Direct-location versus verbal report methods for measuring auditory distance perception in the far field

Pablo E. Etchemendy[1] · Ignacio Spiousas[2,3] · Esteban R. Calcagno[1] ·
Ezequiel Abregú[1] · Manuel C. Eguia[1] · Ramiro O. Vergara[1]

Abstract In this study we evaluated whether a method of direct location is an appropriate response method for measuring auditory distance perception of far-field sound sources. We designed an experimental set-up that allows participants to indicate the distance at which they perceive the sound source by moving a visual marker. We termed this method Cross-Modal Direct Location (CMDL) since the response procedure involves the visual modality while the stimulus is presented through the auditory modality. Three experiments were conducted with sound sources located from 1 to 6 m. The first one compared the perceived distances obtained using either the CMDL device or verbal report (VR), which is the response method more frequently used for reporting auditory distance in the far field, and found differences on response compression and bias. In Experiment 2, participants reported visual distance estimates to the visual marker that were found highly accurate. Then, we asked the same group of participants to report VR estimates of auditory distance and found that the spatial visual information, obtained from the previous task, did not influence their reports. Finally, Experiment 3 compared the same responses that Experiment 1 but interleaving the methods, showing a weak, but complex, mutual influence. However, the estimates obtained with each method remained statistically different. Our results show that the auditory distance psychophysical functions obtained with the CMDL method are less susceptible to previously reported underestimation for distances over 2 m.

Keywords Auditory perception · Distance perception · Psychoacoustics · Response methods · Cross-modal information · Direct location · Verbal report

Pablo E. Etchemendy and Ignacio Spiousas contributed equally to this work.

✉ Ramiro O. Vergara
ramirovergara@lapso.org

[1] Laboratorio de Acústica y Percepción Sonora, Escuela Universitaria de Artes, CONICET, Universidad Nacional de Quilmes, Roque Sáenz Peña 352, Bernal, Buenos Aires B1876BXD, Argentina

[2] Douglas Mental Health University Institute and Department of Psychiatry, McGill University, Montreal, Canada

[3] BRAMS Laboratory, Centre for Research on Brain, Language and Music (CRBLM), Montreal, Canada

Introduction

Auditory distance perception (ADP) has been studied since the early 20th century (Angell & Fite, 1901; Gamble, 1909; Starch & Crawford, 1909) using different experimental methodologies, acoustic environments (both real and virtual) and stimuli. Today we know that ADP relies on the integration of a number of different cues: sound level, direct-to-reverberant energy ratio (DRR), spectral content, binaural differences and previous knowledge of the sound source, among others (see Fluitt, Mermagen, & Letowski, 2014; Kolarik, Moore, Zahorik, Cirstea, & Pardhan, 2016a; Zahorik, Brungart, & Bronkhorst, 2005, for a complete review).

In most ADP studies, participant's distance reports display large biases (i.e., departures from the true source location) and high variability (i.e., inconsistency between responses in successive trials). Typically the reported distances are overestimated for sources located closer than 2 m, while they are substantially and progressively underestimated for greater distances (Bronkhorst & Houtgast, 1999; Fontana & Rocchesso, 2008; Kearney, Gorzel, Rice, & Boland, 2012; Loomis, Klatzky, Philbeck, & Golledge, 1998; Parseihian,

Jouffrais, & Katz, 2014; Zahorik, 2001, 2002). The compressive non-linear relationship between source location and response is reflected in an exponent smaller than one when the distance estimates are fitted by a power-law function (mean = 0.54, reported by Zahorik et al., 2005).

One consequence of this compressive behavior is that, in reverberant environments, perceived distances can be discriminated up to a certain threshold, which gives the maximum limit for perceptual judgments of distance (the so-called "auditory horizon", see Bronkhorst & Houtgast, 1999; Zahorik et al., 2005). One of the hypotheses posed by the literature is that this compression has a perceptual origin (Zahorik et al., 2005) rather than one related to the response method. A candidate explanation is based on the flattening in the decay of the DRR with the source distance (Larsen, Iyer, Lansing, & Feng, 2008). This effect takes place when the energy of the direct sound is small compared with that of the reverberant sound (large negative values of the DRR) and yields to compressive distance perception beyond the auditory horizon. In support of this idea, Loomis et al. (1998) showed this compressive nature using two different response methods: verbal report (VR), in which the subject has to indicate verbally the response using explicit distance scales, such as meters or feets; and direct-location (DL), in which the subject performs an action to indicate the perceived distance, such as walking with eyes covered, pointing, throwing, etc. Finally, experiments of visual distance perception (VDP) have shown that perceived distance responses collected using similar methods showed lower bias and variability than that observed in ADP experiments, supporting the idea that the origin of this effect is due to perceptual factors specific to the auditory domain (Loomis et al., 1998; Loomis, Philbeck, & Zahorik, 2002).

However, a number of recent studies suggest that the response method influences the response obtained in both ADP and VDP. For example, Andre and Rogers (2006) compared VR and blind-walking estimates for VDP, and showed that the latter are consistently more accurate than VRs. In addition, Ashmead et al. (1995) used blind-walking estimates to measure ADP curves in the far field (distance > 1 m) and obtained more accurate responses than those normally reported in other studies using VR. Finally, Brungart et al. (2000) performed both VDP and ADP experiments in the near field (distance < 1 m) comparing VR with three other direct- and indirect-location methods. The DL method consisted of reporting the perceived distance by placing an electromagnetically tracked sensor, mounted on the tip of a wand, at the perceived location of the target. The authors found the DL method to be superior (showing smallest bias and variability) than VR in both modalities. They proposed that the DL method appears to be a natural response, since no mental transformation of the target location is required and subjects can use their own anatomical reference points to determine the target's location.

Unlike the near field, where a variety of DL methods were used, the most used method in far-field ADP experiments is VR (Calcagno, Abregú, Eguía, & Vergara, 2012; Kolarik, Pardhan, Cirstea, & Moore, 2013; Kolarik et al., 2016a; Spiousas, Etchemendy, Eguía, Calcagno, Abregú, & Vergara, 2017; Zahorik, 2001, 2002). The lack of use of DL methods to measure ADP in the far field could be motivated by the fact that the distances to be estimated are beyond hand reach and, therefore, the procedure to measure the estimated distance becomes more complicated and slower to implement and perform, compared to the near field. Among DL methods, a widely used method to measure distance perception in the far field is blind-walking, in which participants view or hear a target, then cover their eyes and attempt to walk without vision to the remembered target location (Ashmead et al., 1995; Creem-Regehr, Willemsen, Gooch, & Thompson, 2005; Loomis et al., 1998; Philbeck, Loomis, & Beall, 1997; Rieser, Ashmead, Talor, & Youngquist, 1990; Thomson, 1983; Wu, Ooi, & He, 2004). Although performance measures with blind-walking show small average systematic errors, this method is subject to certain limitations that make it difficult to use. First, it requires space. In order to report the distance to a real target in the far field, it is necessary to have an experimental room of great dimensions, or an open space in which to carry out the experiments, which can introduce other logistical challenges, especially in ADP experiments (lack of reverberation, excessive background noise, lack of electrical infrastructure for plugging audio devices, among others). Second, each response requires the completion of several steps. The participant has to get up from the chair, walk to the estimated distance, and then return to the origin with the help of the experimenter. Moreover, in each trial the experimenter must move the target away from its position to prevent the listener from colliding with it while responding. Finally, the experimenter has to measure the distance traveled by the participants which, in experiments performed in the dark, is a cumbersome process. These drawbacks are maintained even when virtual sources are used since, although virtual sources can be located at different distances from the listener regardless of the actual experimental room, the subject requires enough real space to report the source location, and the response procedure remains the same.

Many of these practical limitations may be avoided by asking participants to provide VRs of perceived distance. VR reduces time consumption since the task is relatively simple and, under certain conditions (targets located at distances < ~3 m and up to ~5 m in the presence of multiple visual cues), is a precise method for estimating distances (Calcagno et al., 2012; Philbeck & Loomis, 1997). VR also significantly reduces the resource requirements since it can be implemented, in its simplest version, with only a loudspeaker and an audio player. Another great advantage of VR compared to DL methods is that the response is not limited by the boundaries

of the environment where the task is performed, avoiding biases or distortions in reported judgments introduced by floor and ceiling effects. This last aspect makes VR a well-suited alternative for both VDP and ADP experiments carried out in virtual environments, as it allows testing virtual sources located at large distances without the need for a large real space to report the perceived distance. However, VR has some disadvantages with respect to DL methods. First, as we mentioned previously, VDP and ADP experiments showed less accurate responses with VRs (Andre & Rogers, 2006; Brungart et al., 2000; Loomis, Da Silva, Fujita, & Fukusima, 1992; Thomson, 1983), especially for distances greater than ~3 m, for which VRs are typically underestimated (Anderson & Zahorik 2014; Andre & Rogers, 2006; Cutting & Vishton, 1995; Kelly, Loomis, & Beall, 2004; Loomis et al., 1998; Toye, 1986; Zahorik, 2001, 2002). Second, VRs were shown to be more affected by the environmental context (Andre & Rogers, 2006; Iosa, Fusco, Morone, & Paolucci, 2012). For example, a comparative study by Andre and Rogers (2006) showed that blind-walking estimates of visual egocentric distance are consistently more accurate than VRs, and that the type of environment (indoor vs. outdoor) selectively influences VRs but not blind-walking. Finally, VRs require number manipulation, which may render the method unfit for the study of perception in certain populations (e.g., children).

