Landscape and Urban Planning 148 (2016) 216–231

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Research paper

From 3D landscape visualization to environmental simulation: The contribution of sound to the perception of virtual environments

Mark Lindquist[a,*], Eckart Lange[b], Jian Kang[b]

[a] University of Michigan, United States
[b] University of Sheffield, United Kingdom

HIGHLIGHTS

- Sound significantly alters perceptual responses to 3D landscape visualizations.
- Realism and preference are moderated by congruency of visual and sound content.
- Eye level Google Earth visualizations receive low realism ratings.
- Aural-visual survey data collected via the web is comparable to laboratory data.
- Sound and visuals that are spatiotemporally congruent are recommended for simulations.

ARTICLE INFO

Article history:
Received 28 November 2014
Received in revised form
19 December 2015
Accepted 29 December 2015
Available online 29 January 2016

Keywords:
Aural-visual interaction
Environmental perception
Landscape
Visualization
Soundscape
Virtual landscapes

ABSTRACT

This research investigated the perceptual interaction of combining sound with 3D landscape visualizations. Images sourced from Google Earth at St. James's Park, London, UK, showing terrain only, terrain with built form or terrain with primarily vegetation were paired with four sound conditions using recordings from the park (i.e. 'no sound', anthropogenic, mechanical and natural). Perceived realism and preference were evaluated using a survey delivered via the Internet and in a controlled laboratory environment ($N = 199$ total). Analysis using repeated measures ANOVA indicated the interaction of sound and 3D visualizations significantly alters environmental perception both positively and negatively. Sounds and visuals that are congruent receive higher realism and preference ratings while the more incongruent the combination is, the lower the corresponding ratings. The lowest realism and preference ratings are given to visualizations showing terrain only combined with speech. The highest realism ratings overall correspond to visualization with built form combined with speech, and visualizations showing primarily vegetation paired with a birdcall. The absolute highest realism rating was for the visualization with primarily vegetation and some built form paired with speech, while the highest preference ratings correspond to visualizations showing vegetation paired with birdcall or no sound. Aural-visual data collected via the web-based survey was comparable to data collected in the laboratory and overall realism ratings for the Google Earth visualizations were low (e.g. below 3 on a 1–5 likert type scale). The results suggest there is an opportunity to increase experiential authenticity of 3D landscape visualizations with sound.

1. Introduction

Three-dimensional digital visualization of landscapes offers many advantages over conventional methods of representation, particularly when communicating complex spatial arrangements to non-designers (Bishop, 2005; Kwartler, 2005). To date preparers and presenters of 3D visualizations have primarily focused on visual aspects of landscapes, in part based on the dominance of the human visual system (Lange & Bishop, 2005). However, purely visual approaches to landscape experience have been criticized. For example the complex multi-sensory appreciation of individuals for landscapes has been demonstrated (Scott, Carter, Brown, & White, 2009) as has the important impact of sound on the evaluation of outdoor environments (e.g. Anderson, Mulligan, Goodman, & Regen, 1983; Carles, Barrio, & de Lucio, 1999). The research presented here aims to foreground multisensory landscape experience

* Corresponding author at: Mark Lindquist, 2532 Dana Building, School of Natural Resources and Environment, University of Michigan, 440 Church St Ann Arbor, MI 48109-1041, United States.
E-mail addresses: marklin@umich.edu (M. Lindquist), e.lange@sheffield.ac.uk (E. Lange), j.kang@sheffield.ac.uk (J. Kang).

http://dx.doi.org/10.1016/j.landurbplan.2015.12.017

in the design, planning and evaluation of landscape by analysing the impact of sound on the perception of 3D visualizations.

1.1. Soundscape

The impact of sound on the perception of environments is increasingly under scrutiny. Interest by government and policy-makers is growing in this area, particularly in the regulation and abatement of sound in the form of environmental noise from road traffic, aircraft, railway and machinery and their impact on health and safety (Directive 2002/49/EC, 2002). The concept of soundscape as first suggested by Schafer (1977) has developed into an area of research concerned with studying the impact of sound both positively and negatively on an environment and its perception. Soundscapes defined as the "acoustic environment as perceived or experienced and/or understood by a person or people, in context" (ISO 12913-1 2014, 2014). A growing body of knowledge is emerging through empirical study of soundscape in the urban environment, particularly urban plazas (Yang & Kang, 2005a) and green spaces (Irvine et al., 2009; Nilsson & Berglund, 2006).

Landscape architects and planners have recently turned their attention to designing and planning soundscapes. Auditory concepts for soundscape design and planning have been outlined (Hedfors & Berg, 2003) and used to create a toolkit for professionals (Hedfors, 2003). In the area of landscape planning and management there have been advocates for audio design (Brown & Muhar, 2004) as well as auditory planning (Brown, 2004). More recently the soundscape approach has been applied to early stage urban planning (De Coensel et al., 2010) with frameworks for future research and practical needs outlined (Kang, 2010).

1.2. 3D landscape visualization

Landscape visualizations are made up of fundamental landscape elements that are rendered in 3D to approximate the visual qualities of a landscape and usually include terrain, vegetation, built form and water, and can be expanded to incorporate animals (including people) and atmosphere (Ervin, 2001). These elements contribute differently to perceived realism of the landscape being represented, as does their relative distance from an observers view. For example,research has shown that foreground scenes are rated more realistic than middle ground or background scenes at the same level of detail (Lange, 2001). In addition, the inclusion of texture maps on both terrain and built form can greatly increase perceived realism when compared to simple geometry (Appleton & Lovett, 2003; Lange, 2001; Oh, 1994), while the landscape elements that most affect perceived realism have been shown to be built form and vegetation (Bishop & Rohrmann, 2003). Different user characteristics have shown to alter the perception of 3D visualizations: familiarity with a site can alter preference ratings (Lange, Hehl-Lange, & Brewer, 2008) as well as affecting perceived realism of a visualizations (Appleton & Lovett, 2005; Karjalainen & Tyrväinen, 2002; Lange, 2001). Professional background can also alter realism and preference ratings, with built environment and environmental professionals differing from laypeople in their ratings for both realism and preference (Lange et al., 2008; Lange, 2001). In addition, familiarity with 3D graphics in general and experience with 3D graphics in a planning context can both result in higher realism ratings of visualizations (Appleton & Lovett, 2003).

1.3. Empirical soundscape and landscape research

Senses beyond vision can have a significant impact on our perception of and interaction with our environment. Empirical studies have demonstrated a correlation between the intensity of sound pressure level (SPL) as measured in decibels (dB) and subjective evaluation (Yang & Kang, 2005b). Soundscape preference is also related to the meaning of the sound; research has demonstrated a preference for natural over mechanical sounds (Porteous & Mastin, 1985). In addition, preference has been shown to vary by age, cultural background and long-term environmental experience (Yang & Kang, 2003), finding that align with research on noise sensitivity demonstrating considerable individual variation and ability to adapt to noise (Weinstein, 1978).

Landscape perception research has shown that the interactions of audio and visual stimuli have a significant impact on visual and audio–visual responses to both real and photographed settings.In one study participants rated how enhancing or detracting a particular sound was to a landscape both in situ, via photography and description, the results of which indicated sounds congruent with visuals were most enhancing (Anderson et al., 1983). Sound has also shown to influence overall environmental evaluation both negatively and positively, with combination of natural sounds and images preferred over mechanical (Carles et al., 1999; Carles, Bernáldez, & de Lucio, 1992). In addition, the combination of motion and sound have demonstrated to be important for reliable judgment of dynamic landscapes via scenic beauty assessment (Hetherington, Daniel, & Brown, 1993). Aircraft noise in a natural park setting was shown to have a negative impact on responses to scenic beauty, landscape preference, naturalness and solitude (Mace, Bell, Loomis, & Hass, 2003). In another study the presence of any anthropogenic sound was shown to negatively impact landscape preference ratings, while natural sounds had no impact (Benfield, Bell, Troup, & Soderstrom, 2010). Sound has also been shown to alter the subjective perception of tranquil spaces in neuroscience based research using fMRI (Hunter et al., 2010). Through self-reported and physiological measures these earlier studies made valuable research contributions on the effects of sound on perception of real landscapes using photographs and videos. However, by focusing on real landscapes they do not directly inform processes for the evaluation of future landscape change that 3D visualizations offer. A framework has been proposed by the authors for combining sound with 3D landscape visualizations (omitted for BLIND review) and the current study aimed to provide empirical evidence of the contribution to perceived realism and preference evaluations of a computer simulated environment when pairing real sounds with 3D landscape visualizations.

1.4. Research questions and hypotheses

The objectives of this study were to empirically evaluate the effects of sound on realism and preference ratings of virtual landscapes using real sounds with 3D visualizations and to generate hypotheses concerning the impact of different user characteristics on those ratings. Two research questions were addressed in the study ("How do different landscape elements in 3D visualizations (i.e. terrain, vegetation and built form) interact with real sounds to alter perceived realism of, and preference for, the environment being simulated?" and "What user characteristics interact with combined aural–visual stimuli to alter perception of realism and preference for the environment being simulated?") while a third emerged out of the experimental design ("How effective is the Internet for aural–visual data collection compared to a laboratory setting?"). We hypothesized that for realism congruent aural–visual stimuli would result in the highest ratings with the lowest ratings attributed to the lowest level of visual detail and for preference aural–visual combinations with natural sounds would be the most preferred, with the highest preference being for congruent natural combinations, followed by anthropogenic and lowest for mechanical sounds. User characteristics that could influence ratings were identified from the literature for realism (3D graphic experience, 3D graphic familiarity, site familiarity and

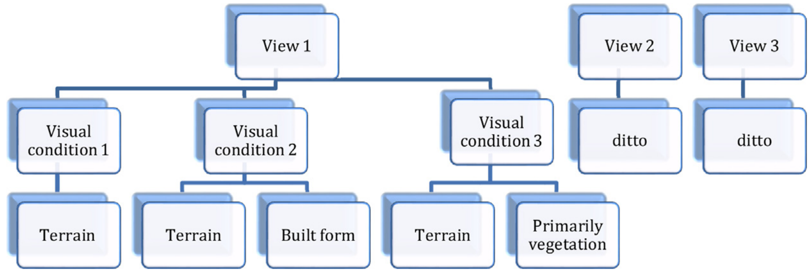

Fig. 1. Experiment image combination, per view.

professional background) and preference (age, cultural background, noise sensitivity and site familiarity). Finally we hypothesized that online results would not differ significantly from laboratory results.

2. Methods

2.1. Participants

A total of 252 respondents participated in the experiment with 199 completing the entire survey (114/57.3% female) aged 18–71 (M = 31.31, SD 13.19). Previous studies examining the perception of aural–visual stimuli in a lab setting have used as few as 20 participants (e.g. Hong & Jeon, 2014) though the number of overall participants is similar to previous studies that evaluated web-based landscape preference data collection (e.g. Wherrett, 2000). In our study 71 participated in the laboratory-based experiment; 128 online. Laboratory-based participants were recruited by the University subject pool email list and by personal contacts, and were eligible to be entered in a draw to win a £50 voucher. Online participants were recruited via Facebook and the University subject pool email list and offered the opportunity to be entered into a draw to win a £25 voucher. Mean time to complete the survey was 13.37 min with a range from 6.83 to 27.95 min. All research conducted was approved through the appropriate institutional ethics procedures.

