The Speed, Precision and Accuracy of Human Multisensory Perception following Changes to the Visual Sense

Sara Elisa Garcia
University College London

Thesis submitted for the degree of
Doctor of Philosophy, March 2016

I, Sara Elisa Garcia, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Human adults can combine information from multiple senses to improve their perceptual judgments. Visual and multisensory experience plays an important role in the development of multisensory integration, however it is unclear to what extent changes in vision impact multisensory processing later in life. In particular, it is not known whether adults account for changes to the relative reliability of their senses, following sensory loss, treatment or training. Using psychophysical methods, this thesis studied the multisensory processing of individuals experiencing changes to the visual sense. Chapters 2 and 3 assessed whether patients implanted with a retinal prosthesis (having been blinded by a retinal degenerative disease) could use this new visual signal with non-visual information to improve their speed or precision on multisensory tasks. Due to large differences between the reliabilities of the visual and non-visual cues, patients were not always able to benefit from the new visual signal. Chapter 4 assessed whether patients with degenerative visual loss adjust the weight given to visual and non-visual cues during audio-visual localization as their relative reliabilities change. Although some patients adjusted their reliance on vision across the visual field in line with predictions based on cue relative reliability, others - patients with visual loss limited to their central visual field only - did not. Chapter 5 assessed whether training with either more reliable or less reliable visual feedback could enable normally sighted adults to overcome an auditory localization bias. Findings suggest that visual information, irrespective of reliability, can be used to overcome at least some non-visual biases. In summary, this thesis documents multisensory changes following changes to the visual sense. The results improve our understanding of adult multisensory plasticity and have implications for successful treatments and rehabilitation following sensory loss.

Contents

Index of Figures and Tables

Acknowledgments

I am immensely grateful to Marko Nardini and Gary Rubin for being outstanding supervisors, providing support, encouragement, guidance and a wealth of knowledge throughout. Thank you also to UCL Impact and the NIHR Biomedical Research Centre for funding this research. I am also very grateful to everyone who helped along the way, particularly Karin Petrini, Lyndon da Cruz, and Michel Michaelides for their technical and conceptual advice. Special thanks to Pete Jones for advice on designing and coding experiments, and for endless discussions about findings and psychophysics. Thanks to all the patients who participated and to Moorfields Eye Hospital.

Published Work

Chapter 3 is published in PLoS One (Garcia, Petrini, Rubin, Da Cruz, & Nardini, 2015); Results from Chapters 2, 3 and 4 have been presented at conferences and published in abstract form (Garcia, Jones, Rubin, & Nardini, 2015; Garcia, Petrini, da Cruz, Rubin, & Nardini, 2014). Chapters 4 and 5 are currently under review.

Chapter 1: Introduction

1.1 Introduction

In daily life, humans rely on multisensory information to perceive the physical world. Furthermore, humans can combine information from multiple senses about the same physical event or stimulus to improve the speed, precision or accuracy of their perceptual judgments. To do so, the human brain requires knowledge about the relationship between different sensory representations. The visual sense is considered fundamental for accurately perceiving space, aligning non-visual spatial representations and, consequently, for developing the ability to combine multisensory information. However, it is unclear to what extent changes to the visual sense later in life, following visual restorative treatment or gradual visual loss, impact multisensory processing.

The first section of this chapter (section 1.2) will describe how multisensory combination can benefit the speed, precision and accuracy of perception. The next sections will review studies that have assessed the impact of blindness on non-visual processing (section 1.3) and multisensory interactions (section 1.4). Next, the ability of visual treatment to restore typical visual, non-visual and multisensory functions will be discussed (section 1.5). Research exploring non-visual and multisensory processing in individuals with partial visual loss will then be reviewed (section 1.6) and finally, an overview of important research areas that this thesis will address will be provided (section 1.7).

1.2 Perceptual Benefits of Multisensory Combination

As many aspects of the physical environment are experienced by more than one sense, integrating information from multiple senses is fundamental to constructing a unified representation of the world. By combining complementary sensory signals, humans can fully characterize their environment. For example, the matching of visual lip movements to sounds disambiguates speech from other noises in the environment. Moreover, multisensory signals can provide the same (‘redundant’) information about a physical event or stimulus, and therefore, in line with signal detection and Bayesian decision theories, humans could use this redundancy to reduce the uncertainty in their behavioural responses. Much research has used psychophysics to study whether humans combine signals according to theoretical ideal observer models, and thereby improve the speed and/or precision of their behavioural responses.

