that fitting method and stimulus

### TABLE 1. Participant identifier, age, gender, the test ear, whether the participant had previous experience with hearing aids, and audiometric thresholds of the test ear (in dB HL)

<table>
<thead>
<tr>
<th>Participant Identifier</th>
<th>Age (Yrs)</th>
<th>Gender</th>
<th>Ear</th>
<th>Hearing Aid Experience</th>
<th>0.125</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>68</td>
<td>M</td>
<td>L</td>
<td>Y</td>
<td>20</td>
<td>20</td>
<td>35</td>
<td>35</td>
<td>40</td>
<td>40</td>
<td>45</td>
<td>60</td>
<td>65</td>
<td>65</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>P2</td>
<td>74</td>
<td>M</td>
<td>R</td>
<td>N</td>
<td>20</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>25</td>
<td>25</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>60</td>
<td>60</td>
<td>55</td>
</tr>
<tr>
<td>P3</td>
<td>67</td>
<td>M</td>
<td>L</td>
<td>N</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>14</td>
<td>10</td>
<td>16</td>
<td>28</td>
<td>38</td>
<td>46</td>
<td>46</td>
<td>58</td>
<td>62</td>
</tr>
<tr>
<td>P4</td>
<td>71</td>
<td>M</td>
<td>L</td>
<td>Y</td>
<td>15</td>
<td>15</td>
<td>25</td>
<td>30</td>
<td>25</td>
<td>40</td>
<td>45</td>
<td>45</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>P5</td>
<td>63</td>
<td>F</td>
<td>L</td>
<td>N</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>20</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>P6</td>
<td>73</td>
<td>M</td>
<td>R</td>
<td>Y</td>
<td>30</td>
<td>30</td>
<td>35</td>
<td>45</td>
<td>50</td>
<td>50</td>
<td>65</td>
<td>65</td>
<td>70</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>P7</td>
<td>69</td>
<td>F</td>
<td>R</td>
<td>N</td>
<td>20</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>35</td>
<td>40</td>
<td>40</td>
<td>50</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>P8</td>
<td>76</td>
<td>M</td>
<td>R</td>
<td>N</td>
<td>30</td>
<td>30</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>35</td>
<td>40</td>
<td>55</td>
<td>65</td>
<td>65</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>P9</td>
<td>66</td>
<td>M</td>
<td>L</td>
<td>Y</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>25</td>
<td>25</td>
<td>20</td>
<td>40</td>
<td>80</td>
<td>80</td>
<td>75</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>P10</td>
<td>68</td>
<td>F</td>
<td>R</td>
<td>N</td>
<td>20</td>
<td>15</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>10</td>
<td>25</td>
<td>35</td>
<td>45</td>
<td>60</td>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td>P11</td>
<td>72</td>
<td>M</td>
<td>L</td>
<td>Y</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>P12</td>
<td>64</td>
<td>M</td>
<td>R</td>
<td>N</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>30</td>
<td>50</td>
<td>50</td>
<td>45</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>P13</td>
<td>69</td>
<td>F</td>
<td>R</td>
<td>Y</td>
<td>15</td>
<td>15</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>55</td>
<td>55</td>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>P14</td>
<td>63</td>
<td>M</td>
<td>R</td>
<td>N</td>
<td>25</td>
<td>15</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>22</td>
<td>30</td>
<td>40</td>
<td>45</td>
<td>65</td>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td>P15</td>
<td>54</td>
<td>M</td>
<td>R</td>
<td>N</td>
<td>18</td>
<td>20</td>
<td>26</td>
<td>28</td>
<td>30</td>
<td>22</td>
<td>20</td>
<td>28</td>
<td>54</td>
<td>72</td>
<td>72</td>
<td>72</td>
</tr>
</tbody>
</table>

Y, yes; N, no.
RESULTS

Comparison of Insertion Gains for CAM2 and NAL-NL2

Figure 1 shows the IGs recommended by CAM2 and NAL-NL2 for an SSN input signal with levels of 50, 65, and 80 dB SPL. Symbols show mean values across participants and error bars indicate ± 1 SE. Symbols for the two methods are slightly offset horizontally to avoid overlap. The IGs for input levels of 65 and 50 dB SPL are shifted upward by 10 and 20 dB, respectively, to avoid overlap of the data points. IGs, insertion gains.