In this work, we tested a DL method that bypasses many of the above-described disadvantages of its kind. In the proposed method, participants use a hand-held control to move a visual marker (made of two green LEDs) along a line parallel to the line joining the listener and the sound source, in order to report the perceived distance (Fig. 1). We termed this method Cross-Modal Direct Location (CMDL) since the response procedure involves the visual modality while the stimulus is presented through the auditory modality. The CMDL method is very similar to that used in a previous ADP study carried out by Fontana and Rocchesso (2008). However, we introduced several changes that in our view facilitate and quicken the task for both the participant and the experimenter: the visual marker is illuminated, it is moved by an electric motor, and the data are automatically collected and stored by a personal computer, allowing us to carry out experiments in complete darkness (see General methods). In contrast, in Fontana and Rocchesso (2008) the response marker (a blue napkin) had to be moved manually by means of a pulley system and the reported distance was measured by the experimenters using a measuring tape; because of these reasons, experiments had to be carried out in a lit environment. Unfortunately, Fontana and Rocchesso (2008) studied the effect of exaggerating the acoustical cue of reverberation while keeping the intensity cue practically constant with source distance, making it difficult to compare their results with previous results obtained by other methods.

As shown above, several VDP studies have shown that verbal estimates tend to underestimate distances greater than ~3 m (Cutting & Vishton, 1995; Kelly et al., 2004; Loomis et al., 1998; Toye, 1986), yet DL methods seem to be more accurate (Thomson, 1983; Loomis et al., 1992). Given this background, we expect that using CMDL to measure ADP will decrease the response biases (underestimation for sources farther than ~3 m) typically obtained using VRs. In order to test this hypothesis, we conducted a series of experiments comparing the ADP response obtained using the classical VR versus the proposed CMDL method under natural listening conditions.

General methods

Testing environment

All experiments were performed in a semi-reverberant room (length: 12 m; width: 7 m; height: 3 m) with walls covered by sound-absorbing panels (pyramid polyurethane acoustic foam, 50 mm), the floor by carpet, and the ceiling by fiberglass acoustic panels. The participant was comfortably seated at 2 m distance from one of the short walls, and slightly offset from the central line of the room in the perpendicular direction. His/her ears were at a height of approximately 1.2 m. The average reverberation time of the room (T_{30}, A-weighting measured with the MLS method) was 0.45 s at the participant's position. The background noise of the room at that position was 19 dBA (measured with a RION NL-32 sound level meter).

Participants

A total of 32 volunteers (26 men, age range 18–35 years; mean age = 28.8 years) participated in the study. Sixteen subjects participated in Experiment 1, eight in Experiment 2, and the remaining eight in Experiment 3. None of them took part in more than one experiment. The majority of the volunteers (~75%) were undergraduate or graduate music students at the Universidad Nacional de Quilmes. All participants reported normal or corrected-to-normal vision and normal hearing. None of the participants had prior knowledge of the testing room nor its dimensions.

Experimental set-up, auditory and visual stimuli

The experimental set-up is shown in Fig. 1. The sound source (Genelec 8020B bi-amplified 50W, Fig. 1A) was located in front of the participant, 1.2 m above floor level. The source was free to move suspended along a 6-m long metal rail. The displacement of the source was done manually by one experimenter, allowing the stimulus to be played at different distances from the
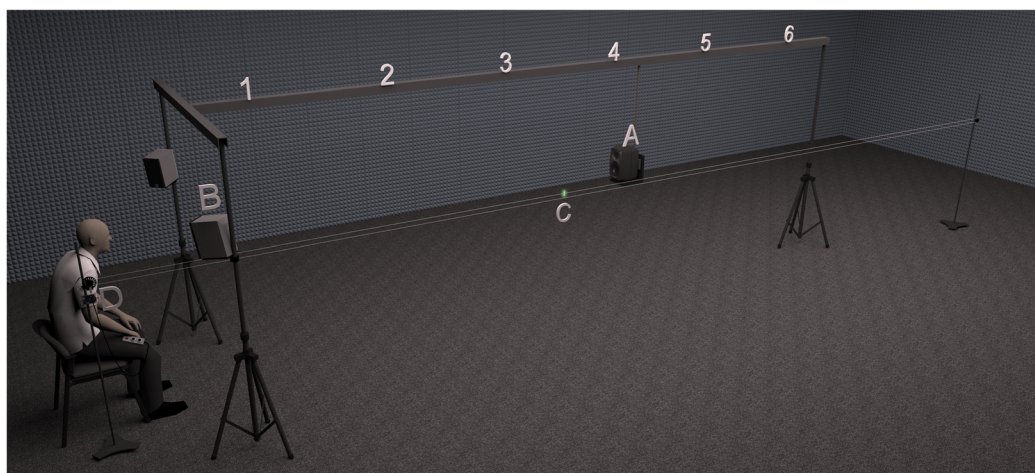
Fig. 1 Three-dimensional model of the experimental set-up. (**A**) Mobile speaker suspended from a metal rail, (**B**) masking system, (**C**) visual mobile marker formed by a pair of green LEDs (standard, 3 mm) located vertically 4 cm apart from each other, remotely controlled by the subject, and (**D**) optical encoder for the mobile marker position. The set of source distances are indicated in meters above the metal rail. The mobile marker could be displaced to a maximum distance of 8.5 m measured from the listener seat

participant (D = 1, 2, 3, 4, 5, and 6 m). The speaker was controlled by a stereo soundcard (PreSonus AudioBox 2×2 USB).

Auditory stimulus consisted of 500-ms white noise clips (measured bandwidth 0.05–20 kHz) with onset and offset ramped by a 50-ms raised cosine. The stimulus bandwidth (white noise) was chosen to maximize the availability of acoustical distance cues yielding to more accurate responses (Spiousas et al., 2017). The duration of the sound clips was set on 500 ms to have a temporally limited clip (in order to reduce the duration of the procedure and prevent subject fatigue), but long enough to minimize the onset and offset influence. This value is above the duration for which the perceived loudness is dependent on the stimulus duration (the so-called "critical duration", approximately 150 ms; Stévens & Hall, 1966), making the results of this study comparable with studies carried out using longer stimuli. The sound level of the stimulus was fixed to a comfortable level of approximately 70 dBA measured at the participant's position with the sound source located at D = 1 m.

Between trials, a masking sound was presented through two loudspeakers located at both sides of the participant (Fig. 1B) at a similar level to that of the stimulus. This masker served the purpose of masking any possible noise related to the speaker movement procedure. Participants reported that these sounds were properly masked by the masker sound. Two seconds after the end of the masking sound, the auditory stimulus was presented through the test speaker. See Calcagno et al. (2012) and Spiousas et al. (2017) for a thorough description of the experimental set-up.

The CMDL response method consisted of a system whereby the participant could move (using a hand-held remote control) a visual marker that runs parallel to the possible positions of the loudspeaker (20 cm to the right) at the height of the listener's head (Fig. 1C). The possible maximum value of the response was 8.5 m measured from the listener. The visual

marker consisted of a pair of green LEDs (standard, 3 mm) located vertically, and separated 4 cm from each other. The distance from the participant to the visual marker was measured with an optical encoder (Fig. 1D) connected directly to an Arduino Duemilanove micro-controller interface. The system allowed a spatial resolution of approximately 5 mm.

It is worth highlighting that all the experiments were performed in complete darkness and that participants could not see any object in the room (walls, ceiling, floor, sound source, etc.) during the procedure. The intensity of the LEDs was adjusted to prevent illumination of the room surfaces. In addition, in order to avoid reflections from the target speaker, the speaker was covered with a non-reflective opaque material, and the LEDs were covered with a sheath that ensured illumination only towards the participant's position.