2.2. Apparatus and materials

2.2.1. Selection of study area, visual and aural stimuli

One overarching aim of the current research was to assess the perception of visualizations sourced from tools used by a variety of spatial and scientific researchers and practitioners. Google Earth was used as a source for the visual stimuli because it is one of the most ubiquitous 3D visualization tools and is being used in many environmental evaluation contexts (e.g. Schroth et al., 2011) but is lacking in depth perceptual evaluation. As a result the study area selection was informed by physical aspects (e.g. accessibility, suitability of the environment) as well as technological aspects (e.g. landscape elements available in Google Earth). St. James's Park, London, UK was selected for the study site because it had: (a) photorealistic vegetation within the park; (b) photorealistic built form surrounding the park for context; (c) relatively high detail in the terrain image mapping; and (d) rigorous surveys and counts on visitor numbers and user satisfaction. St. James's Park is one of the eight Royal Parks of London located in the City of Westminster. The park

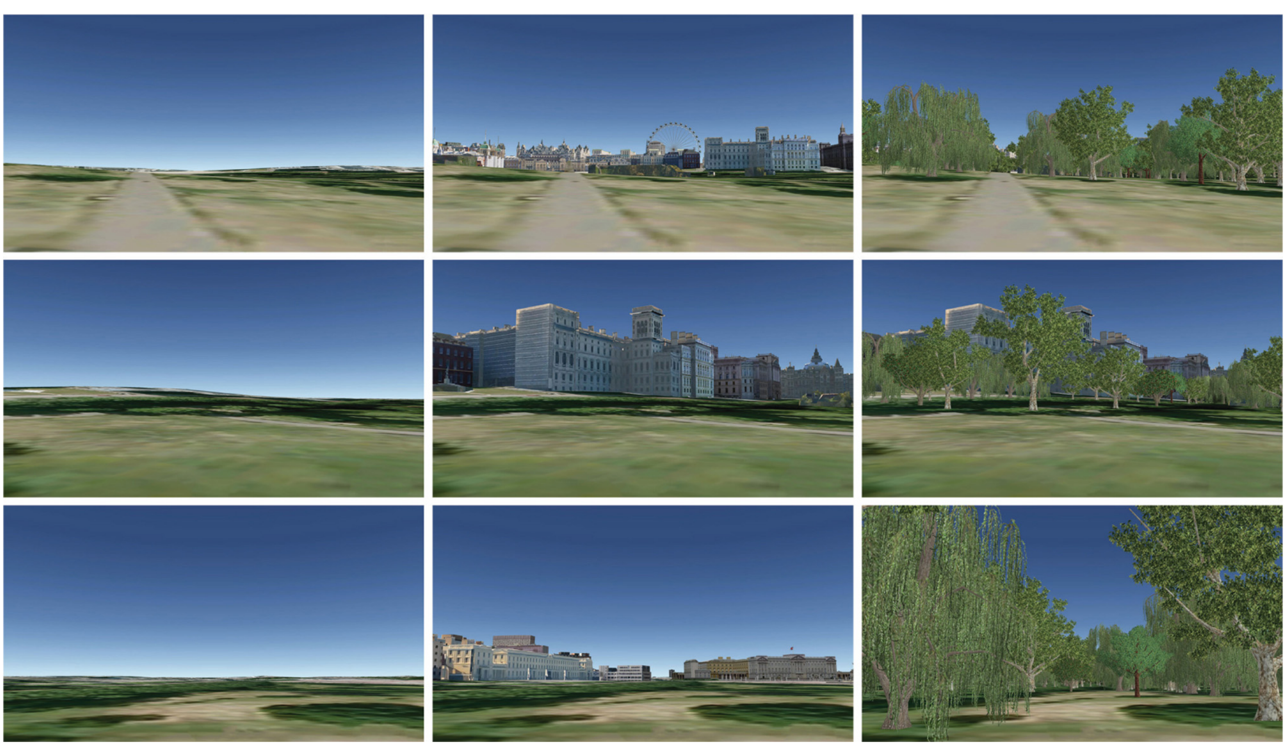

Fig. 2. Views and landscape elements used in the research: view 1 (top row); view 2 (middle row); view 3 (bottom row); by visual condition (1 left column; 2 middle column; 3 right column) (©Google Earth).

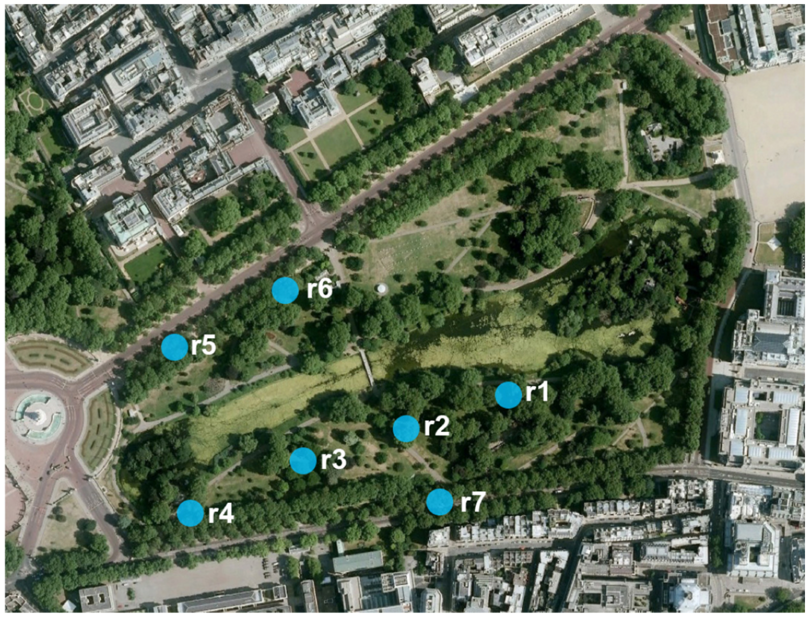

Fig. 3. Recording and measurement locations in St. James's Park.

covers 23 hectares and is bounded by Buckingham palace, The Mall and St. James's Palace, Horse Guards and Birdcage Walk. It includes a small lake with meandering paths and a bridge (SJPaerial 1 kml). St. James's Park was the second most visited of the Royal Parks in 2007 with 6.4 million visitors to Hyde Park's 7.1 million visitors (Hitchcock, Curson, & Parravicini, 2007).

2.2.2. Visual stimuli

The images chosen for the study were drawn from a database of 100 images taken from 3D eye-level views (i.e. foreground, 0–800 m (U.S. Forest Service, 1995)) within Google Earth on the study site. Viewpoints were selected according to the following criteria: (a) showing a representative cross-section of the site; and (b) varying in each image the visual amount of built form, terrain and vegetation (e.g. Appleton & Lovett, 2003). Three views representative of the site were identified (view 1: primarily vegetation with a path; view 2: primarily vegetation with some built form; view 3: primarily vegetation). Three visual representation variables were chosen for this research that have been identified as having a significant impact on perception of visualizations in previous studies: terrain, built form (e.g. Lange, 2001); and foreground vegetation (e.g. Appleton & Lovett, 2003). Each image was exported from Google Earth Pro at a 1080 p HDTV ratio, at a 4800 × 2700 pixel .jpg file to be scalable to fill a 1080 p (1920 × 1080 pixel) monitor. The combination of elements contributing to the different visual conditions used in the experiment are illustrated in Fig. 1 and the nine specific views are illustrated in Fig. 2 (view 1 kml, view 2 kml, view 3 kml).

2.2.3. Acoustic stimuli

Four sound conditions were used in the experiment: no sound, anthropogenic sound (human speech), mechanical sound (road traffic), and natural sound (bird call of a Coot *Fulicaatra*) (e.g. Liu, Kang, Behm, & Luo, 2014; Liu, Kang, Luo, & Behm, 2013). The sounds were drawn from 24 recordings recorded in St. James's Park at four times (0700, 1200, 1700 and 2200) over two days (17–18 July 2012) across sixsites (with a seventh site on the periphery of the park also recorded for a comparison of sound level) (Fig. 3). The sites were chosen after an initial walk-through of the entire park and were evenly distributed along a common route.

Sounds were recorded with an Edirol R-44 4-channel portable recorder in hi-fidelity (48 kHz sampling rate, 24-bit resolution) using 1 channel, a mono microphone,with "Low cut" and the limiter switched on to compensate for wind noise. LAeq was simultaneously measured with a 01 dB Solo Sound Level Meter, which was calibrated using a 01 dB Cal 01 SL that played a 94 dB 1000 Hz tone. The calibrator was also used to record a 10 s segment at the beginning of each recording enabling calibration with the HEAD Analyzer ArtemiS 11.0.200 psychoacoustic analysis software (Head Acoustics, 2012). Recordings were between 140 s and 150 s (120 s recording; 10 s calibration;and 5–10 s removing the calibrator and fitting microphone wind guard). The sounds used for the experiment were selected by identifying sounds exhibiting the most extreme differences by analysis of Leq, Lmax and four psychoacoustic variables of sharpness, fluctuation strength, loudness and roughness (Kang, 2007) and selecting sound different sound content. The intent in selecting the divergent sound types was to provide a temporal soundscape variation approximate to what was uncovered through a soundscape analysis of St. James's Park conducted by one of the authors rather than specifically matching sound and 3D visualization view locations. The location, time and properties of the sounds used in the experiment are shown in Table 1.

Table 1
Location, time, acoustic and psychoacoustic properties of the 3 sounds used in the experiment.

Loc	Time	Leq	Lmin	Lmax	StdDev	Sharpness (acum)	Fluctuation (vacil)	Loudness (sone)	Roughness (asper)	Content
4	700	60.8	55.3	64.6	2	2.39	0.009	37.1	3.32	mechanical
2	1200	56.1	54.3	58.7	0.9	1.99	0.013	30.5	2.86	anthropogenic
6	2200	50.9	49.7	52.2	0.4	1.62	0.007	23.2	2.43	natural

Table 2
Online experiment participant hardware (Display; Audio device).

Display size	Freq	(%)	Audio device	Freq	(%)
Monitor: 13″–21”	93	72.7	Built-in laptop speakers	37	28.9
Monitor: 22”–27”	21	16.4	Earbuds	22	17.2
Unsure	5	3.9	In-ear monitors	20	15.6
Monitor: less than 13”	4	3.1	Desktop computer speakers	17	13.3
iPad/tablet	3	2.3	On-ear headphones	10	7.8
Monitor: larger than 27”	2	1.6	Built-in monitor speakers/soundbar	9	7.0
Total	128	100	Over-ear headphones	9	7.0
			High quality speakers	2	1.6
			Other	2	1.6
			Total	128	100

2.2.4. Questionnaire design

In the main section of the survey respondents were asked to indicate their ratings for realism and preference on a 5 item likert-type scale (“How realistic is your experience of this environment?”—1: not, 2: slightly, 3: somewhat, 4: quite, 5: very; “How much do you like this environment?”—1: not at all, 2: a little, 3: moderately, 4: quite a bit, 5: very much). Verbal labels were assigned following guidelines by Rohrmann (2007) for evenly spaced linguistic separation. Site familiarity was gaugedas recommended by Gale, Golledge, Halperin, & Couclelis (1990) using a two value rating(familiarity and frequency of visits) with respondents shown both an aerial and ground level photograph of St. James’s Park and asked two questions: (1) Are you familiar with St. James’s Park?; and (2) Approximately how often do you visit St. James’s Park?. Noise sensitivity was assessed using a validated 5-item survey developed by Benfield et al. (2012) and direct questions asked to assess 3D graphics familiarity, experience with 3D graphics in a design or planning process, professional background and cultural background.

2.2.5. Apparatus

The laboratory-based part of the experiment used a dual workstation setup. Each workstation was identical except for the audio playback hardware: colour calibrated Dell Ultra sharp 2209WA monitor connected to a PC with the monitor and viewing environment adhering to the ISO 3664 2009 specifications with ambient light at each workstation set to 50 lux (+/−0.7) and light output of each monitor measured at 155 (+/− 1) lumens (ISO 3664 2009, 2009). The resulting on-screen stimulus image size was 31.04 cm width × 17.46 cm height (880 × 495 pixels). Workstation 1 used ‘high quality’ audio hardware (Sennheiser HD 598 over-ear headphones); workstation 2 used ‘low quality’ audio hardware (first generation Apple earbuds with no remote or mic). SPL of the room was 34.1 dB(A) and the SPL from each set of audio hardware were matched using a Neumann KU 100 Dummy Head to measure and set playback level.