Increasing frequency above 4 kHz, whereas the IGs for NAL-NL2 remained roughly constant. The CRs for the two fitting methods were similar. For example, at 2 kHz, the mean IG for CAM2 decreased from 22 to 9 dB as the input level increased from 50 to 80 dB SPL, corresponding to a CR of 1.8. The IGs for NAL-NL2 decreased from 19.4 to 5.2 dB as the input level increased from 50 to 80 dB SPL, corresponding to a CR of 1.9.

Relative Preferences for Sound Quality

To give an overall view of the quality preferences and of individual variability in preferences, the preference scores for sound quality were averaged across all sounds (music and speech) and across all input levels. The results for each participant and the group mean are shown in Figure 2. The standard error of the mean preference score for individual participants was typically approximately 0.1 scale unit. In what follows, scores within the range −0.2 to +0.2 are taken as indicating no clear preference. Preference scores were below −0.2, suggesting a preference for CAM2 relative to NAL-NL2, for 9 of the 15 participants. The others showed no clear preference. The mean score of −0.31 was significantly different from zero, based on a t test (t[2159] = −13.95, p < 0.001). The effect size, specified as Cohen’s d, was 0.30, indicating what is conventionally described as a “small” effect. A within-subjects analysis of variance (ANOVA) on the preference scores based on overall sound quality was conducted with factors: level, compression speed, and signal type (four types of music plus male and female speech). There were no significant main effects or interactions. Thus the general preference for CAM2 relative to NAL-NL2 did not vary significantly across level, compression speed, or type of signal.

Figure 3 shows preference scores averaged across participants for the overall quality of speech. The results are shown separately for the two compression speeds, female and male target talkers, and three input levels. The mean scores were negative, indicating a preference for CAM2 relative to NAL-NL2, for both compression speeds, both talker genders, and all three levels, although the difference was small for the female talkers and fast compression at the 80 dB SPL input level. The mean score of −0.36 was significantly different from zero, based on a t test (t[719] = −9.21, p < 0.001). Cohen’s was 0.34,
indicating a “small” effect. A within-subjects ANOVA on the preference scores for speech, with factors level, compression speed, and talker gender gave no significant main effects or interactions.

Figure 4 shows preference scores averaged across participants for the overall quality of music; results were similar for all four music signals, so they have been averaged across signal type. The results are shown separately for the two compression speeds and three input levels. CAM2 was preferred relative to NAL-NL2 for both compression speeds, and all three levels, although the effect was very small for the input level of 80 dB SPL. The mean score of −0.29 was significantly different from zero, based on a t test (t[1439] = −10.6, p < 0.001). Cohen’s was 0.28, indicating a “small” effect. A within-subjects ANOVA on the preference ratings for music, with factors level, compression speed, and type of music (four types) gave no significant main effects or interactions.

As shown in Figure 2, the relative preference for CAM2 or NAL-NL2 varied across participants. To assess whether the individual differences were related to the severity of the hearing loss, the participants were divided into three groups, each with five participants, based on the severity of their hearing loss averaged over audiometric frequencies from 3 to 10 kHz. For the group designated “Mild,” the mean hearing loss over this range was less than 55 dB. For group “Moderate,” the mean hearing loss was between 55 and 65 dB. For group “Severe,” the mean hearing loss was greater than 65 dB.

An ANOVA was conducted on the preference scores with group (Mild, Moderate, or Severe) as a between-subject factor and compression speed, input level, and type of signal as within-subject factors. There was no significant effect of group and no significant interaction of group with any other factor. Thus, for overall quality, the relative preference for CAM2 over NAL-NL2 does not seem to be related to the degree of hearing loss at high frequencies.