1.2.1 The ideal observer minimizes reaction time

In order to react as quickly as possible to the onset of a physical stimulus, an observer relies on sensory information signalling that the stimulus has been detected. According to signal detection theory, evidence for a physical stimulus (or event) is accumulated over time until a criterion is reached (Green & Swets, 1966; Ratcliff & Smith, 2004; Smith & Ratcliff, 2004). Due to both external (physical) noise and internal (neural) variability, reaction times to a physical stimulus vary each time, (as noisy evidence is accumulated). Consequently, when more than one sensory signal is available, the ideal observer can make use of this redundancy to reduce reaction times (referred to as the redundant signals effect), through probability summation. For example, when an observer is asked to respond to an audio-visual target, the observer has both auditory and visual information to signal the target's presence. Assuming that both sensory signals are processed concurrently in independent channels, the observer can respond based on the sensory channel that finishes processing (or reaches the decision criterion) first (referred to as 'the race model', Raab, 1962). Since the probability of either of two stimuli yielding a fast reaction time is larger than that from either stimulus alone, this leads to a reaction time advantage. Thus, according to probability summation, for every time value (*t*), the observed reaction time distributions should satisfy the 'race model inequality' (Miller, 1982):

$$P\left(RT \le t \middle| S_A \cap S_V\right) = P\left(RT \le t \middle| S_A\right) + P\left(RT \le t \middle| S_V\right) - P\left(RT \le t \middle| S_A \cap RT \le t \middle| S_V\right) \quad \textbf{(1)}$$

where $P\left(RT \le t \middle| S_A\right)$ and $P\left(RT \le t \middle| S_V\right)$ are the auditory and visual cumulative probabilities (respectively) that a response with latency $RT \le t$ has been triggered. Assuming statistically independent response latencies of auditory and visual signals:

$$P\left(RT \le t \middle| S_A \cap RT \le t \middle| S_V\right) = P\left(RT \le t \middle| S_A\right) \times P\left(RT \le t \middle| S_V\right) \quad \textbf{(2)}$$

Since $P\left(RT \leq t \middle| S_A \cap RT \leq t \middle| S_V\right) \geq 0$, Miller's inequality reduces to:

$$P\left(RT \leq t \middle| S_A \cap S_V\right) \leq P\left(RT \leq t \middle| S_A\right) + P\left(RT \leq t \middle| S_V\right) \tag{3}$$

(Note that this is a special case of Boole's inequality: for any finite set of events, the probability that at least one of the events occurs is no greater than the sum of the probabilities of the individual events). Hence, according to the race model, the ideal observer's bimodal reaction time cumulative density function (CDF) is given by the sum of the single modality CDFs.

Interestingly, however, research has frequently reported redundancy gains that exceed those predicted by probability summation, i.e. gains that violate the race model inequality (e.g. Forster, Cavina-Pratesi, Aglioti, & Berlucchi, 2002; Girard, Collignon, & Lepore, 2011; Gondan, Lange, Rosler, & Roder, 2004; Mahoney, Li, Oh-Park, Verghese, & Holtzer, 2011; Miller, 1982; Molholm et al., 2002; Schroger & Widmann, 1998). According to Miller's inequality (Eq. 1), bimodal reaction time distributions cannot exceed the sum of single modality distributions where there are no interactions between signals. Therefore, violations of the model have led to the suggestion that multisensory reaction times that exceed race model predictions reflect integration of sensory signals, whereby evidence for both sensory signals is pooled together to reach a decision, and it is this combination that triggers faster responses ('the coactivation model', Miller, 1982). Instead, Otto and Mamassian (2012) have suggested that violations may be accounted for by trial history effects that invalidate the assumption of statistically independent reaction times to single trials. Specifically, they found response latencies were dependent on the signal presented on the previous trial, with faster responses recorded on trials preceded by the same stimulus (e.g. faster responses to auditory signals following an auditory signal trial). Observed reaction times were well-predicted by a model that accounted for the correlation between latencies in single conditions, but that also assumed increased noise in bimodal conditions. The authors suggested that, assuming evidence accumulation is achieved by increased neuronal firing over time, on bimodal trials two pools of neurons may be necessary to accumulate evidence for distinct signals separately, which could lead to increased internal noise. Hence, violations of the race model indicate that human observers use information from different sensory modalities to react as quickly as possible to stimuli. However, exactly how such sensory signals are processed – whether independently or co-actively – is not yet known.