Experiment 1

Procedure

The purpose of this experiment was to compare VR and CMDL as response methods to measure ADP in the far field. Sixteen subjects participated in this experiment. One half of the participants (Group A) employed VR as response method, using a scale of meters with a precision of one decimal. The other half (Group B) employed the CMDL method previously described. Before entering the test room, each participant was instructed on the task to be performed. Then, the participant was blindfolded and led into the test room, where s/he was seated in a chair located at the zero point. Finally, the lights were turned off and the blindfold was removed.

The procedure consisted of presenting the auditory stimulus (500-ms long, white noise clips) at one of the six distances (D

= 1, 2, 3, 4, 5 and 6 m), and then asking the participant to indicate the apparent distance from the listener to the sound source. The exact wording used for the instructions (translated from Spanish) was "The task consists of reporting the distance between you and the sound source, as you perceive it. This means that you do not have to guess the real location of the source, but instead the distance at which you sense the source is located". We also emphasised the fact that there were no correct or wrong responses. We checked that the subjects understood the task by asking them whether they had any doubt or question. In the case of Group B (CMDL), subjects were also instructed to move the visual marker from the zero point to the distance where they perceived the sound source. After each trial, the participants were instructed to move the visual marker back to the zero point. Before performing the test, the subject was asked to explore the complete range of positions (8.5 m) for the visual marker. Each testing distance was repeated three times in random order, giving a total of 18 trials per block. Only one response was collected per trial, and the participants did not receive any feedback regarding their responses. The experiment lasted approximately 9 min for each group.

Results

Figure 2a displays the mean subjective distance judgments (±SEM) obtained with both response methods, as a function of the physical distance (mean values and confidence intervals are shown in Table S1, Experiment 1 CMDL and VR). Data from one participant of Group B were excluded from the analysis due to a failure in the recollection of responses during the experiment. The perceived distance under both response conditions shows minor differences for $D \leq 3$ m that increase progressively for farther targets. Furthermore, the response with VR shows a plateau for targets located at $D > 3$ m while for CMDL the response is linear for the full range of target positions.

We analyzed the differences between both curves by means of a split-plot ANOVA[1] applied to the response, with "target distance" (six levels, within-subjects) and "response method" (two levels, between-subjects) as fixed factors. Consistently with the visual inspection of the results, the test yielded a significant effect of both factors [distance: $F(2.86, 37.2) = 70.6$, $p = 3.5 \times 10^{-15}$, $\eta_p^2 = 0.85$; response method: $F(1, 13) = 10.0$, $p = 7.4 \times 10^{-3}$, $\eta_p^2 = 0.44$] and their interaction [$F(2.86, 37.2) = 10.6$, $p = 4.5 \times 10^{-5}$, $\eta_p^2 = 0.45$]. Due to the presence of a strong interaction among factors, we also tested the difference for each distance separately; we found significant differences among methods for the two farthest targets [two-tailed, two-sample t-test with sequential Holm-Bonferroni correction for six comparisons; $D = 5$ m: $t(13) = 3.60$, $p = 0.0032 < 0.01$,

[1] Throughout the paper, the Greenhouse-Geiser correction was employed to correct sphericity violations. In those cases we reported the corrected degrees of freedom for the F-statistic along with the corresponding p-value.
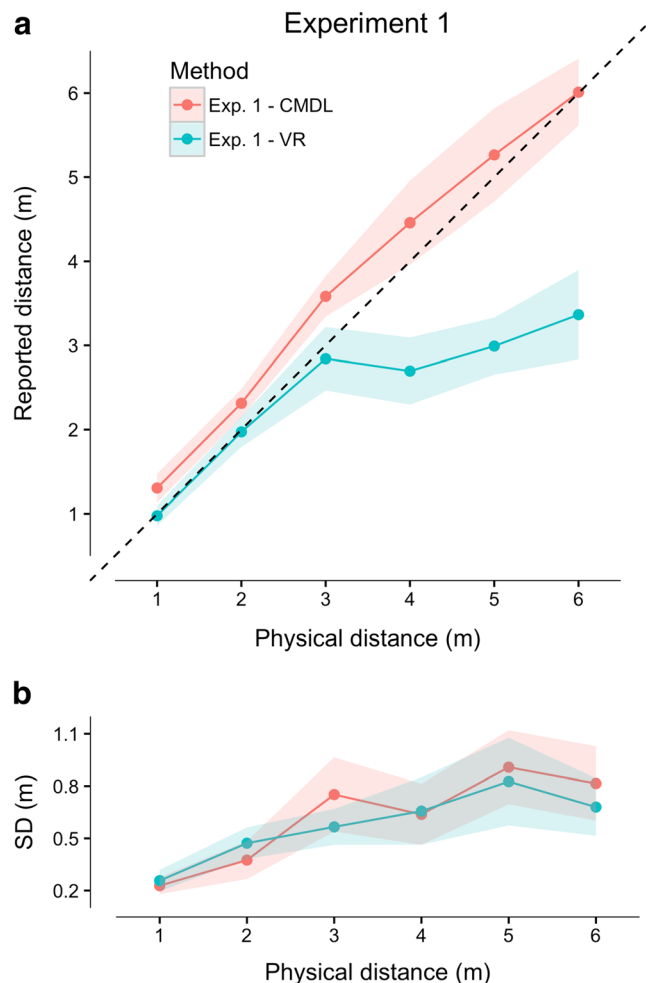
Fig. 2 Results of Experiment 1. (**a**) Auditory distance perception mean subjective responses (±SEM) obtained with CMDL (red) and VR (cyan) methods, as a function of source distance. Black dashed line indicates perfect performance (response = true distance). (**b**) Between-subjects average (±SEM) of the individual standard deviation for each distance and response method

Cohen's $d_s = 1.86$; $D = 6$ m: $t(13) = 3.89$, $p = 0.0018 < 0.0083$, Cohen's $d_s = 2.02$].

In order to further explore these differences, we calculated the percentage response range (see Table 1), defined as the difference between the maximum and minimum mean individual reported distances, normalized by the physical distance range. We analyzed three physical distance ranges: 1–6 m, 1–3 m, and 3–6 m. These ranges were chosen based on the apparent change with distance of the slope in the VR response curve, and are in line with previous studies on ADP that showed increased underestimation after 3 m (Calcagno et al., 2012). We found significant differences (two-sample t-test, Holm-Bonferroni corrected for three comparisons) for the full response range (1–6 m) [$t(13) = 3.96$, $p = 1.6 \times 10^{-3} < 0.025$, Cohen's $d_s = 2.05$], with wider response ranges for CMDL than for VR ($M = 95.1\%$ vs. $M = 52.0\%$). Regarding the two shorter ranges (1–3 m and 3–6 m), we found significant differences for the 3–6 m range [$t(13) = 5.85$, $p = 5.7 \times 10^{-3} <$

Table 1 Percentage response range

Experiment	1–6 m: M (CI) (%)	1–3 m: M (CI) (%)	3–6 m: M (CI) (%)
Exp. 1 VR	52.0 (35.6–68.3)	94.3 (65.1–123.4)	37.2 (27.3–47.2)
Exp. 1 CMDL	95.1 (82.0–108.1)	113.9 (96.2–131.7)	84.6 (72.0–97.3)
Exp. 2 VR	55.3 (37.2–73.5)	74.1 (39.4–108.7)	59.7 (40.9–78.6)
Exp. 3 VR	62.0 (29.9–94.0)	75.6 (37.4–113.9)	57.5 (30.2–84.8)
Exp. 3 CMDL	95.4 (72.0–118.9)	104.7 (73.7–135.7)	92.1 (64.9–119.2)

CI 95% confidence interval

0.017, Cohen's d_s = 3.03], with a wider response range for CMDL than for VR (M = 84.6% vs. M = 37.2%), but not for 1–3 m [$t(13)$ = 1.09, p = 0.30, Cohen's d_s = 0.56]. These results support the observation that the differences in accuracy for both methods increase largely after 3 m, evidencing differences in response compression.

The compression of the response can be addressed by fitting the individual responses with a power function of the form $Y = aX^b$, where X is the physical distance, Y the perceived distance, and the exponent b accounts for the compression rate. The individual values were averaged across subjects (see Table 2). For the VR method, we obtained a highly compressive relation between response and source distance (R^2 = 0.857; a = 1.114 and b = 0.654). On the other hand, the CMDL data resulted in a more linear, yet slightly compressive, relationship (R^2 = 0.969; a = 1.306 and b = 0.875). Interestingly, both methods differed in the compression of the response [$t(13)$ = 2.86, p = 0.013, Cohen's d_s = 1.50].