Relative Preferences for the Clarity of Speech in Noise

To give an overall view of the preferences for the clarity of speech in noise and of individual variability in preferences, the preference scores were averaged across all target talkers and across all input levels. The results for the SNR of 0 dB for each participant and the group mean are shown in Figure 5;
results were similar for the SNR of –3 dB (not shown). Five of 15 participants (P1, P5, P7, P13, and P15) preferred CAM2 relative to NAL-NL2, one (P4) showed the opposite pattern, and the remainder showed no clear preference. The mean score of –0.173 was significantly different from zero, based on a t test (t4319 = −13.88, p < 0.001). Cohen’s was 0.21, indicating a “small” effect. A within-subject ANOVA was conducted with factors level, compression speed, and SNR. There were no significant main effects or interactions, indicating that the preference for CAM2 relative to NAL-NL2 was similar for all levels, both compression speeds, and both SNRs.

As shown in Figure 5, the relative preference for CAM2 or NAL-NL2 varied across participants. To assess whether the individual differences were related to the severity of the hearing loss, the participants were divided into three groups, as before. An ANOVA was conducted with group (Mild, Moderate, or Severe) as a between-subject factor, and compression speed, input level, SNR, and talker gender as within-subject factors. There were no significant main effects or interactions. Thus, for the clarity of speech in noise, the relative preference for CAM2 over NAL-NL2 does not seem to be related to the degree of hearing loss at high frequencies.

Relative Preferences for the Clarity of Speech in a Background Talker

Because there were some differences in results for the female and male target talkers, the results were analyzed separately for the two talker genders. To give an overall view of the preferences for speech clarity and of individual variability in preferences, the preference scores were averaged across input levels and compression speed. The results for the female talkers for the SBR of 0 dB for each participant and the group mean are shown in Figure 6; results were similar for the SBR of –3 dB (not shown). Seven of 15 participants preferred CAM2 relative to NAL-NL2, two (P9 and P14) showed the opposite pattern, and the remainder showed no clear difference. The mean preference score of –0.10 was significantly different from zero, (t4319 = −8.02, p < 0.001), but Cohen’s d was only 0.12, indicating a very small effect. A within-subject ANOVA was conducted on the data for the female talkers with factors level, compression speed, SBR, and talker identity. There was a significant effect of talker identity (F[5, 70] = 7.73, p < 0.001), indicating that the strength of preference varied across talkers. No other main effects were significant.

To assess whether the individual differences were related to the severity of the hearing loss, the participants were divided...
into three groups, as before. An ANOVA was conducted with group (Mild, Moderate or Severe) as a between-subject factor and compression speed, input level, SBR and talker identity as within-subject factors. There was no significant main effect of group. The only significant two-way interaction involving group was the interaction with talker identity: \( F(5, 60) = 9.0, p < 0.001 \). Although mean scores for all talkers were negative (indicating a preference for CAM2) for the Mild and Moderate groups, mean scores were positive (indicating a preference for NAL-NL2) for five of the six talkers for the Severe group.

Individual results for the male target talkers are shown in Figure 7. Eight of 15 participants preferred CAM2 relative to NAL-NL2, 1 (P9) showed the opposite pattern and 6 (P2, P4, P6, P11, P12, and P14) showed no clear preference. The pattern of individual preferences was similar to that for the female target talkers. The mean preference score of −0.10 was significantly different from zero (\( t[4319] = −7.38, p < 0.001 \)), but Cohen’s was only 0.11, indicating a very small effect. A within-subject ANOVA was conducted on the data for the male talkers with factors level, compression speed, SBR, and talker identity. There was a significant effect of talker identity (\( F(5, 70) = 3.65, p = 0.005 \)), indicating that the strength of preference varied across talkers. No other main effects were significant.

To assess whether the individual differences were related to the severity of the hearing loss, the participants were divided into three groups, as before. An ANOVA was conducted with group (Mild, Moderate or Severe) as a between-subject factor and compression speed, input level, SBR and talker identity as within-subject factors. There was no significant main effect of group. The only significant two-way interaction involving group was the interaction with talker identity: \( F(5, 60) = 3.63, p = 0.006 \). Although mean scores for all talkers were negative (indicating a preference for CAM2) for the Mild and Moderate groups, mean scores were positive (indicating a preference for NAL-NL2) for all six talkers for the Severe group.