1.2.2 The ideal observer minimizes uncertainty

Similar to the variability in reaction times, variability in the sensory estimate of an environmental property, due to external (physical noise) and internal (neural) variability, exists too. As a result, when measuring the same environmental property multiple times, the estimate provided by a particular sensory modality will vary slightly each time. Again, it is due to this uncertainty that combining redundant information from multiple sensory modalities can be advantageous.

As an example, consider an observer using visual and haptic systems to provide an estimate of the size of a ball. The probability of the observer estimating the ball to have a particular size X, given visual (V) and haptic (H) cues, is described by the posterior probability distribution, $p\left(X\middle|V,H\right)$. This can be computed using Bayes theorem:

$$p\left(X\middle|V,H\right)=\frac{p\left(V,H\middle|X\right)p\left(X\right)}{p\left(V,H\right)} \tag{4}$$

where $p\left(V,H\middle|X\right)$ is the likelihood of sensing the estimates $\left(V,H\right)$ given that X is true, and $p\left(X\right)$ is the prior probability of different values of X. Assuming that the noise sources in the visual and haptic systems are statistically independent, the likelihood function $p\left(V,H\middle|X\right)$ can be defined as the product of the independent visual and haptic likelihood functions:

$$p\left(V,H\middle|X\right)=p\left(V\middle|X\right)p\left(H\middle|X\right) \tag{5}$$

Therefore, the posterior probability distribution is proportional to the product of the likelihood functions associated with each cue and the prior probability function. The ideal Bayesian observer maximizes the posterior probability to form the best estimate, referred to as the maximum a posteriori (MAP) estimate. Assuming Gaussian likelihood functions and a uniform prior over X (in addition to independent noise) the MAP estimate is the sum of the visual and haptic cues each weighted by its reliability (the inverse of variance). In this case, because the prior is assumed uniform, the MAP estimate is equivalent to the maximum likelihood estimate (MLE):

$$\hat{s}_{VH} = w_V \hat{s}_V + w_H \hat{s}_H \qquad \textbf{(6)}$$

where $\hat{s}_V$ is the visual estimate, $\hat{s}_H$ is the haptic estimate, and w_V and w_H are the relative weights for each modality, inversely proportional to their variances:

$$w_V = \frac{1/\sigma_V^2}{1/\sigma_V^2 + 1/\sigma_H^2} = \frac{\sigma_H^2}{\sigma_V^2 + \sigma_H^2} \qquad \textbf{(7)}$$

The MLE is statistically optimal because it combines single sensory estimates weighted by their relative reliability (Eq. 7) to minimize the variance, thereby producing the 'most reliable' estimate (in that it has the lowest possible variance; see also Fig. 1):

$$\sigma_{VH} = \frac{\sigma_V^2 \sigma_H^2}{\sigma_V^2 + \sigma_H^2} \leq min\left(\sigma_V^2, \sigma_H^2\right) \qquad \textbf{(8)}$$

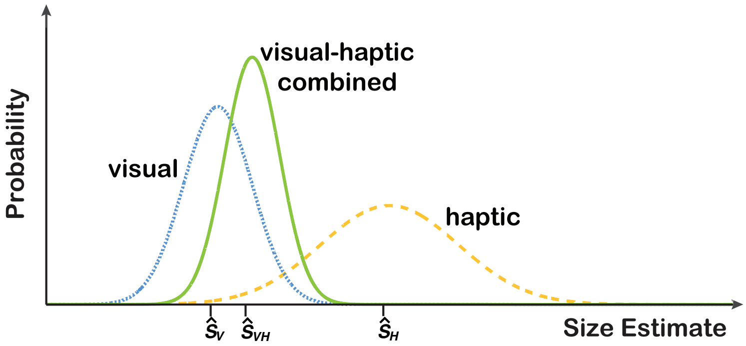


Figure 1: Diagram depicting the likelihood functions of the individual visual and haptic size estimates and the combined estimate, computed according to the MLE (Eq. 6).