The apparatus used for the online study was self-reported by each participant during the demographic part of the online survey for display size and audio device (Table 2). The experiment was delivered via web browser using an online questionnaire (SurveyGizmo, 2012) with an example of the on screen interface shown in Fig. 4.

2.3. Experimental procedure

The overall flow of the survey is illustrated in Fig. 5. In the laboratory condition participants read an information sheet and signed a consent form, then were randomly assigned to either the high or low quality headphone condition by selecting one of two pieces of paper. All participants: (a) answered preliminary user

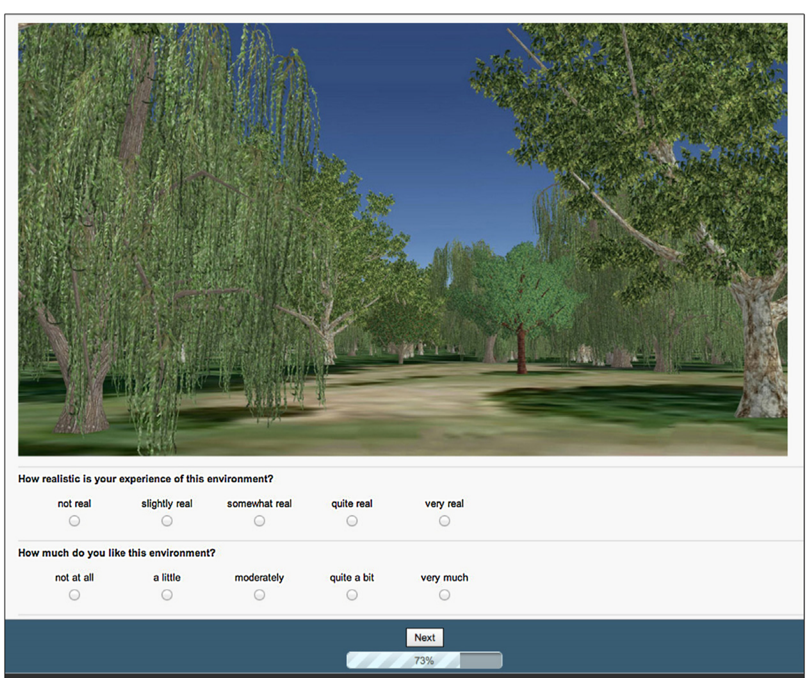

Fig. 4. Example of the online survey.

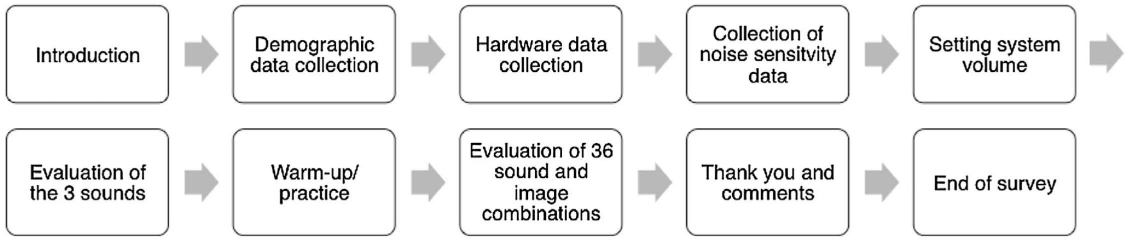

Fig. 5. Flow diagram of the web questionnaire.

characteristic questions; (b) rated the three sounds for the perceived level of loudness and their preference for that sound; (c) were presented with 4 sample sets of stimuli (two image/sound conditions, two image only conditions) to familiarize themselves with the procedure (Lange et al., 2008; Reips, 2002); and (d) were then presented with the aural–visual combinations. The 36 different aural–visual combinations were randomly assigned to participants according to a 3 (visual) × 3 (view) × 4 (sound) factorial design. Sounds were played for 8 s duration, followed by participants indicating their perceived realism and preference.Online participants followed the above procedure with the exception of: (1) The consent form was displayed and agreed to on screen; (2) Being instructed to maximize their browser window; (3) Indicating their monitor size and audio hardware; (5) Being instructed to use a sample audio clip to adjusting their volume to a comfortable level. A pilot study revealed potential confusion from participants relating to terminology of headphone types as well as when asked to set their volume therefore visual cues were added to those questions for assistance. The online survey can be accessed at https://www.surveygizmo.com/s3/2314737/VSS-p1-p56-ONLINE-LAUP.

To reduce dropout and the negative impact of dropout on the survey data six measures were used that were adapted from Roth (2006) which are presented in Table 3. Inclusion criteria were that participants had to exhibit attentiveness by not having an excessively long duration on the main survey questions (i.e. less than 30 s for the 36 sound/image combinations).

2.4. Statistical analysis

The analysis was completed in three parts. Part one employed a mixed repeated measures ANOVA to evaluate the effect of experimental condition and audio visual hardware on responses with between subject factors being “experiment condition” (2 levels), “audio device” (2 levels), and “display size” (2 levels). Part two employed a repeated measures ANOVA to evaluate the effect of the within subject independent variables on realism and preference ratings, as well as their interactions (within subjects factors were “view” (3 levels), “visual condition” (3 levels) and “sound type” (4 levels)). Part three used a mixed repeated measures ANOVA to inform hypothesis generation for future research areas. Post-hoc tests were used to determine if any significant differences existed between groups in the mixed ANOVA with a Bonferroni correction used to control the error rate. There is some controversy surrounding effect size for ANOVAs as there is evidence that eta squared has been misreported in the past in communications research (Levine & Hullett, 2002) which is likely to be the case for landscape preference studies. Alternatives effect sizes have been developed (e.g. Bakeman, 2005; Olejnik & Algina, 2003) however not for the current studies design. As a result, effect size is reported here in partial eta squared ($\eta_p{}^2$) as this is consistent with previous landscape preference research and recommended when other methods are not feasible (Lakens, 2013), which allows for comparison with studies that use a similar design. Analysis was conducted in SPSS 21 (version 21.0.0.2). The data was assessed for normality by inspection of absolute values for skew (<2.0) and kurtosis (<4.0) (West, Finch, & Curran, 1995). No substantial departure from normality was indicated for the data. Violations of sphericity are controlled using the Greenhouse-Geisser correction value, which is included in all following tables when the sphericity test is significant.Additionally the mixed ANOVA were analysed for homogeneity of variance using Levene’s test (Levene, 1960), finding that there was homogeneity of variance for the majority of responses for question research 2 (479/576, 83.2%) and research question 3 (204/216, 94.4%). This was deemed acceptable as ANOVA is robust against reasonable violations of variance, especially when group sizes are relatively equal (Howell, 2012).

3. Results

3.1. Online vs. laboratory participation

In order to determine the suitability of the Internet for aural–visual experiments 3 factors were analysed: experimental condition, audio hardware and display size. The analysis focused on the main effect of each between-subjects factor, and the interaction of the between-subjects factor with each independent variable of view, visual condition and sound. For experimental condition the combined data were analysed for differences occurring between laboratory ($N = 71$) and online participants ($N = 128$) revealing no significant main effects or interactions ($p > .05$ for all). In the laboratory condition participants were randomly assigned to either the high quality ($N = 37$) or low quality ($N = 34$) audio hardware apparatus, the analysis of which revealed no significant main effects or interactions ($p > 0.05$ for all). The combined data ($N = 199$) was analysed to determine any effect of differing video display hardware size on results. All laboratory-based participants used a 22” monitor, while online participants indicated the display size as part of the survey (Table 2). The frequency of display sizes used in the experiment is illustrated in Fig. 6.

The majority of participants used a monitor between 13” and 21” (93/199 or 46.7%) followed closely by a monitor between 22” and 27” (92/199 or 46.2%). As a result the analysis was limited to

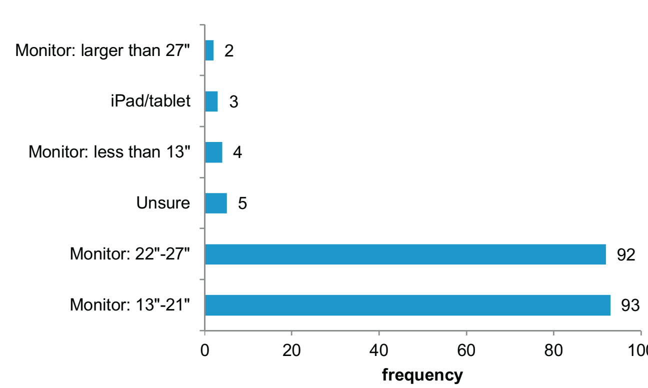

Fig. 6. Combined online and laboratory-based display size frequency.

Table 3
Measures taken to reduce dropout and to reduce the negative impact of dropout (adapted from Roth, 2006).

Measure	Description
High-hurdle technique	The demographic data (personalization) were collected before the evaluation of the stimuli
Warm-up technique	The collection of personal data and practicing stimulus rating before the real experiment start ensures that the data collected in the experimental phase comes from the mostly highly committed participants
Incentive	Participants were given the opportunity to be entered in a draw for a gift certificate (£25 online, £50 laboratory participation)
No plug-ins	No plug-ins are needed for the user's PC, the survey works with all modern web browsers.
Two-item-one-screen design	Each rating takes place on a separate web page. The results are transferred and saved to database immediately after clicking the submit button. If the participant drops out, the former results and the point of time of dropout can be examined
Record of response time per page	The response time is recorded for each web page/each rating. If data quality suffers from interruptions of the experiment, this can be identified

those two variables as there were not enough responses in the other categories to meet the assumption of the mixed ANOVA. The main effect and interactions of display size on realism and preference ratings was not significant ($p > 0.05$ for all) apart from the interaction of display size by view for preference: $F(2,362) = 4.71$, $p = 0.010$, $\eta_p^2 = 0.025$. Visual inspection indicated that participants with displays between 13" and 21" rated all views similarly, while those with displays between 22" and 27" rated preference significantly higher for views 1 and 2 than those with monitors between 13" and 21". This manifested in a mean differences of 0.09 on the 5-point rating scale and a very small effect. Because the effect was small it was concluded that the independent variables in this instance did not dramatically affect results, and further analysis was conducted on the combined sample.

3.2. Mean realism and preference scores by all participants

Participants rated the three sounds without visuals for perceived loudness and preference at the beginning of the survey. During incremental prototyping participants were asked to rate the sounds for realism and preference, however, many indicated some confusion being asked to rate the sound for realism so this was omitted from the main survey. The responses are shown in Table 4 and were in line with previous studies indicating higher preference for natural sounds over anthropogenic sound, with mechanical sounds least preferred. Interestingly the sound content had an effect on perceived loudness as all sounds used in the experiment were matched for loudness.

Table 5 presents all means and standard deviations for realism and preference ratings for all 36 combinations of view, visual condition and sound, as well as the correlation between realism and preference ratings. The overall results can be summarized as follows:

The lowest realism and preference ratings both correspond to visual condition 1 (terrain only) + speech with mean scores for realism ranging from 1.33 to 1.42 and for preference from 1.51 to 1.54.

The highest realism ratings consistently correspond to two combination: visual condition 2 (built form) + speech; and visual condition 3 (primarily vegetation) + natural sound,with mean scores ranging from 2.62 to 2.82. One notable exception is view 2 + visual condition 3 + speech, which happens to have the highest mean score for realism at 2.85 while the same combination for view 3 scores 2.26. This can possibly be explained by the small amount of built form visible behind the vegetation that when viewed with the speech sound indicates the presence of humans somewhere out of the scene.

The highest preference ratings consistently correspond to two combinations: visual condition 3 (primarily vegetation) + no sound; and visual condition 3 + natural sound, with mean scores ranging from 2.95 to 3.13 for both combinations. Finally, realism and preference scores were positively correlated for all 36 combinations with r ranging from 0.399 to 0.832, all significant at $p < 0.01$.

Table 4
Mean loudness and preference ratings of sounds used in the experiment.