We also measured the signed and unsigned percentage error (SPE and UPE, respectively), which are defined as: $SPE = (Y/X - 1) \times 100\%$ and: $UPE = \text{abs}(SPE)$. The sign of the SPE is an indicator of overall overestimation ($SPE > 0$) or underestimation ($SPE < 0$) in the response. The UPE, on the other hand, is a positive-definite magnitude, which measures the degree of consistency of the bias. A value of UPE greater than SPE indicates that subjects combined overestimation and underestimation in the response. We first computed the individual mean errors averaged across distance, and then

averaged them across subjects. The SPE for VR showed underestimation in the response (M = −20.9%, 95% CI [−38.1, −3.75]), while, for the CMDL method, the SPE showed a minor overestimation (M = 13.8%, 95% CI [−4.34, 31.8]). The SPE differed significantly across response methods [two-sample t-test, $t(13)$ = 2.72, p = 0.017, Cohen's d_s = 1.41]. Nevertheless, the UPE was similar in both cases [VR: M = 32.6%, 95% CI [24.6, 40.6]; CMDL: M = 24.1%, 95% CI [14.8, 33.4]; two-sample t-test, $t(13)$ = 1.37, p = 0.19, Cohen's d_s = 0.71], indicating that both methods displayed the same degree of consistency in bias across the whole target distance range.

Finally, we analyzed the variability in the response across trials. Fig. 2b displays the standard deviation averaged across subjects (±SEM) as a function of source distance. Although the number of trials per condition was low (three trials), the number of subjects (eight) and experimental conditions (six distances × two response methods) can compensate for the low degrees of freedom in the measurement of the SD for each target. The data shows a clear pattern, which is confirmed by ANOVA (same test as with the mean response): the intra-subject variability increases with distance [$F(5, 65)$ = 5.37, p = 3.4 × 10^{-4}, η_p^2 = 0.29] and is not statistically different across response methods [$F(1, 13)$ = 0.081, p = 0.78, η_p^2 = 6.2 × 10^{-3}]. The variabilities show no apparent interaction [$F(5, 65)$ = 0.307, p = 0.91, η_p^2 = 0.23]. In conclusion, the response with both methods resulted in the same intra-subject variability.

Table 2 Power function parameters

Experiment	a (CI)	b (CI)	R^2
Exp. 1 VR	1.114 (0.886–1.342)	0.654 (0.539–0.770)	0.857
Exp. 1 CMDL	1.306 (1.042–1.571)	0.875 (0.788–0.962)	0.969
Exp. 2 VR	1.111 (0.804–1.418)	0.670 (0.518–0.821)	0.878
Exp. 3 VR	1.263 (0.685–1.840)	0.728 (0.456–1.000)	0.917
Exp. 3 CMDL	1.916 (1.191–2.641)	0.737 (0.546–0.928)	0.968

CI 95% confidence interval

R^2 is the goodness of fit of the individual responses by a power function of the family $Y = aX^b$, where X is the physical distance and Y the perceived distance

Discussion

The results obtained in Experiment 1 show that CMDL responses were more veridical (less biased and less compressed) with respect to those obtained with VR (Fig. 2a). VRs were consistent with previous results obtained in our laboratory under the same experimental conditions (Calcagno et al., 2012), with responses displaying a linear increase for short distances, and then becoming almost constant when the source distance D is increased beyond 3 m. The main difference between response methods across methods is that, for $D > 3$ m, VRs showed a strong negative bias whereas CMDL did not. This is reflected in

the significant effect of the interaction between response method and target distance on the response and, more specifically, in the differences on the responses for $D = 5$ and 6 m. A possible cause of this effect can be found in the decrease of the values of both the 3–6 m response range and the power-function exponent of VR compared to that obtained with CMDL (see Table 1). The compression of the response observed with VR was similar to that reported by Zahorik et al. (2005) in a meta-analysis of 21 previous ADP studies ($M = 0.65$, vs. $M = 0.54$, respectively), and also to that obtained in Calcagno et al. (2012) ($M = 0.55$, 95% CI [0.54, 0.56]). Conversely, CMDL showed an average response with minimum bias and compression. These results are compatible with many ADP (Ashmead et al., 1995; Brungart et al., 2000) and VDP (Andre & Rogers, 2006; Brungart et al., 2000) studies in which DL methods produced more veridical responses than VR.

One aspect of the CMDL method should be considered carefully. By imposing a maximum value on the response (8.5 m in our experiments), the participants are constrained to respond within a limited range of distances, potentially limiting or distorting CMDL judgments for the farthest sources. First, having an upper bound for the possible responses could induce participants to scale their CMDL judgments across the full range of available distances, locating the farthest perceived sounds close to the farthest available response positions, and spreading their judgments for the nearer sounds across the full range. Second and alternatively, participants could have been unable to report perceived distances beyond the limit imposed by the upper bound of the response range (so-called "ceiling effect", see Kopčo & Shinn-Cunningham, 2011). This drawback is common to most DL methods (since they are performed in real space) and, therefore, in order to avoid it, it is necessary to perform experiments in spaces whose maximum distance far exceeds the maximum distance to which the target is located (Bidart & Lavandier, 2016). This disadvantage is not present in VRs since no limit is imposed on participants' judgments, thus avoiding the potential confound generated by the environment boundaries.

Although we cannot rule out that these effects have affected the CMDL response in Experiment 1, we argue that, in case they existed, their influence was weak. On the one hand, if the participants scaled their perceived distances to the upper distance limit of the CMDL device, the maximum perceived distance would have been expected to be close to that limit (8.5 m). However, participants perceived, on average, the maximum distance near the actual distance of the sound source (6 m). In fact, the CMDL response shows a slight compression for source distances greater than 4 m, and in no case were the responses near the boundary imposed by the CMDL device. On the other hand, if the participants overestimated the location of the farthest source (6 m) beyond the boundary imposed by the CMDL device, it would be expected the location of the

4-m and 5-m sources to be overestimated too. Had this happened, we would observe a clutter of responses for the farthest sources near the 8.5-m limit, and therefore a strong compression in the overall pattern of CMDL responses, which did not happen. Moreover, a previous study (Kopčo & Shinn-Cunningham, 2011) showed that a ceiling effect may induce a reduction in the response variability near the farthest limit, which was not observed in our data either (see Fig. 2b). Finally, Bidart and Lavandier (2016) performed far-field ADP virtual experiments employing a response method which imposed a restricted range of possible auditory distances responses (subjects had to move a cursor on a horizontal continuous linear scale graduated and labelled with distance values every 1 m between 0 and 15 m, displayed on a computer screen). Their results showed that responses were not affected by a ceiling effect. In their work, the farthest virtual source was located at 10 m, while the maximum allowed response was 15 m (a relative difference of 50%). In our study, the relative difference between both magnitudes (6 and 8.5 m, respectively) was similar (42%), and therefore a similar result would be expected. Another aspect to be considered is that far-field ADP is characterized by underestimation of the perceived distance. For this reason, it is unlikely that such a ceiling effect would arise as long as the response range is larger than the target range. On the contrary, this aspect should be carefully taken care of in near-field ADP studies, as near-field responses have a tendency to be overestimated.

Previous studies showed that visual environmental information reduces both bias and compression of verbal ADP responses (Calcagno et al., 2012; Zahorik, 2001). In Experiment 1, the participants who employed the CMDL method obtained visual spatial information through the movement of the visual marker. The presence of an upper bound for the CMDL responses could have provided the subjects with information related to the room dimensions, in particular, a lower bound for the distance to the far wall. In contrast, the participants of the VR group did not have any visual information about the room dimensions. In this context, we do not know how much of the improvement in the response accuracy observed with the CMDL method is due to the response method itself, or to the additional spatial visual information available when using the CMDL device. Furthermore, we do not know for sure how accurate listeners are at estimating the distance to the mobile visual marker. That is, do listeners know the actual distance to which the visual marker is located when they respond with CMDL? We conducted a second experiment in order to answer both questions.

Experiment 2

Experiment 2 was divided in two parts, each one aiming to respond to each of the questions posited in the previous

section. First, we studied the visual perception of distance to the CMDL mobile target. To this end, participants were instructed to move the visual marker to a distance indicated verbally by the experimenter, allowing us to measure the relationship between the physical distance indicated by the experimenter and the distance estimated by the participant using the mobile visual marker. Immediately after completion of the first task, participants performed an ADP task using VR as response method. This second task aimed to reveal whether the spatial visual information provided by the visual marker (during the first task) can influence VR estimates of auditory distances.