To assess whether human adults combine redundant signals to reduce the uncertainty in their final estimate, experimenters have typically used two alternative forced choice paradigms in which participants are asked to make a discrimination judgment using unimodal or bimodal information. For example, Ernst and Banks (2002) asked adult observers to judge which of two successively-presented stimuli (a fixed standard and a variable comparison) was taller, using vision-only, touch-only, or both vision and touch simultaneously. They then plotted the proportion of trials in which the observer indicated that the comparison stimulus (variable height) appeared taller than the standard stimulus (fixed height), as a function of the height of the comparison stimulus, for vision-only, haptic-only and visual-haptic data. Cumulative Gaussian psychometric functions were fit to these data, and used to compute the discrimination threshold (see example Fig. 2.A). Results indicated that adult observers improved the sensitivity of their bimodal judgments, exceeding the sensitivity of the best unimodal cue, and thus providing evidence that adults use information from both sensory cues.

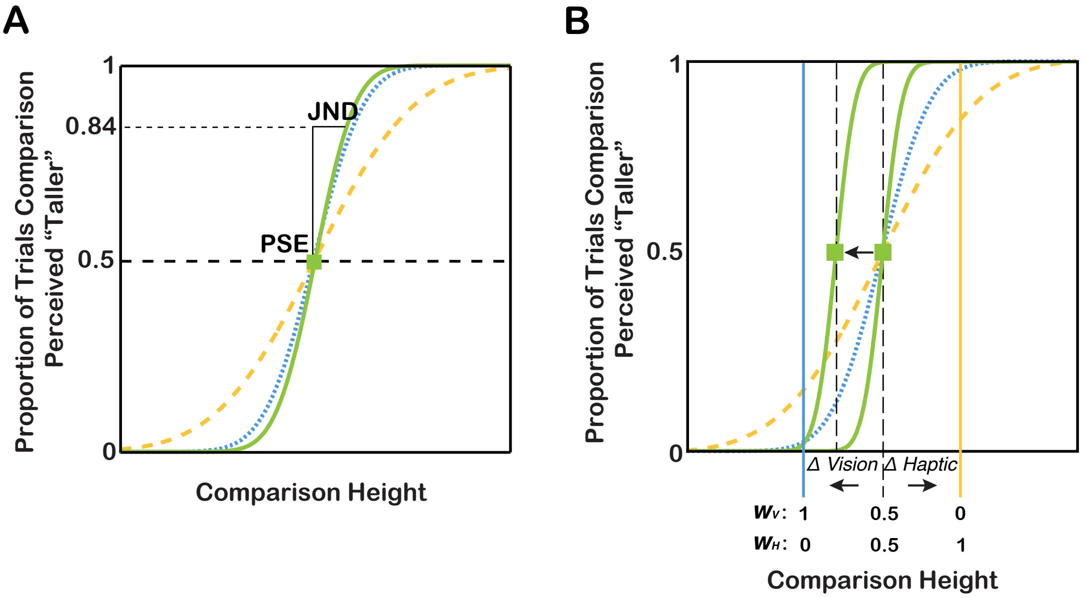


Figure 2: (A) Example of psychometric functions that could be obtained during a 2-AFC task. The just noticeable difference (JND), also referred to as the discrimination or difference threshold, defines the amount by which the standard and comparison stimulus must differ for the observer to differentiate between these ($JND = \sqrt{2}\sigma$). A decrease in the JND reflects an increase in sensitivity. The point of subjective equality (PSE) defines the point at which comparison and standard are perceived as the same. (B) Diagram showing a change in PSE when a conflict is introduced on bimodal trials. In the example, on bimodal trials visual and haptic stimuli are shifted in opposite directions by amount Δ. Since vision is more reliable than audition, reflected by the steeper slope of the psychometric curve, the PSE of the bimodal psychometric curve shifts toward vision.

Similar improved bimodal precision has been shown for other cross-sensory judgments, including: audio-visual spatial localization (Alais & Burr, 2004), visual-haptic shape discrimination (Helbig & Ernst, 2007) and visual-haptic distance estimation (Gepshtein & Banks, 2003). Improvements in estimation precision have also been found when using multiple cues from within the same modality, for example: using both textual and disparity visual cues improves slant estimation precision (Hillis, Watt, Landy, & Banks, 2004; Knill & Saunders, 2003), and using both orientation and spatial frequency or contrast visual cues improves texture edge localization precision (Landy & Kojima, 2001).