Source	Loudness			Preference		
	Mean	SD	*N*	Mean	SD	*N*
Traffic	2.68	0.95	199	2.02	1.02	199
Speech	2.42	0.98	199	2.24	0.93	199
Nature	2.09	0.87	199	2.97	1.04	199

3.3. Analysis of variance of mean realism and preference scores

The means were analysed via a three factor within respondents ANOVA with the factors defined as "view" (3 levels); visual condition (3 levels) and sound (4 levels). The overall results for realism and preference are shown in Table 6. All main effects and interactions were significant ($p < 0.05$ for all) with visual condition having the largest effect for realism and preference as well as the interaction having a relatively strong effect. The data was analysed further by isolating each visual condition by the level of the variables "view" and "sound", separately for each dependent variable "realism" and "preference", with post hoc contrasts used to identify significant effects. Fig. 7 illustrates the mean realism and preference ratings by view, visual condition and sound type, with the full ANOVA and contrast tables for realism shown in Table 7 and preference in Table 8.

3.3.1. ANOVA and contrasts for realism scores

ANOVA and contrasts for realism are shown in Table 7 analysed separately for each visual condition. For visual condition 1 the main effect of 'sound' was significant ($p < 0.001$) though not the main effect of 'view' ($p < 0.307$) or the interaction ($p < 0.301$) therefore the data was collapsed across view with mean realism ratings computed for each level of the independent variable sound. Contrasts revealed significant differences between 4 of the 6 sound conditions ($p < 0.001$, Table 7), the exception being 'no sound vs. traffic' ($p < 0.179$) and 'traffic vs. nature' ($p < 0.82$). The relationship is illustrated in Fig. 7, showing that 'speech' receives the lowest overall realism rating (M = 1.40, SD = 0.62) which was significantly lower than all the other sound conditions (no sound: $M = 1.80$, SD = 0.95, $p < 0.001$; traffic: $M = 1.93$, SD = 0.91, $p = 0.001$; nature: $M = 2.01$, SD = 0.90, $p < 0.001$), with the largest difference between 'speech' and 'nature' ($\eta_p^2 = 0.378$).

For visual condition 2 the main effects of 'view' and 'sound' were significant ($p < 0.001$ for both) though not the interaction ($p < 0.684$). Contrasts (Table 7) revealed that realism ratings did not differ significantly between views 1 and 2 ($p < 0.132$) but differed significantly between views 1 and 3 ($p < 0.001$) and views 2 and 3 ($p < 0.027$), therefore view 3 was considered in isolation and realism ratings for views 1 and 2 merged. The effect of 'view' is explained because view 3 was rated significantly lower than views

Table 5
Mean realism and preference ratings for each combination of view, visual condition and sound type with correlations between realism and preference.

View	Vis Cond	Sound	Realism		Preference		Pearson Correlations
			Mean	SD	Mean	SD	
1	1	No sound	1.73	0.99	1.87	1.08	0.787**
		Traffic	1.94	1.06	1.72	0.91	0.669**
		Speech	1.41	0.73	1.54	0.77	0.444**
		Nature	1.98	0.97	2.15	1.05	0.731**
	2	No sound	2.37	0.99	2.62	1.02	0.649**
		Traffic	2.51	1.04	2.25	1.01	0.480**
		Speech	2.82	1.11	2.56	1.00	0.509**
		Nature	2.41	1.06	2.63	1.03	0.673**
	3	No sound	2.57	1.12	3.01	1.10	0.670**
		Traffic	2.43	1.07	2.39	0.97	0.472**
		Speech	2.39	1.07	2.36	0.92	0.550**
		Nature	2.76	1.15	3.07	1.07	0.635**
2	1	No sound	1.84	1.03	2.06	1.11	0.789**
		Traffic	1.94	1.03	1.89	0.99	0.620**
		Speech	1.42	0.76	1.56	0.78	0.518**
		Nature	2.02	1.03	2.29	1.12	0.719**
	2	No sound	2.29	1.04	2.44	1.00	0.651**
		Traffic	2.51	0.99	2.11	0.93	0.399**
		Speech	2.76	1.17	2.49	0.95	0.542**
		Nature	2.26	1.03	2.45	1.02	0.680**
	3	No sound	2.63	1.04	2.95	1.03	0.600**
		Traffic	2.59	1.09	2.44	0.97	0.423**
		Speech	2.85	1.07	2.72	1.01	0.563**
		Nature	2.75	1.07	3.00	1.04	0.619**
3	1	No sound	1.8	1.04	1.97	1.08	0.832**
		Traffic	1.89	1.08	1.83	1.03	0.677**
		Speech	1.33	0.68	1.51	0.75	0.459**
		Nature	2.02	1.02	2.23	1.14	0.779**
	2	No sound	2.22	0.97	2.22	0.90	0.613**
		Traffic	2.41	0.97	2.01	0.85	0.438**
		Speech	2.62	1.08	2.29	0.84	0.510**
		Nature	2.19	0.97	2.28	0.99	0.612**
	3	No sound	2.63	1.12	3.13	1.07	0.615**
		Traffic	2.24	1.10	2.24	1.01	0.557**
		Speech	2.26	1.10	2.35	0.97	0.529**
		Nature	2.81	1.10	3.06	1.06	0.693**

Table 6
ANOVA results for realism and preference, all participants.

Source	SS	df	MS	*F*	Sig.	η_p^2	ε[a]
Realism							
Main effects							
view	17.45	2, 394	8.72	17.14	**<0.001**	0.080	
visual condition	874.35	1.56, 307.72	559.75	126.93	**<0.001**	0.392	0.781
sound	22.50	2.53, 498.43	8.89	6.35	**0.001**	0.031	0.843
Two-way interactions							
view × vis cond.	15.43	4, 788	3.86	7.49	**<0.001**	0.037	
view × sound	14.33	6, 1182	2.39	6.79	**<0.001**	0.033	
viscond × sound	231.27	5.23, 1029.38	44.26	46.54	**<0.001**	0.191	0.871
Three-way interaction							
view × viscond × sound	19.03	10.31, 2030.98	1.85	4.56	**<0.001**	0.023	0.859
Preference							
Main effects							
view	15.29	2, 392	7.65	13.57	**<0.001**	0.065	
visual condition	840.16	1.53, 299.81	549.26	145.71	**<0.001**	0.426	0.765
sound	291.36	2.74, 536.23	106.50	55.84	**<0.001**	0.222	0.912
Two-way interactions							
view × viscond	33.84	3.74, 732.01	9.06	16.32	**<0.001**	0.077	0.934
view × sound	5.75	5.63, 1103.20	1.02	2.87	**0.011**	0.014	0.938
viscond × sound	131.18	5.57, 1091.61	23.55	37.14	**<0.001**	0.159	0.928
Three-way interaction							
view × viscond × sound	20.90	10.85, 2125.76	1.93	5.30	**<0.001**	0.026	0.904

SS = Type III Sum of Squares df = degrees of freedom; MS = Mean Square *F* = F ratio; Sig. = significance; η_p^2 = partial eta squared (effect size).
[a] A value in this column indicates Mauchly'ssphericity test was significant, and the indicated Greenhouse-Geisser correction applied to the results. Significance (at 0.05) is in **bold**.

1 and 2 across all stimuli combinations for visual condition 2, with mean scores ranging from 0.10 to 0.18 lower, while the effect of each sound condition was consistent across views (e.g. Fig. 7). Contrasts revealed similar significant differences between 5 of the 6 sound conditions for both views ($p < 0.05$, Table 7), the exception being 'no sound vs. nature' condition, which was not significant for either view 1&2 or view 3 ($p = 1.000$, for both). 'Speech' was rated most realistic (e.g. $M = 2.80$, SD = 1.02 views 1&2), significantly

Table 7
ANOVA and contrasts for realism by visual condition.

Source	SS	df	MS	F	Sig.	$\eta_p{}^2$	ε[a]
Visual condition 1							
Main effects							
view	0.85	2, 396	0.42	1.19	0.307	0.006	
sound	132.24	3, 594	44.08	45.5	**<0.001**	0.187	
Interactions							
view × sound	1.93	5.58, 1104.85	0.35	1.21	0.301	0.006	0.93
Contrasts[b]							
no sound vs traffic	3.49	1, 198	3.49	4.79	0.179	0.024	
no sound vs speech	31.89	1, 198	31.89	43.59	**<0.001**	0.18	
no sound vs nature	8.86	1, 198	8.86	15.58	**0.001**	0.073	
trafficvs speech	56.46	1, 198	56.46	83.44	**<0.001**	0.296	
trafficvs nature	1.23	1, 198	1.23	2.23	0.82	0.011	
speechvs nature	74.39	1, 198	74.39	120.56	**<0.001**	0.378	
Visual condition 2							
Main effects							
view	11.54	2, 396	5.77	10.81	**<0.001**	0.052	
sound	78.5	2.85, 563.29	27.59	30.33	**<0.001**	0.133	0.948
Interactions							
view × sound	1.47	5.61, 1110.07	0.26	0.64	0.684	0.003	0.934
Contrasts[b]							
Views							
view 1 vs view 2	1.02	1, 198	1.02	4.11	0.132	0.02	
view 1 vs view 3	5.72	1, 198	5.72	20.67	**<0.001**	0.095	
view 2 vs view 3	1.91	1, 198	1.91	6.95	**0.027**	0.034	
View 1&2 collapsed							
sound	27.67	3, 594	9.22	25.56	**<0.001**	0.114	
no sound vs traffic	6.33	1, 198	6.33	9.58	**0.014**	0.046	
no sound vs speech	42.07	1, 198	42.07	59.01	**<0.001**	0.23	
no sound vs nature	0.01	1, 198	0.01	0.02	1	0	
trafficvs speech	15.76	1, 198	15.76	24.52	**<0.001**	0.11	
trafficvs nature	5.81	1, 198	5.81	7.14	**0.049**	0.04	
speechvs nature	40.7	1, 198	40.7	47.61	**<0.001**	0.19	
View 3							
sound	23.6	2.855	8.27	16.29	**<0.001**	0.076	
no sound vs traffic	7.29	1, 197	7.29	7.46	**0.034**	0.036	
no sound vs speech	31.52	1, 197	31.52	30.82	**<0.001**	0.135	
no sound vs nature	0.18	1, 197	0.18	0.22	1	0.001	
trafficvs speech	8.49	1, 197	8.49	9.58	**0.013**	0.046	
trafficvs nature	9.78	1, 197	9.78	10.81	**0.007**	0.052	
speechvs nature	36.49	1, 197	36.49	30.39	**<0.001**	0.134	
Visual condition 3							
Main effects							
view	20.54	1.93, 380.62	10.63	15.94	**<0.001**	0.075	0.966
sound	41.77	2.76, 543.91	15.13	13.91	**<0.001**	0.066	0.92
Interactions							
view × sound	30.05	5.58, 1098.34	5.39	12.55	**<0.001**	0.06	0.959
Contrasts[b]							
Views							
view 1 vs view 2	5.58	1, 197	5.58	18.19	**<0.001**	0.085	
view 1 vs view 3	0.48	1, 197	0.48	1.72	0.191	0.009	
view 2 vs view 3	9.33	1, 197	9.34	24.56	**<0.001**	0.111	
View 1							
sound	16.5	2.88, 567.08	5.73	9.58	**<0.001**	0.046	0.96
no sound vs traffic	3.96	1, 197	3.96	2.98	0.516	0.015	
no sound vs speech	6.55	1, 197	6.55	5.25	0.138	0.026	
no sound vs nature	6.91	1, 197	6.91	5.58	0.115	0.028	
trafficvs speech	0.32	1, 197	0.32	0.37	1	0.002	
trafficvs nature	21.34	1, 197	21.34	18.3	**<0.001**	0.085	
speechvs nature	26.91	1, 197	26.91	25.98	**<0.001**	0.117	
View 2							
sound	8.52	2.84, 562.08	3	5.13	**0.002**	0.025	0.946
no sound vs traffic	0.25	1, 197	0.25	0.19	1	0.001	
no sound vs speech	9.78	1, 197	9.78	8.15	**0.029**	0.04	
no sound vs nature	3.16	1, 197	3.16	3.38	0.404	0.017	
trafficvs speech	13.14	1, 197	13.14	14.08	**0.001**	0.067	
trafficvs nature	5.17	1, 197	5.17	4.34	0.199	0.022	
speechvs nature	1.82	1, 197	1.83	1.65	1	0.008	
View 3							
sound	46.76	2.85, 564.95	16.39	23.37	**<0.001**	0.106	0.951
no sound vs traffic	29.94	1, 197	29.94	21.29	**<0.001**	0.098	
no sound vs speech	26.91	1, 197	26.91	15.87	**0.001**	0.075	
no sound vs nature	6.55	1, 197	6.55	5.09	0.151	0.025	
trafficvs speech	0.08	1, 197	0.08	0.07	1	0	
trafficvs nature	64.49	1, 197	64.49	55.6	**<0.001**	0.22	
speechvs nature	60.01	1, 197	60.01	43.62	**<0.001**	0.181	

df = degrees of freedom; F = F ratio; Sig. = significance; $\eta_p{}^2$ = partial eta squared (effect size).