Procedure

Eight subjects participated in this experiment, none of which took part in Experiment 1. Before entering the test room, each participant was instructed on the task to be performed. Then, s/he was blindfolded and led into the test room, in which s/he was seated in a chair located at the zero point. The experiment was conducted in complete darkness, and the participant kept his eyes uncovered so that s/he could only see the visual mobile marker (during the first task). Before the first test, the participant was asked to employ the hand-held control to move the visual marker along its full range (8.5 m). The first task consisted of moving the visual marker from the zero point ($D = 0$) to a distance verbally indicated by the experimenter (who was seated at the right side of the participant, at a distance of 1 m). The second task consisted of providing VR estimates of ADP, under

the same conditions as in Experiment 1, and was performed immediately upon completion of the first task, without leaving the seat. For both tasks, target distances were $D = 1, 2, 3, 4, 5$, and 6 m. Each target distance was repeated three times in random order, giving a total of 18 trials per block. Both the experimental set-up and the targets' locations were the same as in the previous experiment. Each block lasted approximately 9 min.

Results

In Fig. 3a the final positions of the visual markers are displayed as a function of the distance indicated by the experimenter. Each data point corresponds to the mean across subjects (±SEM). The average response shows a minimal bias, as can be confirmed by adjusting a power function of the form $Y = aX^b$ to the data. The resulting parameters correspond closely to the veridical relation: $Y = X$ ($R^2 = 0.986$; $a = 1.242$ and $b = 0.909$). This response was less biased than those obtained in previous studies of visual perception of fixed objects located in the dark (Calcagno et al., 2012; Philbeck & Loomis, 1997). A possible cause for such a difference is that subjects exploited some of the additional dynamic cues available when the visual marker is moving directly away from or towards the observer. Also, it is possible that the use of the hand-held control has brought additional temporal information. The fact that the velocity of the mobile marker was relatively constant implies that the response duration was proportional to the displacement of the marker. This temporal information may have helped listeners to improve their performance by
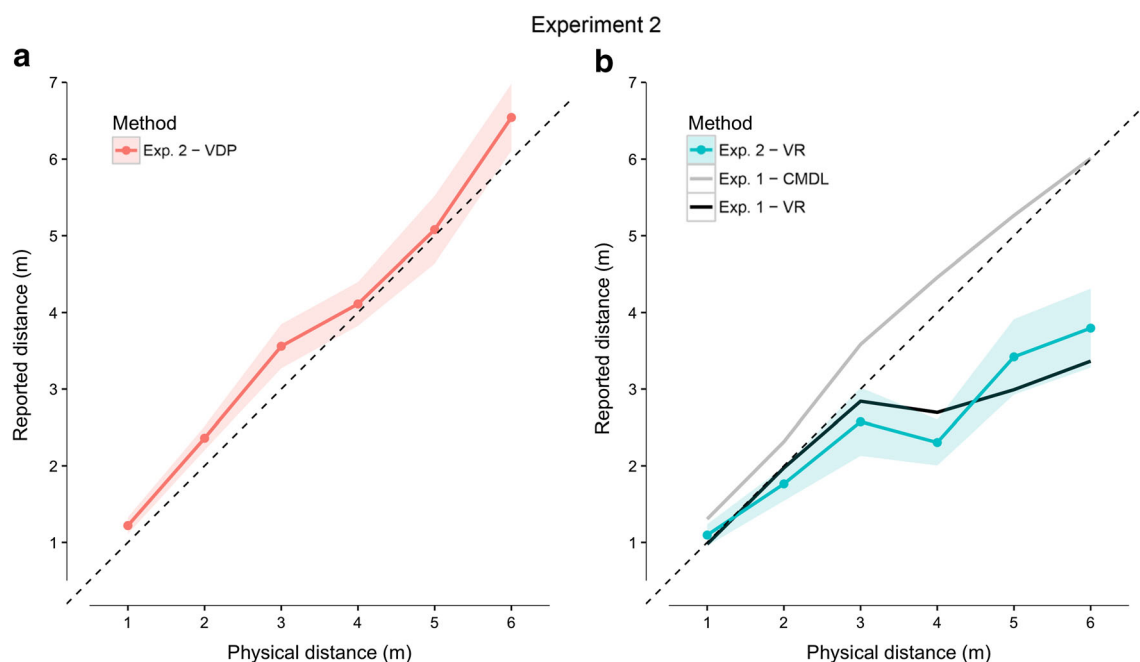
Fig. 3 Results of Experiment 2. (**a**) Results of the VDP task for each distance indicated by the experimenter. (**b**) ADP responses obtained in Experiment 2 with VR (cyan) compared to CMDL (grey) and VR (black)

and responses obtained in Experiment 1. Both panels show the mean subjective response ± SEM. Black dashed lines indicate perfect performance (response = true distance)

estimating the time needed to bring the visual marker to shorter distances, where the response is more precise, and then extrapolating to larger distances. If this was the case, subjects may have required a certain number of trials in order to acquire the information of the marker velocity, and, at the same time, they may have improved the response as they became familiar with the motion characteristics of the device. In turn, this would have resulted in a change (recalibration) of the response across trials. In order to test for this hypothesis, a two-way, within-subjects ANOVA was applied to the response, with fixed factors "target distance" (six levels) and "trial number" (three levels). The test showed a non-significant effect of neither "trial number" [$F(2, 14) = 1.20$, $p = 0.33$, $\eta_p^2 = 0.15$] nor the interaction between "trial number" and "target distance" [$F(3.84, 26.9) = 0.995$, $p = 0.42$, $\eta_p^2 = 0.12$], suggesting that there was no recalibration of the response during the experiment. Furthermore, a majority of subjects responded by making several approaches (between three and four) of the visual marker before reaching its final position, which is in contradiction with a response governed purely by temporal information. Although we did not measure the detailed structure of the response, this qualitative observation, combined with the lack of an effect of "trial number" on the response, argues in favor that visual (instead of temporal) cues were the main source of information employed by the subjects during the task.

Figure 3b displays the mean subjective auditory distance judgments (±SEM) obtained with VR (during the second task) as a function of source distance (mean values and confidence intervals are shown in Table S1, Exp. 2 VR). The response follows closely that obtained with VR in Experiment 1. In order to test whether the visual information available in the first part of the experiment influenced the VRs, we compared the results obtained with VRs in Exps. 1 and 2 (split-plot ANOVA with "target distance" and "experiment" as within- and between-subjects fixed factors, respectively). We found no significant differences among experiments [experiment: $F(1, 14) = 0.002$, $p = 0.96$, $\eta_p^2 = 1.3 \times 10^{-4}$; experiment × target distance: $F(2.09, 29.3) = 1.40$, $p = 0.26$, $\eta_p^2 = 0.091$], indicating that the information provided by the previous calibration task (both the visual cues and the calibration itself) did not modify the ADP response obtained immediately after with VR. We also analyzed the compression of the ADP response by fitting power-law functions to the individual responses. The resulting best estimates of the parameters were similar to that obtained for the response in Experiment 1 ($R^2 = 0.878$; $a = 1.111$ and $b = 0.670$). The exponent showed no significant differences when compared to the one obtained in Experiment 1 using VR [$t(14) = 0.157$, $p = 0.88$, Cohen's $d_s = 0.078$]. The only difference found (two-sample t-test, Holm-Bonferroni corrected for three comparisons) between these two conditions was for the percentage response range for higher distances (3–6 m), where there was an increase in

Experiment 2 compared to Experiment 1 [$M = 59.7\%$ vs. $M = 37.2\%$; $t(14) = 3.23$, $p = 0.0060 < 0.017$, Cohen's $d_s = 1.11$].

Finally, we analyzed the intra-subject variability of the VRs (by means of the individual SD). The pattern is very similar to that obtained in the previous experiment, with the SD increasing with distance in a seemingly linear fashion. We compared the results with the VR of Experiment 1 (ANOVA with the same characteristics as for the mean response). We found a significant effect of the target distance [$F(2.97, 41.6) = 4.66$, $p = 6.9 \times 10^{-3}$, $\eta_p^2 = 0.25$], but there were no significant differences across experiments [experiment: $F(1, 14) = 0.87$, $p = 0.37$, $\eta_p^2 = 0.058$); experiment × target distance: $F(2.97, 41.6) = 0.117$, $p = 0.95$, $\eta_p^2 = 8.3 \times 10^{-3}$]. Therefore, we conclude that the information provided by the CMDL method during the first task did not modify the intra-subject variability of the VRs in the following ADP task.

Discussion

The results of the first task show that the CMDL device gives the participants the necessary information to know how far the visual marker is located. The participants were able to accurately locate the visual mobile marker at a distance previously indicated by the experimenter. A similar result was obtained by Kolarik, Pardhan, Cirstea, and Moore (2016b) in a blind-walking task, in which both sighted and blind subjects were able to accurately walk to a distance previously indicated by the experimenter.