To assess whether humans combine cues optimally to minimize uncertainty in their bimodal estimates, in addition to bimodal sensitivity, it is important to also measure the weights given to each sensory cue during bimodal judgments and compare these to ideal observer predictions. This is because combining cues with suboptimal cue weights can also lead to improvements in bimodal precision (albeit reduced relative to the optimal observer), and so by measuring bimodal sensitivity alone, it is difficult to conclude whether humans are combining cues optimally. Alternatively, if measuring cue weighting alone, humans who do not combine cues, but instead make decisions based on one cue, may appear to weight sensory cues optimally if they alternate between cues according to the optimal relative reliability ratio.

Whilst the MLE can be computed from the unisensory data (Eq. 6 & 8), to measure cue weights researchers have devised methods to introduce a minor conflict between the sensory information constituting the standard bimodal stimulus on some bimodal trials. For example, Ernst and Banks (2002) used a random-dot stereogram portraying a bar of specified size as their visual stimulus, whereas the haptic stimulus was generated separately, using two haptic force-feedback devices (presented beneath a mirror displaying the random-dot stereogram reflection from a CRT monitor). This allowed them to introduce a consistent small discrepancy (conflict) between the visual and haptic size information represented by the standard stimulus, on some bimodal trials. This in turn, enabled them to measure how observers weighted each sensory cue during bimodal judgments, by measuring the difference in the point of subjective equality (PSE) on bimodal conflict versus no conflict trials.

For example, in Figure 2.B a conflict was introduced to the standard stimulus on bimodal trials, such that visual and haptic cues to size were shifted in opposite directions by the same amount Δ. Accordingly, if weighting both visual and haptic cues equally ($w_V = 0.5$, $w_H = 0.5$) the size of the standard on conflict trials would correspond to the size of the standard on no conflict trials, and there would be no shift in the point of subjective equality (PSE). If instead the observer relied entirely on vision, the perceived size of the standard on conflict trials would be smaller than that on no conflict trials, and consequently the conflict PSE would shift in the direction of the visual cue by amount Δ. Hence, the difference in the PSE between conflict and no conflict trials provides a measure of how each cue is being weighted. In Figure 2.B, the visual cue is more reliable than the haptic cue, and consequently the PSE for conflict trials has shifted toward vision, but by less than amount Δ, indicating some reliance on haptic information too. If weighting cues optimally, the difference between the conflict and no conflict PSEs ought to reflect the optimal cue weighting.

Additionally, assessing sensory combination for cues with different relative reliabilities directly tests whether humans are indeed weighting cues according to their reliability (as opposed to a fixed ratio that happens to coincide with the experimental optimal weighting). By using a random-dot stereogram to portray the size of a bar as the visual stimulus, Ernst and Banks (2002) were able to vary the reliability of the visual information by adding noise to the depth of the dots. This allowed them to calculate the reliability of the visual information under different noise levels, relative to the haptic information, (which could then be used to compute the optimal visual weighting in bimodal trials, Eq. 7). They found that as the relative visual reliability decreased (due to increased noise), the perceived bimodal estimate was increasingly determined by the haptic size estimate (reflected by a shift in measured PSEs toward the haptic estimate as per predictions).