**DISCUSSION**

**Magnitude of the Preference Differences**

The magnitude of the preference scores was generally small, typically being less than one scale unit on a scale where a score of −3 or +3 would indicate a perfectly consistent and strong preference for one fitting over the other. The small preference scores may reflect three factors:

1. Variability in the responses. Participants were not completely consistent in their judgments. Because the maximum absolute value of the score on a single trial was 3, any variability leads to a mean score above −3 and below 3.
2. Participants are usually reluctant to use the extremes of a rating scale (Poulton 1979; Moore & Tan 2003). Hence, scores of −3 or 3 were very rare.
3. Preferences may have been relatively weak, because the perceptual differences between the fittings were small.

Given that the measured preferences were small, it might be thought that they are not relevant clinically. However, the direction of the effects for overall sound quality was consistent across the great majority of participants, and it occurred consistently over a range of sound levels. Therefore, we are of the opinion that the differences in preference between CAM2 and NAL-NL2 are meaningful, even though they are small.

**Differences in Audibility for CAM2 and NAL-NL2**

To assess the extent to which the audibility of the speech stimuli differed when they were amplified using NAL-NL2 and CAM2, we calculated audibility in different frequency bands for speech at 50 and 65 dB SPL. The calculations were based on the average audiogram across the 15 participants. The audiometric thresholds in dB HL were converted to the minimum audible pressure (MAP) for monaural listening in dB SPL, by...
adding the mean MAP values for normal hearing predicted by the loudness model of Moore and colleagues (Moore et al. 1997; Glasberg & Moore 2006; Moore & Glasberg 2007). The insertion gains prescribed for 50- and 65 dB SPL speech were averaged across participants for each of the two fitting procedures. Levels were calculated in frequency bands that were 2 ERB<sub>N</sub> wide, to allow for the fact that auditory filters are usually somewhat broader for hearing-impaired than for normal-hearing participants (Glasberg & Moore 1986). The speech spectrum was taken as the average for male and female talkers, as published by Moore et al. (2008). Figure 8 shows the results for the CAM2 procedure for speech at 65 dB SPL (top) and 50 dB SPL (bottom). All levels are expressed as sound pressure level at the eardrum. The solid line shows the MAP for normal hearing. The dashed line with asterisks shows the MAP corresponding to the average audiogram of the participants in this study. The dotted line without symbols shows the root-mean-square (RMS) level in 2-ERB<sub>N</sub>-wide bands after application of the insertion gains prescribed by CAM2. For the 65 dB SPL speech, the aided RMS level remains above the absolute threshold for frequencies up to 10 kHz, although at 10 kHz the RMS level is equal to the absolute threshold, meaning that only approximately half of the dynamic range of the speech would be audible; the “effective” dynamic range is usually assumed to extend from 15 dB below to 15 dB above the RMS level (ANSI 1997). The “full” dynamic range of the speech would have been audible from 0.2 to 4.2 kHz. For the 50 dB SPL speech, the RMS level approaches the absolute threshold at approximately 6.5 kHz. The full dynamic range of the speech would have been audible from approximately 0.27 to 3 kHz.

Figure 9 shows corresponding results for NAL-NL2. For the 65 dB SPL speech, the aided RMS level approaches the absolute threshold at 6 kHz. The full dynamic range of the speech would have been audible from approximately 0.2 to 3.8 kHz. For the 50 dB SPL speech, the RMS level approaches the absolute threshold at approximately 4.5 kHz. The full dynamic range of the speech would have been audible from approximately 0.27 to 1.8 kHz.

It is clear that for speech with medium and low levels, CAM2 leads to markedly higher audibility at medium and high frequencies than NAL-NL2 does, consistent with the analysis of Johnson and Dillon (2011).