[a] A value in this column indicates Mauchly'ssphericity test was significant, and the indicated Greenhouse-Geisser correction applied to the results. Significance (at 0.05) is in **bold**.

[b] Contrasts adjusted for multiple comparisons with Bonferroni adjustment.

Table 8
ANOVA and contrasts for preference by visual condition.

Source	SS	df	MS	F	Sig.	$\eta_p{}^2$	ε[a]
Visual condition 1							
Main effects							
view	7.327	1.90, 375.21	3.87	10.773	**<0.001**	0.052	
sound	148.017	3, 594	49.34	50.853	**<0.001**	0.204	
Interactions							
view × sound	2.116	5.60, 1108.15	0.378	1.26	0.275	0.006	0.93
Contrasts							
Views							
view 1 vs view 2	3.663	1, 198	3.66	18.92	**<0.001**	0.087	
view 1 vs view 3	0.95	1, 198	0.95	5.10	0.075	0.025	
view 2 vs view 3	0.882	1, 198	0.88	6.77	**0.03**	0.033	
View 1&3 collapsed							
sound	46.081	3, 594	15.36	43.90	**<0.001**	0.181	
no sound vstrafic	3.94	1, 198	3.94	5.38	0.129	0.026	
no sound vs speech	32.161	1, 198	32.16	40.60	**<0.001**	0.17	
no sound vs nature	13.851	1, 198	13.85	22.59	**<0.001**	0.102	
trafficvs speech	13.588	1, 198	13.59	22.34	**<0.001**	0.101	
trafficvs nature	32.564	1, 198	32.56	44.41	**<0.001**	0.183	
speechvs nature	88.222	1, 198	88.22	122.56	**<0.001**	0.382	0.93
View 2							
sound	56.637	3, 594	18.88	34.23	**<0.001**	0.147	
no sound vstrafic	4.829	1, 198	4.83	4.02	0.279	0.02	
no sound vs speech	50.251	1, 198	50.25	44.47	**<0.001**	0.183	
no sound vs nature	10.633	1, 198	10.63	10.56	0.008	0.051	
trafficvs speech	23.925	1, 198	23.93	23.80	**<0.001**	0.107	
trafficvs nature	29.794	1, 198	29.79	27.16	**<0.001**	0.121	
speechvs nature	107.116	1, 198	107.12	91.07	**<0.001**	0.315	
Visual condition 2							
Main effects							
view	39.42	1.90, 371.65	20.79	31.67	**<0.001**	0.139	0.948
sound	44.59	2.86, 561.38	15.57	18.9	**<0.001**	0.088	0.955
Interactions							
view × sound	2.2	6, 1176	0.37	1.11	0.354	0.006	
Contrasts[b]							
Views							
view 1 vs view 2	4.2	1, 196	4.2	13.46	**<0.001**	0.064	
view 1 vs view 3	19.67	1, 196	19.67	52.62	**<0.001**	0.212	
view 2 vs view 3	5.7	1, 196	5.7	22	**<0.001**	0.105	
View 1							
sound	18.63	3, 594	6.21	12.4	**<0.001**	0.059	
no sound vs traffic	27.52	1, 198	27.52	25.9	**<0.001**	0.116	
no sound vs speech	1.29	1, 198	1.29	1.35	1	0.007	
no sound vs nature	0	1, 198	0	0	1	0	
trafficvs speech	16.91	1, 198	16.91	17.89	**<0.001**	0.083	
trafficvs nature	27.52	1, 198	27.52	25.89	**<0.001**	0.116	
speechvs nature	1.29	1, 198	1.29	1.39	1	0.007	
View 2							
sound	17.87	3, 588	5.96	11.31	**<0.001**	0.055	
no sound vs traffic	20.79	1, 196	20.79	18.94	**<0.001**	0.088	
no sound vs speech	0.51	1, 196	0.51	0.44	1	0.002	
no sound vs nature	0.02	1, 196	0.02	0.02	1	0	
trafficvs speech	27.8	1, 196	27.8	30.57	**<0.001**	0.135	
trafficvs nature	22.11	1, 196	22.11	19.71	**<0.001**	0.091	
speechvs nature	0.33	1, 196	0.33	0.29	1	0.001	
View 3							
sound	10.82	2.88, 569.40	3.76	8.52	**<0.001**	0.041	0.959
no sound vs traffic	9.29	1, 198	9.29	10.01	0.011	0.048	
no sound vs speech	1.13	1, 198	1.13	1.62	1	0.008	
no sound vs nature	0.72	1, 198	0.72	0.83	1	0.004	
trafficvs speech	16.91	1, 198	16.91	24.06	**<0.001**	0.108	
trafficvs nature	15.2	1, 198	15.2	16.2	**<0.001**	0.076	
speechvs nature	0.05	1, 198	0.05	0.05	1	0	
Visual condition 3							
Main effects							
view	3.22	1.90, 376.88	1.69	2.51	0.085	0.013	0.952
sound	231.01	2.80, 553.97	82.57	65.51	**<0.001**	0.249	0.933
Interactions							
view*sound	23.14	5.6, 1108.84	4.13	9.99	**<0.001**	0.048	0.933
View 1							
sound	91.46	2.88, 569.16	31.82	44.68	**<0.001**	0.184	0.958
no sound vs traffic	77.27	1, 198	77.27	52.26	**<0.001**	0.209	
no sound vs speech	90.23	1, 198	90.23	56.22	**<0.001**	0.221	
no sound vs nature	0.61	1, 198	0.61	0.54	1	0.003	
trafficvs speech	0.5	1, 198	0.5	0.4	1	0.002	
trafficvs nature	91.58	1, 198	91.58	66.32	**<0.001**	0.251	

Table 8 (*Continued*)

Source	SS	df	MS	F	Sig.	η_p^2	ε[a]
speechvs nature	105.65	1, 198	105.65	78.25	**<0.001**	0.283	
View 2							
sound	38.94	3, 594	12.98	22.41	**<0.001**	0.102	
no sound vs traffic	52.28	1, 198	52.28	47.99	**<0.001**	0.195	
no sound vs speech	10.18	1, 198	10.18	8.3	0.026	0.04	
no sound vs nature	0.41	1, 198	0.41	0.43	1	0.002	
trafficvs speech	16.33	1, 198	16.33	12.9	0.002	0.061	
trafficvs nature	61.92	1, 198	61.92	52.15	**<0.001**	0.208	
speechvs nature	14.65	1, 198	14.65	11.92	0.004	0.057	
View 3							
sound	123.75	2.85, 564.07	43.44	60.09	**<0.001**	0.233	0.95
no sound vs traffic	150.4	1, 198	150.4	88.47	**<0.001**	0.309	
no sound vs speech	120.73	1, 198	120.73	77.04	**<0.001**	0.28	
no sound vs nature	1.29	1, 198	1.29	1.11	1	0.006	
trafficvs speech	1.63	1, 198	1.63	1.25	1	0.006	
trafficvs nature	123.86	1, 198	123.86	92.5	**<0.001**	0.318	
speechvs nature	97.09	1, 198	97.09	82.89	**<0.001**	0.295	

df = degrees of freedom; F = F ratio; Sig. = significance; η_p^2 = partial eta squared (effect size).

[a] A value in this column indicates Mauchly'ssphericity test was significant, and the indicated Greenhouse-Geisser correction applied to the results. Significance (at 0.05) is in **bold**.

[b] Contrasts adjusted for multiple comparisons with Bonferroni adjustment.

higher than 'traffic' (e.g. $M = 2.52$, SD = 0.91, $p < 0.001$, views 1&2) and much higher than 'no sound' (e.g. $M = 2.34$, SD = 0.91, $p < 0.001$ views 1&2) or 'nature' ($M = 2.34$, SD = 0.94, $p < 0.001$ views 1&2), which also had one of the larger effects ($\eta_p^2 = 0.23$). The combinations with 'no sound' and 'nature' resulted in almost identical ratings (e.g. $M = 2.34$, $p = 1$, views 1&2).

For visual condition 3 the ANOVA revealed that the main effects and interactions were all significant ($p < 0.001$ for all). Contrasts (Table 7) revealed that realism ratings differed significantly between views 1 and 2 ($p < 0.001$) and views 2 and 3 ($p < 0.001$) but not views 1 and 3 ($p < 0.101$). Further contrasts were conducted for each view by sound type, which revealed varied significant differences between realism ratings depending on the view (Table 7). All views were rated similarly with 'no sound' ($M = 2.57$–2.63) and'nature' ($M = 2.75$–2.81), which were not significantly different from each other for any view ($p > 0.05$ for all). 'Traffic' and 'speech' had nearly identical ratings for view 1 (traffic $M = 2.43$, SD = 1.06; speech $M = 2.39$, SD 1.06) and view 3 (traffic $M = 2.24$, SD = 1.10; speech $M = 2.26$, SD = 1.09), however for view 2 ratings were higher for traffic ($M = 2.59$, SD = 1.09) and significantly higher for speech ($M = 2.85$, SD = 1.07) (e.g. Fig. 7). As view 2 was the only image in visual condition 3 that contained built form this seems to indicate that the presence of any built form can have an important moderating effect on perceived realism with different sound types.

3.3.2. *ANOVA and contrasts for preference scores*

ANOVA and contrasts for preference are shown in Table 8 analysed separately for each visual condition. For visual condition 1 the main effect of 'view' and 'sound' were significant ($p < 0.001$) though not the interaction ($p < 0.275$). Contrasts (Table 8) revealed that preference ratings differed significantly between views1 and 2 ($p < 0.001$), and views 2 and 3 ($p = 0.030$), while views 1 and 3 did not differ significantly ($p = 0.075$); therefore view 2 was considered in isolation and views 1 and 3 merged. Fig. 7 illustrates the relationship: view 2 was more preferred than views 1 and 3 relatively consistently for all sounds, scoring from 0.03 to 0.14 higher. Contrasts show the effect of the different sounds was consistent for the 4 conditions, with significant differences between preference means for 5 of the 6 sound combinations ($p < 0.05$, Table 8), the exception being the 'no sound vs. traffic' condition, which was not significant for views 1 and 3 ($p = 0.129$) or view 2 ($p = 0.279$). As with realism 'speech' resulted in the lowest preference ratings for visual condition 1 (view 1&3: $M = 1.53$, SD = 0.65; view 2: $M = 1.56$, SD = 0.78).

For visual condition 2 the main effect of 'view' and 'sound' were significant ($p < 0.001$) though not the interaction ($p = 0.354$). Contrasts revealed that preference ratings differed significantly between all views ($p < 0.001$ for all, Table 8); therefore each view was considered independently. Contrasts revealed significant differences between preference ratings for 3 of the 6 sound combinations that involved traffic ($p < 0.001$ for all except view 3, 'no sound' vs. 'traffic' $p = 0.011$), while the remaining three were not significant ($p = 1.000$ for all). There was a clear downward trend in preference ratings between views; view 1 was most preferred (means ranging from 2.25 to 2.62), followed by view 2 (2.11–2.49) and view 3 least preferred (2.01–2.30) which all differed significantly ($p < 0.001$ for all), which could be attributed to the Ferris wheel visible in the background of view 1 and more detail of the building in view 2 compared to view 3. In addition, 'traffic' with visual condition 2 resulted in significantly lower preference scores for all views ($p < 0.05$ for all) compared to the other sounds resulting in mean scores between 0.29 and 0.38 lower.