Although much research has similarly shown that human adults adjust cue weights according to changing reliability (Alais & Burr, 2004; Gepshtein & Banks, 2003; Helbig & Ernst, 2007; Hillis et al., 2004), some studies have shown that cue weights do not always meet optimal model predictions (Battaglia, Jacobs, & Aslin, 2003; J. S. Butler, Smith, Campos, & Bulthoff, 2010; Fetsch, Turner, DeAngelis, & Angelaki, 2009; Knill & Saunders, 2003; Oruc, Maloney, & Landy, 2003; Rosas, Wagemans, Ernst, & Wichmann, 2005). Fetsch et al. (2009) studied monkey and human combination of visual (optic flow) and vestibular cues for discriminating heading direction. Similar to the method described above, participants were presented trials of visual information only (random-dot cloud simulating optic flow presented via 3D glasses), vestibular information only (physical motion of a platform on which participant chair was fixed) and both (optic flow and platform motion). The reliability of the visual stimulus was varied (by manipulating the motion coherence of the optic flow pattern), and a small discrepancy in heading angle between visual and vestibular information was introduced on some bimodal trials. Both monkeys and humans showed reduced bimodal thresholds (see also similar findings by J. S. Butler et al., 2010; Gu, Angelaki, & Deangelis, 2008), and adjusted relative visual-vestibular weights in the direction predicted by the MLE (i.e. increasing visual weight as visual coherence increased). However, on average, both monkeys and humans tended to significantly over-weight the vestibular cue (or under-weight the visual cue), and bimodal thresholds (though reduced) were significantly greater than optimal predictions. These findings suggest that monkeys and humans were combining visual and vestibular information to heading direction sub-optimally. Specifically, monkeys and humans showed a bias toward relying on vestibular information for heading judgments. The authors suggested that the overweighting of vestibular information may be accounted for by causal inference models (e.g. Kording et al., 2007), in which the ideal observer considers the information provided by each cue but also the probability that the two cues arose from the same source. Since the optic flow information presented may have indicated either self- or environmental-motion, the authors argued that there was a causal ambiguity to resolve that led to the vestibular overweighting observed.

The role of prior knowledge

In addition to combining redundant information from multiple sensory inputs to minimize uncertainty, an ideal observer can reduce uncertainty in their perceptual judgments by combining sensory information with prior knowledge (based on their experience within the environment) in line with Bayes theorem (Eq. 4). Kording and Wolpert (2004) showed that human adults use this optimal Bayesian rule to reduce uncertainty in their estimates of a cursor’s position. Specifically, they asked participants to accurately point toward a visual target using their right index finger, and manipulated the displacement and reliability of the visual feedback. Results indicated that, to generate movements toward the target, participants combined prior knowledge (of the distribution of displacements, learnt during 1,000 training trials) with sensory evidence (of what they saw on each trial) as predicted by Bayesian statistics. Other aspects of human perception have similarly been explained in terms of combined reliance on prior knowledge and sensory information, including for example the combination of prior knowledge about visual scenes with image features for interpreting images (e.g. review by Kersten, Mamassian, & Yuille, 2004; Langer & Bulthoff, 2001). Moreover, a number of perceptual illusions and biases have been found to reflect reliance on prior knowledge (Senna, Parise, & Ernst, 2015; Stocker & Simoncelli, 2006; Welchman, Lam, & Bulthoff, 2008). Prior knowledge accurately represents the statistics of the environment, indicating the likelihoods of certain events or stimuli. Consequently, however, reliance on prior knowledge to reduce uncertainty can introduce perceptual inaccuracies (see also section 1.2.3), since the most likely event/stimulus will not always represent that being perceived. For example, humans tend to underestimate the speed of moving objects, with greater underestimations for low-contrast (noisier) stimuli than high-contrast stimuli (Thompson, 1982). This bias in object speed has been explained by use of a “slow motion prior” reflecting (accurately) that objects in natural visual scenes are most likely to be static (Stocker & Simoncelli, 2006; Welchman et al., 2008). In line with Bayesian statistics, as sensory uncertainty increases (i.e. contrast of stimuli decreases), observers rely increasingly on this slow motion prior, leading to greater underestimations of object speed.

1.2.3 The ideal observer maximizes accuracy

Reliability-weighted cue averaging maximizes precision, but not necessarily accuracy. Precision (or reliability) measures the consistency of a percept elicited by a repeated stimulus, whereas accuracy measures the extent to which the percept truly represents the physical stimulus. Judgments with low variable error are precise, whereas judgments with low constant error are accurate (see Fig. 3). The most precise sensory cue is not necessarily always the most accurate, and consequently, reliability-based cue combination could result in biased perception (e.g. Watt, Akeley, Ernst, & Banks, 2005). For example, in Figure 1, assuming that the visual estimate ($\hat{S}_V$) correctly (accurately) represents the size of the physical stimulus, whereas the haptic size estimate ($\hat{S}_H$) is biased, combining visual and haptic estimates introduces an error into the bimodal size estimate ($\hat{S}_{VH}$), despite improving bimodal precision.