Many previous studies have indicated that visual spatial information can be stored in memory and then used in experiments performed in the dark (Andre & Rogers, 2006; Calcagno et al., 2012; Loomis et al., 1998). If the response observed with CMDL in Experiment 1 was induced by visual spatial information provided by the visual mobile marker, it would be expected that the information obtained in the first part of Experiment 2 would affect the VR response measured in the following ADP task; however, this was not the case.

This result does not match what was reported in previous studies where the stored environmental visual information induced more accurate VR responses (Calcagno et al., 2012). The main difference between the visual conditions used here and in this previous study resides in the number and complexity of the visual cues presented in each case. Perhaps the increased amount of visual references in Calcagno et al. (2012) facilitated the memorization of the visual context, creating a robust memory of the room where events occurred. However, we do not know exactly what the effect of such spatial memory on the visual information provided by the CMDL device is. The scarcity of visual references during the manipulation of the mobile visual marker could have made it more difficult to keep the spatial visual information in memory throughout the whole experiment.

In order to minimize the influence of memory on the VR response, we performed a third experiment where, for each trial, the participants performed a CMDL report immediately followed by a VR.

Experiment 3

The goal of this experiment was to test whether the visual cues provided by the CMDL method influence the VR responses when each method is employed in subsequent trials. To this end, we performed an ADP experiment in which the response methods were interleaved, allowing us to study the influence of the CMDL method on VRs more directly.

Procedure

Eight subjects participated in this experiment, none of which took part in Experiments 1 or 2. Subjects performed an ADP task employing successively the CMDL and VR methods; i.e., in a given trial, the subject responded with CMDL, and in the next trial with VR. This design allowed us to minimize any effects of memory from one response method to the other. Under this constraint, each combination of position and

Fig. 4 Results of Experiment 3. Average ADP responses obtained in Experiment 3 with CMDL (red) and VR (cyan) compared to CMDL (grey) and VR (black) responses obtained in Experiment 1. The panels show the mean subjective response ± SEM. The black dashed line indicates perfect performance (response = true distance)

response method was randomly presented three times, giving a total of 36 trials (six positions × 2 response methods × three trials = 36 trials). Both the experimental set-up and the targets' locations were the same as in previous experiments. The experiment lasted approximately 18 min.
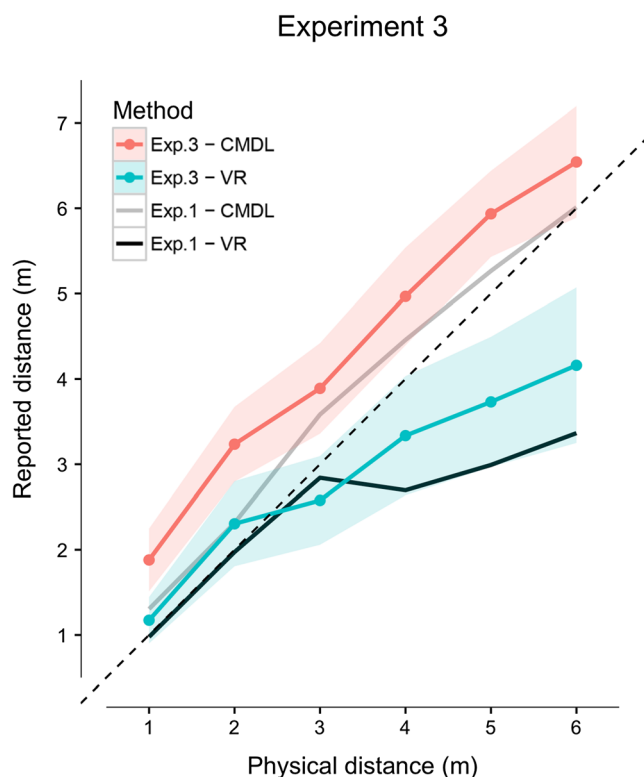
Results

Figure 4 shows the average response (±SEM) obtained in Experiment 3 with both methods (mean and confidence intervals are shown in Table S1, Experiment 3 CMDL and VR). Similar to Experiment 1, the range of the response is greater using CMDL than using VR (VR: $M = 3.10$ m vs. CMDL: $M = 4.77$ m). Although the participants used CMDL and VR sequentially, the responses with both methods show a pattern similar to that observed in Experiment 1. However, CMDL responses appear more overestimated, and VRs seem less compressive, compared to the results of Experiment 1 (grey and black lines in Fig. 4).

We started by analyzing the difference between response methods by means of a repeated-measures ANOVA with "target distance" and "response method" as fixed factors. The analysis showed a significant effect of both main factors [target distance: $F(1.37, 9.60) = 26.9$, $p = 2.4 \times 10^{-4}$, $\eta_p^2 = 0.79$; response method: $F(1, 7) = 8.59$, $p = 0.022$, $\eta_p^2 = 0.55$] and the interaction [$F(5, 35) = 7.19$, $p = 1.2 \times 10^{-4}$, $\eta_p^2 = 0.51$]. Due to the presence of a strong interaction, we compared the response across methods for each distance separately. We obtained significant differences for all but the target located at 2 m [one-tailed, paired-sample t-test with Holm-Bonferroni correction for six comparisons; $D = 1$ m: $t(7) = 3.37$, $p = 0.0060$; $D = 3$ m: $t(7) = 2.57$, $p = 0.018$; $D = 4$ m: $t(7) = 2.96$, $p = 0.011$; $D = 5$ m: $t(7) = 3.05$, $p = 0.0092$; $D = 6$ m: $t(7) = 3.42$, $p = 0.0056$], which is also a major difference with respect to the results of Experiment 1.

Considering that both the CMDL and VR responses appear to differ from the respective results of Experiment 1, we also searched for statistical differences between Experiments 1 (each method in isolation) and 3 (interleaved). We compared each response method separately by means of two split-plot ANOVAs with "target distance" (within-subjects) and "response method" (between-subjects) as fixed factors. For both response methods the test showed no significant differences across experiments [VR: distance: $F(1.58, 22.1) = 25.7$, $p = 6.0 \times 10^{-6}$; experiment: $F(1, 14) = 0.39$, $p = 0.54$; interaction: $F(1.58, 22.1) = 1.13$, $p = 0.33$; CMDL: distance: $F(2.44, 31.7) = 89.7$, $p = 1.2 \times 10^{-14}$; experiment: $F(1, 13) = 1.09$, $p = 0.31$; interaction: $F(2.44, 31.7) = 0.30$, $p = 0.78$].

Next, we analyzed the compression of the response. In this experiment none of the percentage response ranges were significantly different across methods [two-sided, paired t-test

with Holm-Bonferroni correction for three comparisons; 1–6 m: $t(7) = 2.93$, $p = 0.022$; 1–3 m: $t(7) = 1.72$, $p = 0.13$; 3–6 m: $t(7) = 3.00$, $p = 0.020$; the three comparisons result in no significant differences after correcting for multiple comparisons] suggesting a smaller effect on the response compression. Furthermore, both methods were well-fitted by power-law functions of the form $Y = aX^b$ (CMDL: $R^2 = 0.968$; a: 1.916; b: 0.737; VR: $R^2 = 0.917$; a: 1.263; b: 0.728) with non-significant differences between exponents [$t(7) = 0.120$, $p = 0.91$].

Finally, we analyzed the response error. The SPE indicated that the CMDL response was systematically overestimated, while VRs were slightly underestimated [CMDL: $M = 38.6\%$, 95% CI [6.64, 70.5]; VR: $M = -9.04\%$, 95% CI [−43.4, 25.3]; two-tailed, paired-sample t-test, $t(7) = 2.83$, $p = 0.025$, Cohen's $d_z = 0.99$]. The UPE, on the other hand, was very similar for both methods [CMDL: $M = 48.2$, 95% CI [23.3, 73.1]; VR: $M = 49.1$, 95% CI [41.6, 56.7]; two-tailed, paired-sample t-test, $t(7) = 0.084$, $p = 0.93$].