For visual condition 3 the main effect of view was not significant ($p = 0.085$), while the main effect of sound and the interaction of view and sound was significant ($p < 0.0005$ for both). Contrasts revealed similar effects of sound across views, with 'traffic' and 'speech' both resulting in the lowest preference ratings (ranging from 2.26 to 2.35) and 'no sound' and 'nature resulting in the highest preference ratings (ranging from 2.95 to 3.13). The exception to this trend was for view 2 combined with "speech" which received a mean score significantly higher than the other 'traffic' or 'speech' combinations ($M = 2.72$, SD = 1.02). As was the case with realism the built form visible in the background of view 2 suggests the importance of aural-visual congruence and the impact this can have on preference ratings, in this instance up to 0.46 on a 5-point scale.

3.4. *Hypothesis generation for perceptual variation by user characteristic*

For analysis by user characteristic the aim was to generate hypotheses for future research (e.g. McGuire, 1997). Previous research that focused on perceptual variation by user characteristics drew conclusions primarily on the separate evaluation of landscape and soundscape. Informed by this research specific user characteristics that could influence perception were identified (Table 9). These were examined in relation to mean realism and preference ratings using a mixed ANOVA separately for each characteristic, with the respective user characteristic as the

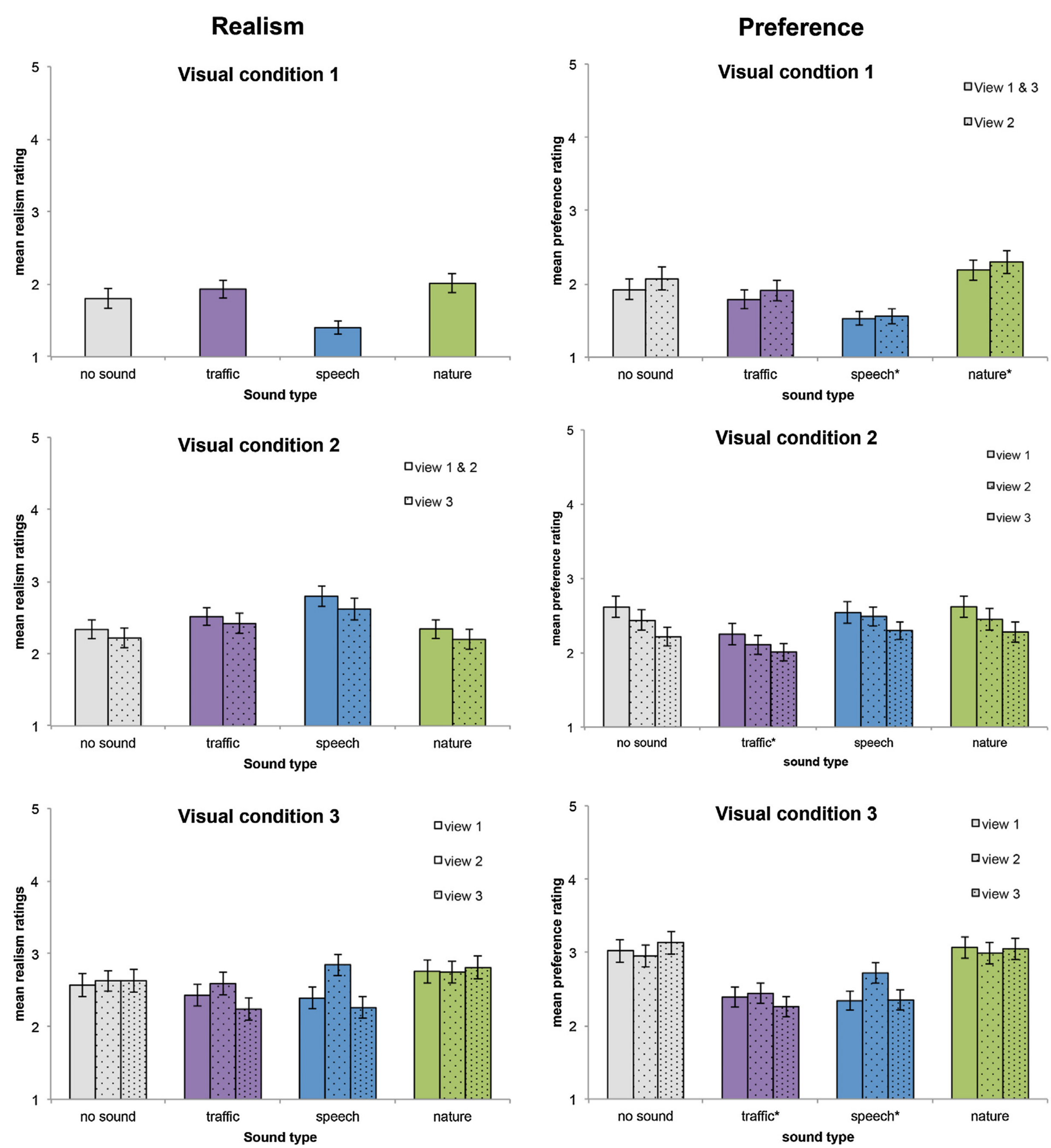

Fig. 7. Mean rating for realism (left) and preference (right) by visual condition: visual condition 1, collapsed across IV of view for realism and views 1&3 for preference (top); visual condition 2, comparing view 1&2 collapsed to view 3 for realism and all view for preference (middle); and visual condition 3, comparing all 3 views for realism and preference. Error bars show 95% confidence intervals.

between-subjects variable and the main effects of view, visual condition and sound as the within subject factors. For site familiarity the responses to the two questions were averaged to provide an aggregate site familiarity score however this resulted in a very large grouping in bottom of responses (e.g. 91.5% of respondents scoring a 1 or 2 out of 5). To satisfy the ANOVA assumptions site familiarity was analysed via 3 groups: moderate to very high familiarity, a little, and no familiarity. Analysis of professional background was going to be conducted across three groups (Landscape, $n = 62$; Built Environment, $n = 16$; and "Other", $n = 94$) however this resulted in very unbalanced groups. As a compromise responses were combined into two categories: Landscape and Built Environment (Architecture, Civil Engineering, Geography, Horticulture, Landscape Architecture and Planning backgrounds) and 'Other', (i.e. expert and layperson). Other characteristics were analysed unaltered.

The analysis revealed a number of small but statistically significant effects that differed based on user characteristics for realism (site familiarity: $F(10.57,1030.12) = 2.31$, $p = 0.009$, partial $\eta^2 = 0.02$; professional background: $F(5.30, 895.02) = 2.23$, $p = 0.046$, partial $\eta^2 = 0.01$) and preference (country: $F(5.58,1088.20) = 2.34$, $p = 0.034$, partial $\eta^2 = 0.01$; professional background: $F(55.57,935.23) = 2.16$, $p = 0.049$, partial $\eta^2 = 0.01$) while the other characteristics did not show any significant effects ($p > 0.05$ for all). Contrasts revealed that the majority of interactions were, while significant, small in all cases (e.g. partial η^2 ranging

Table 9
User characteristic mixed ANOVA groups.

Characteristic [rating type] (group, n)	Total
Site familiarity [realism & preference] (3 groups, n = 198)	
very much/quite a bit/moderately	41 (20.6%)
a little	51 (25.6%)
not at all	106 (53.8%)
3D computer graphics familiarity [realism] (5 groups, n = 199)	
Very much	27 (13.6%)
Quite a bit	36 (18.1%)
Moderately	41 (20.6%)
A little	76 (38.2%)
Not at all	19 (9.5%)
3D graphics experience in design/planning [realism] (2 groups, n = 199)	
No	131 (65.8%)
Yes	68 (34.2%)
Professional background [realism] (2 groups, n = 170)	
Expert	78 (46.8%)
Laypeople	92 (53.2%)
Age groups [preference] (3 groups, n = 193)	
15–24 years	80 (40.2%)
25–44 years	84 (42.2%)
45–64 years	29 (14.6%)
Cultural background [preference] (2 groups, n = 197)	
UK	94 (47.7%)
"Other"	103 (52.3%)
Noise sensitivity [preference] (3 groups, n = 199)	
Low	66 (33.2%)
Medium	109 (54.8%)
High	24 (12.1%)

from 0.05 to 0.09 with means differing by up to 0.26 on the 1–5 scale). The standout was the interaction involving a participant's country of background and preference ratings which indicated that participants who had spent the majority of their life in the UK preferred the combination of visual condition 1 with the natural sound more than non-UK participants ($F(1,195) = 114.82$, $p < 0.001$, partial $\eta^2 = 0.37$) which manifested in a difference of 0.30 between the two groups. This is attributed to UK participants potentially viewing the terrain only visualization as a beach and the bird call originating from a seagull based on anecdotal conversations following the experiment, which is an area for further research. It is worth noting that also the vast majority of significant interaction involved 'speech' as a variable, with the exception being differences by professional background for realism and preference, which was the 'no sound' vs. 'traffic' condition for both. The impact of speech specifically focusing on the inclusion and exclusion of people in visualizations is an area for future research, as is the effect of different types of speech on participants with different cultural backgrounds.

4. Discussion

4.1. Contribution of combining sound and 3D landscape visualizations on the perception of a simulated environment

The results of this study are consistent with the hypotheses that the interactions of aural and visual stimuli have a noticeable and statistically significant effect on realism and preference ratings for a simulated environment, which is primarily influenced by the congruence of the aural and visual stimuli. Fig. 8 illustrates this, showing the positive or negative effect of each sound by visual condition relative to the 'no sound' condition. Overall speech has the largest impact on realism ratings, both positively and negatively, which varies by the visual condition. Combinations that were perceived incongruous have the largest (negative) impact on preference, with anthropogenic and mechanical sounds detracting most from visualizations with predominantly vegetation. This would suggest that "what is expected" has less of a perceptual impact than what is not expected, which aligns with previous research on soundscape preference that indicated that human preference scores correlated with the absence or presence of wanted or unwanted sounds (Lam, Brown, Marafa, & Chau, 2010). This is also consistent with previous research indicating the importance of the congruence of aural and visual stimuli on multi modal perception (e.g. Carles et al., 1999; Zhang & Kang, 2007). While the sounds were recorded at the general locations of the visualizations the majority of participants were not familiar with the site and as such would be responding to the combinations without connecting this to a specific real world location and what they expect to experience there. The results also support previous research that indicates that natural sound has either a positive or small effect on preference of photographs (e.g. Benfield et al., 2010). What is interesting is the moderating effect of any visual anthropogenic indicators (e.g. buildings) on both realism and preference ratings. This is evident in visual condition 3: traffic and speech both reduced perceived realism the most for view 3 (predominantly vegetation) and less for view 1 (mostly vegetation with a building just visible in the distance), while speech actually increased perceived realism for view 2 (showing mostly vegetation with a building clearly visible behind the trees). For preference traffic and speech again are moderated by the imagery containing built form, in a similar pattern to that for realism,

The combination of visual condition 2 (terrain and built form) with the speech sound resulted in the highest overall increase in realism ratings (+.50 compared to no sound) further supporting the importance of congruent stimuli. Even when no people are present in the visualization the presence of anthropogenic visual elements (i.e. buildings) seems to have provided the context for realism. In this instance natural sound did not significantly alter perceived realism, which suggests a stronger influence of anthropogenic sound on visualizations of natural environments than natural sound on visualizations of built environments.Interestingly there is also some preliminary evidence that congruence is a factor even when people perceive sounds and visuals incorrectly as potentially is the case with UK participants potentially hearing a 'seagull' and seeing a 'beach'. This highlights the importance of both context specificity of task in aural–visual environmental simulations.