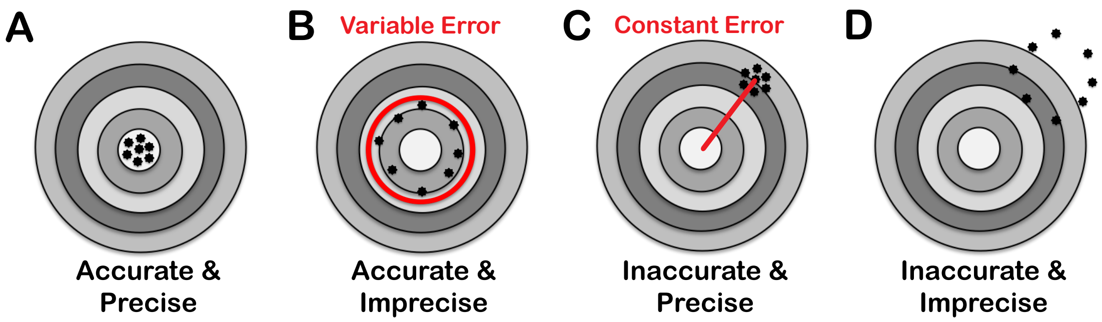


Figure 3: Illustration depicting differences between accuracy and precision.

Since agreement between perception and the physical environment is fundamental to human survival, the optimal observer should maximize both accuracy and precision. Observers acquire prior information regarding the relative accuracy of sensory cues through feedback following interactions with the environment. Accordingly, in an ideal observer framework, maximising accuracy could be modelled by adding a prior toward favouring the more accurate cue (see for e.g. Battaglia et al., 2003). Hence if, for example, observers had prior information that vestibular sensory cues were more accurate than visual sensory cues for heading discrimination, they ought to weight vestibular cues more than predicted based on reliability alone. However, a robust system should be able to use accurate cues to reduce biases, i.e. instead of simply relying less on an inaccurate sensory cue, accurate sensory cues ought to be used to adjust inaccurate cues so that they become accurate in the long-term.

Correspondingly, there is evidence that during development, children learn the correspondences between different sensory cues, with the most accurate sense teaching ('calibrating') the others (Gori, 2015; see section 1.7). Maintaining accuracy is complex because body and/or environmental changes can alter the mapping between sensory modalities, and consequently sensory mappings are continually updated (known as 'recalibration'). For example, wearing prism glasses shifts the visual field, altering the visual-motor mapping, but after a brief period of exposure, humans quickly learn to adapt to the visual displacement. How then do humans learn to recalibrate sensory cues to reduce perceptual errors?

Since perceptual errors can be either systematic, due to a miscalibration between sensory cues, or random, reflecting (internal or external) sensory noise, recalibration must involve a process that averages out random error whilst also accounting for changes in systematic error. Research has aimed to understand how the human system solves the problem of maximising both precision and accuracy. Burge, Ernst, and Banks (2008) found that human adaptation in human reaching was largely well-predicted by an optimal Kalman filter model that combines noisy measurements over time, with prior information of mapping stability, to maximise both accuracy and precision. They asked participants to reach toward a visual target, after which visual feedback was provided. The visual-motor mapping and the reliability of the visual feedback cue were manipulated. As predicted by a Kalman filter, participants adapted more slowly when the position of the feedback was less certain, and adapted more quickly when the uncertainty of the visual-motor mapping was increased.

In a different study, Zaidel, Ma, and Angelaki (2013) instead manipulated the mapping and the reliability of *both* decision cues. As in earlier studies, monkeys and humans reported whether self-motion was to the right/left of straight ahead using visual (optic flow) and/or vestibular (platform motion) information, and were given auditory external feedback regarding their responses (high tone indicating correct response, low tone incorrect). The experiment consisted of three phases: (i) Pre-calibration, in which visual-only, vestibular-only, or combined cues were presented (without feedback) to measure baseline bias and reliability; (ii) Calibration, during which only combined visual-vestibular cues with a discrepancy were presented; and (iii) Post-calibration, in which visual-only and vestibular-only cues were presented to measure adaptation (and combined cues were interleaved to retain calibration during measurement). As expected, when a cue (visual or vestibular) was both more reliable and more accurate, only the less reliable and inaccurate cue shifted. However, when the less reliable cue was accurate (and the more reliable cue was inaccurate), both cues shifted together, i.e. the accurate cue also shifted, and away from the inaccurate cue. The authors suggested that, since external feedback was provided on the combined estimate, when the less reliable cue was more accurate, each individual cue was calibrated in accordance with the combined estimate (referred to as ‘cue yoking’). However, they argued that such yoking is transient, and would ultimately converge on the accurate solution.