### Insertion Gains

The differences between the IGs prescribed by CAM2 and by NAL-NL2 were small (usually < 5 dB) for frequencies up to 4 kHz, but CAM2 prescribed considerably more IG than NAL-NL2 for frequencies above 4 kHz. For the mean across participants, the IG prescribed by CAM2 at 10 kHz was close to 50 dB for an SSN input with a level of 50 dB SPL. Such high IGs are difficult to achieve in practice with current hearing aids, even with “closed” fittings. The situation would be made worse with “open” fittings, which would have been appropriate for the majority of participants tested here. For such fittings, acoustic feedback would have limited the IG that could be applied at high frequencies for both CAM2 and NAL-NL2. In our evaluation of CAM2 using wearable experimental aids (Moore & Füllgrabe 2010), we were able to achieve the gains prescribed by CAM2 for frequencies up to 7.5 kHz only by using closed earmolds. However, our recent work has shown that, although there were benefits for speech intelligibility of providing amplification of frequency components between 5 and 7.5 kHz, there was no additional benefit of amplifying frequency components between 7.5 and 10 kHz (Füllgrabe et al. 2010; Moore et al. 2010a, 2011). Thus, it may be sufficient to apply the IGs prescribed by CAM2 only for frequencies up to 7.5 or 8 kHz. This would probably have only small effects on sound quality (Moore & Füllgrabe 2010) but would reduce the maximum IG required, because the gain prescribed by CAM2 usually increases progressively with increasing frequency, as can be seen in Figure 1.

The simulated hearing aid used here had relatively low compression thresholds, which meant that the IG increased with...
decreasing sound level down to quite low levels. It may be the case that, in practice, the benefits for speech intelligibility of amplification of frequency components above 5 kHz occur mainly for speech in background sounds, because for speech in quiet intelligibility is usually dominated by frequency components below 5 kHz (ANSI 1997). Typically, when speech is heard in noise the overall sound level is relatively high (Pearsons et al. 1976). Hence, it may not be necessary to use low compression thresholds for channels in the hearing aid centered above 4 kHz. The use of higher compression thresholds would reduce the maximum IG required. For example, if the compression threshold at 8 kHz was set so that it was just reached for an SSN input with a level of 65 dB, the maximum prescribed IG at 8 kHz would be close to 30 dB. This is not greatly in excess of what can be achieved with some modern hearing aids, based on our own measurements with such aids.

**Individual Differences in Preferences for Sound Quality**

Averaged across all sounds used, 9 of the 15 participants preferred CAM2 relative to NAL-NL2, and the remainder showed no clear preference; see Figure 2. Of the nine participants who preferred CAM2, three had previous hearing aid experience, whereas six did not. Of the six who showed no clear preferences, three had previous hearing aid experience whereas three did not. Thus, previous experience with hearing aids does not seem to be the critical factor in determining preferences. When the ratings were analyzed with the participants divided into groups (Mild, Moderate, or Severe) based on the amount of high-frequency hearing loss, the interaction of fitting method with group was not significant. Participant P4, who was an experienced hearing aid user with a mild/moderate high-frequency loss, preferred CAM2 for overall quality but preferred NAL-NL2 for the clarity of speech in noise. The origin of the individual differences in preferences remains unclear.

**Effect of Fitting Method and Individual Differences for the Clarity of Speech in a Background Talker**

For the clarity of speech in a background talker, there was a significant effect of fitting method for both talker genders, but the effect was very small. The small effect may have occurred because the intelligibility of speech in a background talker is mainly determined by informational masking rather than by energetic masking (Brungart et al. 2001). The major problem for the listener is in deciding which components of the input signal emanated from the target talker and which components emanated from the background talker. When informational masking is dominant, changes in the audibility of high-frequency components in the speech may have little influence on intelligibility because segregation of the mid frequency components of the target and masker rather than audibility is the limiting factor. Informational masking can be partially or completely overcome by spatial separation of the target and background talkers (Freyman et al. 2001; Moore 2012), and under conditions of spatial separation the audibility of high-frequency components in the speech can affect speech intelligibility (Best et al. 2005; Moore et al. 2010a). Therefore, we anticipate that effects of the fitting method (CAM2 versus NAL-NL2) might become larger for the clarity of speech in a background talker under conditions in which the target and background are spatially separated and the participant is bilaterally aided. That remains to be determined.