4.2. Perceived realism of Google Earth

In addition to the aural–visual interactions this research has also identified an important visual result—the relatively low realism attributed to Google Earth eye-level visualizations. It is cautioned that based on the visualizations used in this research the highest level of realism attributed to a Google Earth visualization is 2.81 (on a scale of 5), which places even the most realistic combinations between 'slightly' and 'somewhat' real on the ratings scale. Researchers and professionals need to be aware of this relatively low level of realism if using Google Earth.

4.3. Using the Internet for aural-visual data collection and landscape evaluation

Our study provides preliminary support for using the Internet as a means to engage stakeholders with aural-visual environmental simulation showing no statistically significant differences between online and laboratory results. While there sults are promising and could facilitate testing preference of potentially large groups of stake holders' further research is needed to evaluate differences between perceptual responses to multi sensory computer simulated environments andreal-world, on site experience. Additionally

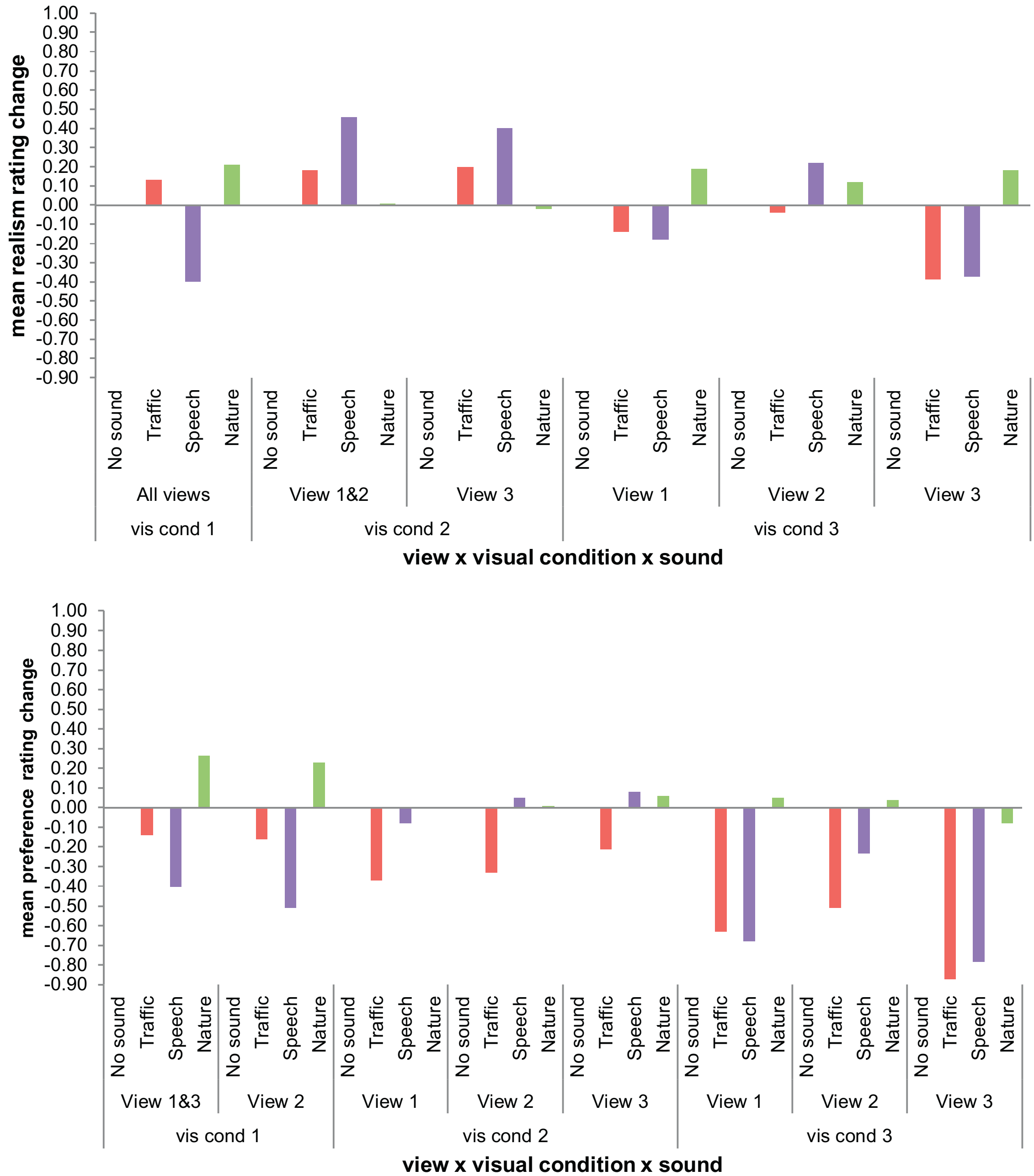

Fig. 8. Mean realism (top) and preference (bottom) rating change relative to the 'no sound' condition.

display size and audio hardware fidelity were shown to not significantly alter realism and preference ratings for the simulated environment. This suggests that when using monitors up to 27" and headphones for sound delivery there is flexibility in the choice of hardware used.

4.4. Limitations of the study

Some important limitations of the study are acknowledged. First, there were a limited number of views used in the experiment, and an uneven operationalization across the views resulting in complex non-linear patterns of results. Second, because of the within-subject design of the study and requiring participants to rate realism and preference together there could be an influence on responses and a relationship between realism and preference occur where it would not were the questions asked separately. Indeed as the realism and preference responses are correlated for each question there is clearly a relationship between them, indicating either that higher realism leads to higher preference, or that participants did not completely distinguish between the two. Some evidence for the former is that visual condition 2 realism scores increased with traffic, while decreasing significantly for preference. This is an area where further research would be beneficial.

4.5. Code of ethics for environmental simulation

As demonstrated in the current research there is obvious potential for misuse of these results, particularly to negatively influence preference for a landscape scenario by using incongruent sounds. This can be addressed largely by preparers and presenters following Sheppard (2001) code of ethics for landscape visualizations. Purposefully using incongruent sounds to illicit negative reactions would be highly unethical, and likely would be called out as a problem by either informed stakeholders or professionals. More problematic is identifying exactly what incongruent sound is—as mentioned, all of the sounds used in the experiment were recorded at St. James's Park and, while not directly matched to visualization location, could be argued to be a valid inclusion with a particular visualization or viewpoint. The selection of view(s), sound(s), seasons and time of day are but a few of the critical decisions to be made, and we propose that *accuracy of experience*be a guiding principle for aural–visual simulations. In the context of ethical conduct preparers and presenters of aural–visual simulations are encouraged to follow Sheppard's code of ethics (2001). In addition they should generally produce simulations that simulate the actual or expected appearance and sound of the landscape and soundscape as closely as possible using spatiotemporally congruent stimuli. This would lend itself to more neutral judgements and transparency in the process.

4.6. Implications for future research

The current study focused specifically on perceptual responses to the combination of aural and visual stimuli in order to inform broad recommendations for aural–visual environmental simulations and to identify future research needed. We suggest the following research is needed to inform aural–visual environmental simulations:

- evaluation of differences between perceptual responses to multisensory computer simulated environments and real-world, on site experience.
- development of procedures to evaluate realism and preference responses to different sounds.
- perceptual effects of different sound sources (e.g. synthesized) with different types of visualizations (e.g. realistic) (see Lindquist, Lange, & Kang, 2014 for an elaboration of options).
- specific techniques of auralization for particular project contexts (e.g. for wind turbines see Pieren, Heutschi, Müller, Manyoky, & Eggenschwiler, 2014).
- assessment of the communication effectiveness of aural–visual simulations.
- the effect of different presentation modes of aural–visual stimuli (e.g. real-time movement vs. static; surround sound vs. mono/stereo).
- the effect of different user characteristics on the perception of aural–visual simulations (specifically on the impact of site familiarity and professional background on perceived realism and cultural and professional background on preference).

4.7. Implications for practice

For landscape architects and planners this signals the importance of considering the total environmental experience beyond how something looks, which can additionally result in positive indirect effects, such as aiding spatial navigation for the visibly impaired by enhancing sound based spatial cues (e.g. Parkin & Smithies, 2012). Preparers and presenters of visualizations need to be conscious that when using sound with visualizations different points of view may vary the landscape elements represented which can alter perceptual responses in unexpected ways. This underscores the importance of multiple viewpoints for evaluation, as well as the potential contribution to multisensory environmental simulation of real-time visualization with accurately modelled spatialized sound. Visualization preparers also need to be conscious of the cross-sensory effect of aural–visual interaction, as a sound that may be obviously one thing to a researcher can be altered by a visualization to be heard as something different by a participant—with potentially unwanted and unexpected results

5. Conclusions

Relying solely on visual representations of environments for design, planning and evaluation does not sufficiently simulate our experience of the world. In this study the contribution of sound to the perception of 3D landscape visualizations was assessed in a laboratory and online experiment. There is a clear opportunity to increase the authenticity of landscape experience when using 3D landscape visualisations by incorporating sound asit significantly alters perception of realism and preference. Eye level Google Earth visualizations receive low realism ratings and aural–visual data collected via a web-based survey was comparable to data collected in the laboratory.Realism and preference vary primarily as a function of aural–visual congruency (e.g. the more congruent the visual landscape element was with the sound the higher the realism and preference was). These findings suggest that coupling the appropriate sound with a corresponding visualization can be an effective way to more accurately simulate environmental experience when using 3D landscape visualization.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.landurbplan.2015.12.017.

References

Anderson, L. M., Mulligan, B. E., Goodman, L. S., & Regen, H. Z. (1983). Effects of sounds on preferences for outdoor settings. *Environment and Behavior*, *15*(5), 539–566. http://dx.doi.org/10.1177/0013916583155001

Appleton, K., & Lovett, A. (2003). GIS-based visualisation of rural landscapes: defining sufficient realism for environmental decision-making. *Landscape and Urban Planning*, *65*(3), 117–131. http://dx.doi.org/10.1016/s0169-2046(02)00245-1

Appleton, K., & Lovett, A. (2005). GIS-based visualisation of development proposals: reactions from planning and related professionals. *Computers, Environment and Urban Systems*, *29*(3), 321–339. http://dx.doi.org/10.1016/j.compenvurbsys.2004.05.005

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. http://dx.doi.org/10.3758/bf03192707

Benfield, J. A., Bell, P. A., Troup, L. J., & Soderstrom, N. C. (2010). Aesthetic and affective effects of vocal and traffic noise on natural landscape assessment. *Journal of Environmental Psychology*, *30*(1), 103–111. http://dx.doi.org/10.1016/j.jenvp.2009.10.002

Benfield, J. A., Nurse, G. A., Jakubowski, R., Gibson, A. W., Taff, B. D., Newman, P., et al. (2012). Testing noise in the field: a brief measure of individual noise sensitivity. *Environment and Behavior*, http://dx.doi.org/10.1177/0013916512454430

Bishop, I. D. (2005). Visualization for participation: the advantages of real-time? In E. Buhmann, P. Paar, I. D. Bishop, & E. Lange (Eds.), *Trends in real-time landscape visualization and participation* (pp. 2–15). Dessau, Germany: Wichmann. Anhalt University of Applied Sciences.

Bishop, I. D., & Rohrmann, B. (2003). Subjective responses to simulated and real environments: a comparison. *Landscape and Urban Planning*, *65*(4), 261–277. http://dx.doi.org/10.1016/s0169-2046(03)00070-7

Brown, A. L. (2004). An approach to soundscape planning. *Acoustics 2004: transportation Noise and vibration: The New Millenium* (pp. 565–570). Retrieved from http://www.acoustics.asn.au/conference_proceedings/AAS2004/ACOUSTIC/PDF/AUTHOR/AC040110.