Discussion

The results of Experiment 3 show that, although each response method was employed in subsequent trials, their estimates remained statistically different. Contrary to our initial hypothesis, the visual references provided by the CMDL trials were insufficient to eliminate the underestimation of the VR judgments for distant sources. Although we did not find significant differences between the estimates obtained in Experiments 1 and 3 for each response method, we unexpectedly found a few evidences of mutual influence. Unlike Experiment 1, in Experiment 3 neither the compression nor the response range showed significant differences between methods. This result may be due to the fact that, compared to Experiment 1, VRs obtained in Experiment 3 appeared to be less compressive ($M = 0.728$ vs. $M = 0.654$, respectively), while CMDL estimates were more compressive ($M = 0.737$ vs. $M = 0.875$ respectively). We hypothesize that interleaving both response methods in successive trials may have caused an association between the acoustical cues related to the source distance (identical for both response conditions) and the perceptual representation inherent to each methodology, influencing the perceptual calibration of the other method (CMDL inducing less compressive VRs, and VRs inducing more compressive CMDL estimates). Another difference between Experiments 1 and 3 is that, while in Experiment 1 only the two farthest sources ($D = 5$ and 6 m) were perceived farther away with CMDL compared to VRs, in Experiment 3 this effect was observed for all but one ($D = 2$ m) of the tested physical distances. Considering Fig. 2, this difference seems to be explained by a greater overestimation of the source distance with CMDL (mainly at close distances, $D = 1$ and 2 m) compared to the estimates obtained when testing this method in isolation (Exp. 1). This

result suggests that, by interleaving the response methods, VRs influenced the CMDL responses, inducing an overestimation of the perceived distance for close sources. This behavior was an unexpected but interesting result. Unfortunately, the present study does not allow us to be conclusive about the causes of this outcome. However, a speculation based on previous studies might be posed.

The overestimation of the responses found in the CMDL distance curves of Experiment 3 resembles that obtained in three previous far-field ADP studies (Bidart & Lavandier, 2016; Cabrera & Gilfillan, 2002; Calcagno et al., 2012). A common aspect of these studies is that all of them involved some sort of visual map of the physical space. In Bidart and Lavandier (2016) the participants indicated the distance by moving a cursor on a horizontal continuous linear scale displayed on a computer screen, In Calcagno et al. (2012) visual markers were fixed at distances known by the participants (2, 4, 6 and 8 m). Finally, Cabrera and Gilfillan (2002) employed a series of labeled pointers placed directly in front of the participant at 1-m intervals, the farthest being 8-m distant. In the first study participants could only access a virtual representation of the physical space while, in the latter, explicit visual anchors were placed in real space. Could the overestimation observed in Experiment 3 have been caused by the internalization of some sort of spatial map? If this hypothesis was true, in Experiment 3 the subjects would associate the VR numerical estimates for the nearest sources (for which VRs show minimum bias and variability) with the respective acoustical distance cues, resulting in an auditory map analogous to the explicit visual maps of the aforementioned studies. If this was the case, these fixed landmarks could have influenced the CDML responses, causing them to be overestimated. However, this is only a speculation, and further research is necessary in order to understand the observed effects presumably caused by the interaction between different response methods in successive trials.

In summary, although in Experiment 3 both methods were used in successive trials, we observed remarkable differences between their estimates. Moreover, the responses obtained in Experiment 3 with each method were not significantly different from the respective responses obtained in Experiment 1. If the differences between VR and CMDL observed in the Experiment 1 were due to the presence of extra visual spatial information during the CMDL task, the successive use of both methods should have markedly improved VRs. However, this was not observed. Despite having the same visual information present during CMDL trials, the results of Experiment 3 showed a small improvement of VRs in terms of the compression of the curve, but this improvement was insufficient to eliminate the bias in the responses for the farthest targets, which remained similar to that of Experiment 1. Also, we found evidence that VRs influenced the CMDL response, inducing overestimation for the nearest targets. Considering this

result along with those obtained in Experiment 2, we conclude that the differences observed in the responses obtained in Experiments 1 and 3 with CMDL and VRs were mainly due to factors inherent to the response methods themselves, and not due to changes in calibration dependent on the visual information provided by the CMDL mobile visual marker.

General discussion

The main goal of this study was to evaluate the suitability of the proposed CMDL method to measure ADP estimates for sources located in the far field. Experiment 1 showed that CMDL responses were significantly less biased and less compressive than VRs. This result is in line with numerous previous studies of both auditory and visual distance perception where VRs tended to underestimate distances longer than ~3 m (Anderson & Zahorik, 2014; Andre & Rogers, 2006; Calcagno et al., 2012; Cutting & Vishton, 1995; Kelly et al., 2004; Loomis et al., 1998; Zahorik, 2001, 2002), while estimates using blind-walking were accurate (Ashmead et al., 1995; Loomis et al., 1992; Thomson, 1983).

Results of Experiment 2 showed that participants were precise at locating the visual marker at a distance previously presented verbally by the experimenter, indicating that, when responding with CMDL, participants seemed to know quite accurately the actual distance to the visual marker. This means that, when using the CMDL device, participants had access to spatial visual information (for example that the room is at least 8.5-m long) that was not available for participants who responded verbally. However, the results of Experiments 2 and 3 showed that this information had little influence on the ADP VRs. In this line, VRs from Experiments 2 and 3 did not show significant differences for any of the variables (response, compression and response range) obtained with the same method in Experiment 1. If the visual information provided by the CMDL device (and not the method itself) was responsible for the more accurate responses observed in Experiment 1, it would have been expected that this information had strongly influenced VRs in both Experiments 2 and 3, which was not observed. The combined results of Experiments 2 and 3 suggest that the observed differences across methods in Experiment 1 were not induced by the extra spatial information obtained while using the CMDL device.

The differences between VR and CMDL responses should not be automatically interpreted as due to changes in the perceived distance induced by the method. The difference could also be caused by changes in the calibration of the reported distance while the perceived distance remained unchanged. The results obtained here do not allow us to be conclusive about this respect, since the present evidence suggests a mixture of both factors. For example, the fact that the response obtained with CMDL and VR shows a significant

interaction (Experiments 1 and 3), suggests that both response methods use the available cues differently and construct functionally distinct underlying representations. The differences in compression and response range observed in Experiment 1 also suggest that both methods are controlled by functionally distinct representations. On the other hand, several factors suggest that the differences between VR and CMDL were due to differences in response calibration. First, the variability observed in both methods was very similar, indicating that the task was equally difficult in both cases. Second, previous studies of VDP have reported that the response accuracy is more affected by environmental settings for VRs than for DL methods (Andre & Rogers, 2006; Woods, Philbeck, & Danoff, 2009), suggesting that VRs need more spatial references to correctly calibrate distance perception. In fact, previous ADP studies showed accurate VR responses in the presence of multiple visual-context information sources (Calcagno et al., 2012; Zahorik, 2001). These results suggest that participants are able to perceive the distance to the source accurately, but need spatial references to report it correctly. One of the main differences between the two methods is that CMDL reports require no mental transformation of the target location, i.e., the participant only needs to locate the mobile marker at the perceived location of the sound source. On the contrary, VR requires the participant to mentally calculate an explicit value for the perceived location of the sound source. This step could lead to errors in the calibration of the response, especially in the dark, where the scarcity of visual reference cues increases the uncertainty in the representation of target location.

According to previous results obtained by Calcagno et al. (2012) and Zahorik (2001), we expected that the spatial information provided by the visual mobile marker would affect the VRs. However, this was not observed in either Experiment 2 or Experiment 3. Several factors could explain these contradictory results. First, in Calcagno et al. (2012) fixed targets located at distances known to the participant were used as visual distance anchors. These targets were lit throughout the experiment and therefore served as a permanent fixed reference to calculate the perceived distance. Second, the Zahorik study was performed under full visual conditions and therefore VRs were influenced by more numerous and complex sources of visual information than what could be obtained here by using the CMDL device.

In Calcagno et al. (2012), we hypothesized that the improvement in the VR responses induced by visual information could be caused by a relation between the perceived size of the room and the spatial calibration of the VR estimates. Later, Kolarik et al. (2013) tested this hypothesis and showed a positive correlation between the ADP response and the size of the room perceived through reverberation cues. The results from the first part of Experiment 2 suggest that, when moving the CMDL visual marker, the participants obtained visual

information about the length of the room and the range of possible source distances. Also, the participants received, in addition to the visual information provided by the visual marker during CMDL, verbal information about the maximum distance to which the visual marker could be carried. However, this information (either through visual or verbal spatial information) was insufficient to induce changes in VRs. Giving this contradictory evidence, we believe that more studies are needed to elucidate the influence of visual cues and room size knowledge on ADP.

CMDL is an interesting method to measure ADP because it appears to be a natural response since no mental transformation of the target location is required and subjects can use their own anatomical reference points (Brungart et al., 2000). Moreover, CMDL has great advantages in relation to other DL methods used in the far field (mainly blind-walking). First, the participant does not have to stand up to indicate the perceived distance of the sound source, which greatly facilitates the task. This is a requirement to measure ADP in people who have difficulty moving, but also makes the trial time much shorter for healthy participants. Second, the CMDL device facilitates the task and allows automatic collection of experimental data, yielding to experiments with better sample size and less noisy data points. Third, CMDL can be easily replicated in identical conditions in different environments. Fourth, CMDL allows for a continuous response, while with VR subjects have a tendency to collapse the response to the nearest meter or half-meter, even when they are allowed to report it with a larger precision (see Supplemental Fig. S2 for an analysis of this effect on our data). Fifth, according to the results presented in this paper, CMDL allows an accurate perception of the position of a sound source, even in the dark, reducing the effect of the testing environment on the perceived distance. Finally, this method comprises an amenable task for participants, reducing the exhaustion and lack of concentration during the experiment.