Seven participants gave higher preference ratings for CAM2 than for NAL-NL2 for both target talker genders, whereas only one (P9) preferred NAL-NL2 for both talker genders. ANOVAs based on the data obtained when participants were split into groups according to the severity of their high-frequency hearing loss showed a significant interaction between group and talker identity for both the female and male target talkers. For the Mild and Moderate groups, CAM2 was preferred overall for all talkers, whereas for the Severe group, NAL-NL2 was mostly preferred. The results suggest that severe high-frequency hearing loss is sometimes associated with a preference for somewhat lower IGs at high frequencies than prescribed by CAM2. However, P13 had a mean high-frequency hearing loss of 80 dB but consistently preferred CAM2 over NAL-NL2, so there is clearly individual variability in the preferences that is not accounted for by high-frequency hearing loss.

**Effects of Input Sound Level and Compression Speed**

It is possible that the general trend for CAM2 to be preferred over NAL-NL2 was related to the greater loudness that would have been produced by the use of CAM2 gains because CAM2 prescribed slightly greater gains than NAL-NL2 for medium frequencies and markedly greater gains for high frequencies. If the judgments were influenced by loudness, then one would expect the preference for CAM2 over NAL-NL2 to be reduced or even reversed for the 80 dB SPL input level because lower loudness would typically be preferred for stimuli with such an input level (Smeds 2004; Smeds et al. 2006). In fact, for all types of preferences judgments—overall quality, clarity of
speech in noise, and clarity of speech in a background talker—there was no significant interaction of the fitting method with input sound level. Thus, the data do not support the idea that relative preferences for CAM2 or NAL-NL2 gains were strongly influenced by loudness. The fact that the relative preferences were similar for all three levels suggests that the relative preferences were mainly based on the different shapes of the frequency-gain characteristics (CAM2 prescribing more gain above 4kHz) rather than on loudness.

It should be noted that, within a pair of sounds to be compared, the compression speed was always the same. We did not directly compare preferences for slow and fast compression. In a previous study in which such a comparison was made (Moore et al. 2011), there was no significant effect of compression speed for any stimulus (music or speech) when the input level was 50 dB SPL. For an input level of 65 dB SPL, there was a significant preference for slow compression for classical music, but not for the other sounds used. For an input level of 80 dB SPL, slow compression was preferred over fast compression for classical music, jazz music, and female speech, but not male speech. The strength and consistency of the preferences varied across participants.

**Generality of the Results**

The participants tested here all had the common pattern of no or little hearing loss at low frequencies with increasing loss at high frequencies. Further studies are clearly needed to compare CAM2 and NAL-NL2 for other types of hearing losses, such as flat losses, reverse-slope losses, or mid-frequency losses, and also for greater degrees of sloping hearing loss.

During testing, there was no opportunity for participants to adjust the gain to achieve a preferred loudness. This simulates the situation of a hearing aid that has no volume control, or for which the volume control is disabled. This is consistent with one of the goals of both CAM2 and NAL-NL2, that the fitting should lead to satisfactory loudness for a large range of input sound levels, without any need for manual adjustments. Of course, many hearing aids do have a volume control, and relative preferences for CAM2 or NAL-NL2 might be affected by the availability of a volume control. This is a topic for future research.

The simulated hearing aid used here had five compression channels. As noted earlier, the choice of number of channels reflects a compromise between having sufficient flexibility to implement the frequency-dependent gains and compression ratios required to match the prescription while avoiding the reduction of spectral and temporal contrast that can occur when many channels are used (Plomp 1988; Bor et al. 2008; Stone & Moore 2008). Accuracy in achieving the target frequency-dependent gains in the simulated hearing aid was greatly increased by our use of a digital filter before compression processing. There is no general consensus as to the “optimum” number of compression channels in a hearing aid, if indeed there is an optimum. It is possible that the pattern of preferences found here would change if simulated or real aids with a different number of channels were used, but this seems unlikely given previous research showing only small effects of the number of channels used (Yund & Buckles 1995; Keidser & Grant 2001).