Brown, A. L., & Muhar, A. (2004). An approach to the acoustic design of outdoor space. *Journal of Environmental Planning and Management*, *47*(6), 827–842. http://dx.doi.org/10.1080/0964056042000284857

Carles, J. L., Barrio, I. L., & de Lucio, J. V. (1999). Sound influence on landscape values. *Landscape and Urban Planning*, *43*(4), 191–200. http://dx.doi.org/10.1016/s0169-2046(98)00112-1

Carles, J. L., Bernáldez, F. G., & de Lucio, J. V. (1992). Audio-visual interactions and soundscape preferences. *Landscape Research*, *17*(2), 52–56. http://dx.doi.org/10.1080/01426399208706361

De Coensel, B., Bockstael, A., Dekoninck, L., Botteldooren, D., Schulte-Fortkamp, B., Kang, J., et al. (2010). *The soundscape approach for early stage urban planning: a case study*. pp. 1–10. Lisbon, Potugal: Proc. Internoise.

Directive 2002/49/EC. (2002). Directive 2002/49/EC of the European parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise. *Official Journal of the European Communities L*, *189*, 12, (18.7.2002).
Ervin, S. M. (2001). Digital landscape modeling and visualization: a research agenda. *Landscape and Urban Planning*, *54*(1–4), 49–62. http://dx.doi.org/10.1016/s0169-2046(01)00125-6
Gale, N., Golledge, R. G., Halperin, W. C., & Couclelis, H. (1990). Exploring spatial familiarity. *The Professional Geographer*, *42*(3), 299–313. http://dx.doi.org/10.1111/j.0033-0124.1990.00299.x
Hedfors, P. (2003). Sonic tools for landscape architecture. In *Conf. proceedings Acoustic Ecology, Australian Forum for Acoustic Ecology/World Forum for Acoustic Ecology* (pp. 19–23). http://www.resonantdesigns.com/proceedings/papers/pHedfors.pdf
Hedfors, P., & Berg, P. G. (2003). The sounds of two landscape settings: auditory concepts for physical planning and design. *Landscape Research*, *28*(3), 245–263. http://dx.doi.org/10.1080/01426390306524
Hetherington, J., Daniel, T. C., & Brown, T. C. (1993). Is motion more important than it sounds?: the medium of presentation in environment perception research. *Journal of Environmental Psychology*, *13*(4), 283–291. http://dx.doi.org/10.1016/S0272-4944(05)80251-8
Hitchcock, M., Curson, T., & Parravicini, P. (2007). *Visitors to the royal parks: results of stead y state count (January–July 2007)*. International Institute for Culture, Tourism and Development, London Metropolitan University.
Hong, J. Y., & Jeon, J. Y. (2014). The effects of audio-visual factors on perceptions of environmental noise barrier performance. *Landscape and Urban Planning*, *125*, 28–37. http://dx.doi.org/10.1016/j.landurbplan.2014.02.001
Howell, D. (2012). *Statistical methods for psychology*. Belmont, CA: Cengage Learning.
Hunter, M. D., Eickhoff, S. B., Pheasant, R. J., Douglas, M. J., Watts, G. R., Farrow, T. F. D., Hyland, D., & Woodruff, P. W. R. (2010). The state of tranquility: subjective perception is shaped by contextual modulation of auditory connectivity. *Neuroimage*, *53*(2), 611–618. http://dx.doi.org/10.1016/j.neuroimage.2010.06.053
Irvine, K. N., Devine-Wright, P., Payne, S. R., Fuller, R. A., Painter, B., et al. (2009). Green space, soundscape and urban sustainability: an interdisciplinary, empirical study. *Local Environment: The International Journal of Justice and Sustainability*, *14*(2), 155–172. http://dx.doi.org/10.1080/13549830802522061
ISO 12913-1 2014. (2014). *ISO 12913-1 Acoustics—soundscape—part 1: definition and conceptual framework*. Geneva, Switzerland: International Standard Organization.
ISO 3664 2009. (2009). *Graphic technology and photography—viewing conditions*. Geneva, Switzerland: International Standard Organization.
Kang, J. (2007). *Urban sound environment*. Abingdon, UK: Taylor & Francis.
Kang, J. (2010). From understanding to designing soundscapes. *Frontiers of Architecture and Civil Engineering in China*, *4*(4), 403–417. http://dx.doi.org/10.1007/s11709-010-0091-5
Karjalainen, E., & Tyrväinen, L. (2002). Visualization in forest landscape preference research: a Finnish perspective. *Landscape and Urban Planning*, *59*(1), 13–28. http://dx.doi.org/10.1016/s0169-2046(01)00244-4
Kwartler, M. (2005). Visualization in support of public participation. In I. D. Bishop, & E. Lange (Eds.), *Visualization in landscape and environmental planning: technology and applications*. London: Taylor &Francis.
Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. http://dx.doi.org/10.3389/fpsyg.2013.00863
Lam, K.-C., Brown, A. L., Marafa, L., & Chau, K.-C. (2010). Human preference for countryside soundscapes. *Acta Acustica United with Acustica*, *96*(3), 463–471. http://dx.doi.org/10.3813/aaa.918299
Lange, E. (2001). The limits of realism: perceptions of virtual landscapes. *Landscape and Urban Planning*, *54*(1-4), 163–182. http://dx.doi.org/10.1016/s0169-2046(01)00134-7
Lange, E., & Bishop, I. D. (2005). Communication, perception and visualization. In I. D. Bishop, & E. Lange (Eds.), *Visualization in landscape and environmental planning: technology and applications* (pp. 3–21). London: Taylor &Francis.
Lange, E., Hehl-Lange, S., & Brewer, M. J. (2008). Scenario-visualization for the assessment of perceived green space qualities at the urban-rural fringe. *Journal of Environmental Management*, *89*(3), 245–256. http://dx.doi.org/10.1016/j.jenvman.2007.01.061
Levene, H. (1960). Robust testes for equality of variances. In I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, & H. B. Mann (Eds.), *Contributions to probability and statistics: essays in honor of harold hotelling* (pp. 278–292). Menlo Park, CA: Stanford University Press.
Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, *28*(4), 612–625. http://dx.doi.org/10.1111/j.1468-2958.2002.tb00828.x
Lindquist, M., Lange, E., & Kang, J. (2014). An assessment of the potential of using visual abstraction & sound for inclusive geodesign abstraction for landscape visualization. In U. Wissen Hayek, P. Fricker, & E. Buhmann (Eds.), *Peer reviewed proceedings of digital landscape architecture 2014 at ETH Zurich* (pp. 170–179). Berlin/Offenbach: Herbert Wichmann Verlag.
Liu, J., Kang, J., Luo, T., & Behm, H. (2013). Landscape effects on soundscape experience in city parks. *The Science of the Total Environment*, *454–455*(0), 474–481. http://dx.doi.org/10.1016/j.scitotenv.2013.03.038
Liu, J., Kang, J., Behm, H., & Luo, T. (2014). Effects of landscape on soundscape perception: soundwalks in city parks. *Landscape and Urban Planning*, *123*(0), 30–40. http://dx.doi.org/10.1016/j.landurbplan.2013.12.003
Mace, B. L., Bell, P. A., Loomis, R. J., & Hass, G. E. (2003). Source attribution of helicopter noise in pristine national park landscapes. *Journal of Park and Recreation Administration*, *21*(3), 97–119. http://js.sagamorepub.com/jpra/article/view/1501
McGuire, W. J. (1997). Creative hypothesis generating in psychology: some useful heuristics. *Annual Review of Psychology*, *48*(1), 1–30. http://dx.doi.org/10.1146/annurev.psych.48.1.1
Nilsson, M. E., & Berglund, B. (2006). Soundscape quality in suburban green areas and city parks. *Acta Acustica United with Acustica*, *92*, 903–911. http://www.ingentaconnect.com/content/dav/aaua/2006/00000092/00000006/art00008
Oh, K. (1994). A perceptual evaluation of computer-based landscape simulations. *Landscape and Urban Planning*, *28*(2–3), 201–216. http://dx.doi.org/10.1016/0169-2046(94) 90008-6
Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447. http://dx.doi.org/10.1037/1082-989x.8.4.434
Parkin, J., & Smithies, N. (2012). Accounting for the needs of blind and visually impaired people in public realm design. *Journal of Urban Design*, *17*(1), 135–149. http://dx.doi.org/10.1080/13574809.2012.646139
Pieren, R., Heutschi, K., Müller, M., Manyoky, M., & Eggenschwiler, K. (2014). Auralization of wind turbine noise: emission synthesis. *Acta Acustica United with Acustica*, *100*(1), 25–33. http://dx.doi.org/10.3813/AAA.918683
Porteous, J. D., & Mastin, J. F. (1985). Soundscape. *Journal of Architectural and Planning Research*, *2*, 169–186.
Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, *49*(4), 243–256. http://dx.doi.org/10.1026//1618-3169.49.4.243
Rohrmann, B. (2007). *Verbal qualifiers for rating scales: sociolinguistic considerations and psychometric data. Project report*. Melbourne, Australia: University of Melbourne. http://www.psych.unimelb.edu.au/staff/br/vqs-report.pdf
Roth, M. (2006). Validating the use of Internet survey techniques in visual landscape assessment: an empirical study from Germany. *Landscape and Urban Planning*, *78*(3), 179–192. http://dx.doi.org/10.1016/j.landurbplan.2005.07.005
Schafer, R. M. (1977). *The soundscape: our sonic environment and the tuning of the world*. New York: Destiny.
Schroth, O., Pond, E., Campbell, C., Cizek, P., Bohus, S., & Sheppard, S. R. J. (2011). Tool or toy? Virtual globes in landscape planning. *Future Internet*, *3*(4), 204–227. http://dx.doi.org/10.3390/fi3040204
Scott, A., Carter, C., Brown, K., & White, V. (2009). Seeing is not everything: exploring the landscape experiences of different publics. *Landscape Research*, *34*(4), 397–424. http://dx.doi.org/10.1080/01426390903009289
Sheppard, S. R. J. (2001). Guidance for crystal ball gazers: developing a code of ethics for landscape visualization. *Landscape and Urban Planning*, *54*(1–4), 183–199. http://dx.doi.org/10.1016/s0169-2046(01)00135-9
SurveyGizmo. (2012). Boulder, Co., USA. Retrieved from https://www.surveygizmo.com/
U.S. Forest Service (1995). *Landscape Aesthetics A Handbook for Scenery Management*. in: U.S. Forest Service, Ed. (Agricultur). U.S. Forest Service. Retrieved from www.fs.fed.us_cdt_carrying_capacity_landscape_aesthetics_handbook_701_no_append.
Weinstein, N. D. (1978). Individual differences in reactions to noise: a longitudinal study in a college dormitory. *Journal of Applied Psychology*, *63*(4), 458–466. http://dx.doi.org/10.1037/0021-9010.63.4.458
West, S., Finch, J., & Curran, P. (1995). Structural equation models with non-normal variables: problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: concepts, issues, and applications* (pp. 57–75). Thousand Oaks, CA: SAGE Publications.
Wherrett, J. R. (2000). Creating landscape preference models using internet survey techniques. *Landscape Research*, *25*(1), 79–96. http://dx.doi.org/10.1080/014263900113181
Yang, W., & Kang, J. (2003). A cross-cultural study of the soundscape in urban open public spaces. In *Proceedings of the 10th International Congress on Sound and Vibration* (pp. 2703–2710).
Yang, W., & Kang, J. (2005a). Acoustic comfort evaluation in urban open public spaces. *Applied Acoustics*, *66*(2), 211–229. http://dx.doi.org/10.1016/j.apacoust.2004.07.011
Yang, W., & Kang, J. (2005b). Soundscape and sound preferences in urban squares: a case study in sheffield. *Journal of Urban Design*, *10*(1), 61–80. http://dx.doi.org/10.1080/13574800500062395
Zhang, M., & Kang, J. (2007). Towards the evaluation, description, and creation of soundscapes in urban open spaces. *Environment and Planning B: Planning and Design*, *34*(1), 68. http://dx.doi.org/10.1068/b31162