However, CMDL has several limitations to consider. First, it has to be performed in complete darkness, while for VR it is enough to occlude the participant's vision. This requirement limits the method to enclosed environments and complicates the experimenter's role during the procedure. Second, as the device employs a visual marker, blind participants would not be able to use this method. Third, the device used here, while simple in construction and programming, requires time and space to be mounted. For this reason, the CMDL device would make it substantially difficult to test ADP in participants' homes for participant groups for which travel to the laboratory may be difficult (e.g. Kolarik, Pardhan, et al., 2016).

The results obtained here show that CMDL could be an interesting method for measuring ADP, especially for far-field sources located in the dark. However, it is difficult to generalize the results obtained here to other experimental conditions. For example, our results were obtained using only one auditory stimulus in a single reverberant room. Moreover, we only tested a limited range of distances. It would be interesting then to study if the results obtained here can be extended to other auditory stimuli (e.g., speech, band-pass noise, etc.), other auditory environments (e.g., free field or rooms with different levels of reverberation), and distances beyond 6 m, where previous work has shown substantial underestimation of ADP judgments.

A general drawback in the study of ADP is the lack of consensus on the methodology used to measure listeners' responses. Such methodological heterogeneity clearly makes it difficult to compare the results obtained across different studies. Unifying the criteria by which ADP is measured would be a very important step forward in the understanding of this research topic, and for this reason we believe that investigating whether the CMDL method (or other methods alike) is robust under different experimental conditions may be of interest to the study area.

Author contributions R.V. designed the study. E.R.C., E.A., and R.V. performed the experiments. P.E.E., I.S., and M.C.E. analyzed the data. P.E.E., I.S., and R.V. interpreted the data and co-wrote the paper.

References

Anderson, P. W., & Zahorik, P. (2014). Auditory/visual distance estimation: Accuracy and variability. *Frontiers in Psychology, 5,* 1097.

Andre, J., & Rogers, S. (2006). Using verbal and blind-walking distance estimates to investigate the two visual systems hypothesis. *Attention, Perception, & Psychophysics, 68*(3), 353–361.

Angell, J. R., & Fite, W. (1901). From the Psychological Laboratory of the University of Chicago: The monaural localization of sound. *Psychological Review, 8*(3), 225.

Ashmead, D. H., Davis, D. L., & Northington, A. (1995). Contribution of listeners' approaching motion to auditory distance perception. *Journal of Experimental Psychology: Human Perception and Performance, 21*(2), 239.

Bidart, A., & Lavandier, M. (2016). Room-induced cues for the perception of virtual auditory distance with stimuli equalized in level. *Acta Acustica United with Acustica, 102,* 159–169.

Bronkhorst, A. W., & Houtgast, T. (1999). Auditory distance perception in rooms. *Nature, 397*(6719), 517–520.

Brungart, D. S., Rabinowitz, W. M., & Durlach, N. I. (2000). Evaluation of response methods for the localization of nearby objects. *Attention, Perception, & Psychophysics, 62*(1), 48–65.

Cabrera, D., & Gilfillan, D. (2002). *Auditory distance perception of speech in the presence of noise. Proceedings of the 2002 International Conference on Auditory Display, Kyoto, Japan, July 2–5, 2002.* http://hdl.handle.net/1853/51331

Calcagno, E. R., Abregú, E. L., Eguia, M. C., & Vergara, R. (2012). The role of vision in auditory distance perception. *Perception, 41*(2), 175–192.

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., & Thompson, W. B. (2005). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual indoor environments. *Perception, 34*(2), 191–204.

Cutting, J. E., & Vishton, P. M. (1995). Information potency and spatial layout. In W. Epstein & S. J. Rogers (Eds.), *Handbook of Perception and Cognition: Vol. 5. Perception of Space and Motion* (2nd ed., pp. 69–117). San Diego: Academic Press.

Fluitt, K., Mermagen, T., & Letowski, T. (2014). Auditory distance estimation in an open space. In *Soundscape Semiotics - Localization and Categorization*. Edited by Herve Glotin, ISBN 978-953-51-1226-6. Publisher: Intech.

Fontana, F., & Rocchesso, D. (2008). Auditory distance perception in an acoustic pipe. *ACM Transactions on Applied Perception (TAP), 5*(3), 16.

Gamble, E. A. (1909). Minor studies from the psychological laboratory of Wellesley College: Intensity as a criterion in estimating the distance of sounds. *Psychological Review, 16*(6), 416.

Iosa, M., Fusco, A., Morone, G., & Paolucci, S. (2012). Walking there: Environmental influence on walking-distance estimation. *Behavioural Brain Research, 226*(1), 124–132.

Kearney, G., Gorzel, M., Rice, H., & Boland, F. (2012). Distance perception in interactive virtual acoustic environments using first and higher order ambisonic sound fields. *Acta Acustica United with Acustica, 98*(1), 61–71.

Kelly, J. W., Loomis, J. M., & Beall, A. C. (2004). Judgments of exocentric direction in large-scale space. *Perception, 33*(4), 443–454.

Kolarik, A. J., Pardhan, S., Cirstea, S., & Moore, B. C. (2013). Using acoustic information to perceive room size: Effects of blindness, room reverberation time, and stimulus. *Perception, 42*(9), 985–990.

Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., & Pardhan, S. (2016a). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics, 78*(2), 373–395.

Kolarik, A. J., Pardhan, S., Cirstea, S., & Moore, B. C. (2016b). Auditory spatial representations of the world are compressed in blind humans. *Experimental Brain Research*. doi:10.1007/s00221-016-4823-1

Kopčo, N., & Shinn-Cunningham, B. G. (2011). Effect of stimulus spectrum on distance perception for nearby sources. *The Journal of the Acoustical Society of America, 130*(3), 1530–1541.

Larsen, E., Iyer, N., Lansing, C. R., & Feng, A. S. (2008). On the minimum audible difference in direct-to-reverberant energy ratio. *The Journal of the Acoustical Society of America, 124*(1), 450–461.

Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 906.

Loomis, J. M., Klatzky, R. L., Philbeck, J. W., & Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Attention, Perception, & Psychophysics, 60*(6), 966–980.

Loomis, J. M., Philbeck, J. W., & Zahorik, P. (2002). Dissociation between location and shape in visual space. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1202.

Parseihian, G., Jouffrais, C., & Katz, B. F. (2014). Reaching nearby sources: Comparison between real and virtual sound and visual targets. *Frontiers in Neuroscience, 8,* 269.

Philbeck, J. W., & Loomis, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance, 23*(1), 72.

Philbeck, J. W., Loomis, J. M., & Beall, A. C. (1997). Visually perceived location is an invariant in the control of action. *Attention, Perception, & Psychophysics, 59*(4), 601–612.

Rieser, J. J., Ashmead, D. H., Talor, C. R., & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception, 19*(5), 675–689.

Spiousas I., Etchemendy P. E., Eguia M. C., Calcagno E. R., Abregú E., & Vergara R. O. (2017). Sound spectrum influences auditory distance perception of sound sources located in a room environment. *Frontiers in Psychology, 8*, 969. doi:10.3389/fpsyg.2017.00969

Starch, D., & Crawford, A. L. (1909). Minor studies from the psychological laboratory of the Wellesley College: The perception of the distance of sound. *Psychological Review, 16*(6), 427.

Stévens, J. C., & Hall, J. W. (1966). Brightness and loudness as functions of stimulus duration. *Perception & Psychophysics, 1*(5), 319–327.

Thomson, J. A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance, 9*(3), 427.

Toye, R. C. (1986). The effect of viewing position on the perceived layout of space. *Attention, Perception, & Psychophysics, 40*(2), 85–92.

Woods, A. J., Philbeck, J. W., & Danoff, J. V. (2009). The various perceptions of distance: An alternative view of how effort affects distance judgments. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 1104.

Wu, B., Ooi, T. L., & He, Z. J. (2004). Perceiving distance accurately by a directional process of integrating ground information. *Nature, 428*(6978), 73–77.

Zahorik, P. (2001). Estimating sound source distance with and without vision. *Optometry and Vision Science, 78*(5), 270–275.

Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *The Journal of the Acoustical Society of America, 111*(4), 1832–1846.

Zahorik, P., Brungart, D. S., & Bronkhorst, A. W. (2005). Auditory distance perception in humans: A summary of past and present research. *Acta Acustica United with Acustica, 91*(3), 409–420.