The present study simulated unilateral aiding. It is possible that the pattern of preferences might change for bilateral fittings. However, this seems unlikely, given that preferred gains and frequency responses do not seem to differ markedly for unilateral and bilateral fittings (Alcántara et al. 2004; Mariage et al. 2004).

The simulated hearing aid used here allowed prescribed gains and CRs to be programmed very accurately, and also allowed rapid switching between the settings for CAM2 and NAL-NL2, thus facilitating paired comparisons. However, participants were not given the opportunity to acclimatize to any specific setting. Further studies using field trials with wearable aids are desirable to assess whether the relative preference for CAM2 over NAL-NL2 found here would also occur in everyday life and whether it would be affected by acclimatization.

The presentation using headphones used here bypasses a common problem when using open-fit hearing aids. With such aids, sound at low and medium frequencies is heard via a direct path from sound source to the eardrum and via the hearing aid.

The aid introduces a time delay that leads to comb-filtering effects and irregularity in the effective frequency response at the eardrum (Stone et al. 2008). The exact pattern of comb filtering is level dependent because the relative level of the direct sound and delayed sound through the aid is itself level dependent as a result of the compression processing in the hearing aid. This makes it difficult, if not impossible, to achieve the target IGs over a wide range of levels. For real hearing aids, as opposed to the simulated aid used here, the failure to match targets may be as important as the differences between methods used for initial fitting.

**SUMMARY AND CONCLUSIONS**

We obtained relative preference judgments for sounds processed using a simulated unilaterally fitted five-channel compression hearing aid with gains and compression ratios selected according to the recommendations of CAM2 or NAL-NL2. Within a given trial, the same segment of sound was played twice, once with CAM2 settings and once with NAL-NL2 settings, in random order. Participants had to indicate which one was preferred and by how much.

Judgments of overall sound quality were obtained for female and male speech in quiet, and for four types of music (classical, jazz, a man singing, and percussion). Judgments of speech clarity were obtained for female and male speech in speech-shaped noise, female speech in a male-talker background, and male speech in a female-talker background. Factors investigated included the compression speed in the simulated hearing aid (slow or fast) and the input sound level (50, 65, or 80 dB SPL).

For judgments of overall sound quality, 9 of the 15 participants preferred CAM2 relative to NAL-NL2, and the remainder showed no clear preference. There was a significant overall preference for CAM2. The preference for CAM2 over NAL-NL2 was present for all types of stimuli, both compression speeds, and all three levels.

For judgments of the clarity of speech in noise, 5 of the 15 participants preferred CAM2 relative to NAL-NL2, 1 showed the opposite preference, and the remainder showed no clear preference. There was a significant overall preference for CAM2. The preference for CAM2 over NAL-NL2 was present for all types of stimuli, both compression speeds, and all three levels.

For judgments of the clarity of speech in a background talker there was a significant overall preference for CAM2 relative to NAL-NL2, but the effect was very small.
ACKNOWLEDGMENTS

The authors thank Michael Stone for writing the software for the hearing aid simulator and for assistance with calculations of audibility; Christian Füllgrabe for help with the preference-rating software and helpful comments on an earlier version of this article; and Brian Glaserbarg for help with statistical analyses. The authors also thank Ben Hornsby for very helpful comments on earlier versions of this article.

This work was supported by Deafness Research UK and by the Medical Research Council (United Kingdom) grant number G0701870.

The authors have no conflicts of interest to disclose.

Address for correspondence: Brian C. J. Moore, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England. E-mail: bcjm@cam.ac.uk

Received January 9, 2012; accepted June 4, 2012.

REFERENCES


Keidser, G., & Grant, F. (2001). The preferred number of channels (one, two, or four) in NAL-NL1 prescribed wide dynamic range compression (WDRC) devices. Ear Hear, 22, 516–527.