Interestingly, in an earlier study, Zaidel, Turner, and Angelaki (2011) found that when no external feedback is given (i.e. cue accuracy is unknown), discrepant sensory cues undergo mutual calibration toward one another, but vestibular adaptation was greater than visual adaptation, irrespective of relative cue-reliability. The authors argued that this was consistent with cue calibration depending on cue-accuracy instead. However, whereas supervised calibration prioritises external accuracy, unsupervised calibration (no feedback) aims for internal consistency instead.

In summary, human adults aim to maximize the accuracy (when feedback is available), and the precision of their perceptual decisions. The process by which these objectives are achieved simultaneously depends on the nature of the feedback provided and the relative reliability and accuracy of each cue.

1.2.4 Possible costs of multisensory combination

Whilst combining sensory information may be advantageous for reducing reaction times, increasing precision and maximizing accuracy, integrating sensory information may also come at a cost. Specifically, integration could be disadvantageous if, as a consequence, the brain were only able to access the combined percept and not the individual sensory information. This is because, since combining cues optimally involves computing the reliability-weighted average, the same combined percept could be achieved when both cues indicate a medium value as when both cues greatly differ but average to a medium value. As a result, it could be impossible for an observer to discriminate between different physical stimuli that yield the same combined percept.

Hillis, Ernst, Banks, and Landy (2002), found evidence for ‘mandatory fusion’ within, but not between senses. In particular, they found visual disparity and visual texture information to slant was indeed lost through combination, but that this was not the case for visual and haptic cues to object size. However, though not fused completely, ‘between’ signals interact since individual visual and haptic sensory estimates have been shown to be biased (in the direction of optimal combined visual-haptic weights) by the accompanying modality (Ernst, Banks, & Bulthoff, 2000). Note, that this is in agreement with the finding by Zaidel et al. (2011) that in the absence of external feedback, sensory cues undergo calibration toward one another to maximize internal consistency (section 1.2.3). Ernst (2006a) accounted for these findings in an ideal observer framework by using a task-dependent ‘coupling prior’ representing the probability distribution of naturally occurring mappings between sensory signals. Having a weaker coupling than visual disparity and texture cues to slant (represented by a wider coupling prior distribution), visual and haptic size cues are able to adapt quickly to changes in their mapping (thereby maximizing accuracy, see section 1.2.3), since (unlike with mandatory fusion) discrepancies between cues can be detected.

Consequently, combining cues, though optimal for reducing uncertainty or maximizing accuracy, can lead to multisensory perceptual errors, or 'illusions' described below.

1.2.5 Multisensory perceptual illusions

Multisensory illusions reflect combined use of information from separate modalities to yield a joint estimate of an external property that (in the case of the illusions) is not a true representation of the physical world. Hence, these multisensory illusions are examples of perceptual errors or costs of multisensory combination. The illusions are in line with the predictions or reliability-weighted cue averaging, and therefore provide support that humans combine cues according to this ideal observer model. Furthermore, multisensory illusions have frequently been used in research studies to assess whether individuals with sensory impairments combine multisensory information in the same way as control participants (see sections 1.4.1 – 1.4.3). There are a variety of multisensory illusions involving interactions between different senses. Brief descriptions of the most frequently reported illusions are provided below.

The Ventriloquist Effect

The ventriloquist effect describes the effect whereby temporally aligned but spatially displaced visual information can bias (or 'capture') the perceived location of a sound. The illusion takes its name from the stage act of ventriloquism, in which a person (the ventriloquist) changes their voice so that it appears to come from a puppet whose lips are made to move. Such visual-auditory binding may be facilitated by top-down knowledge, however the effect is also seen when using neutral stimuli such as light flashes and tones (Bertelson & Radeau, 1981; Wallace, Roberson, et al., 2004). This behaviour is in accordance with that of an ideal observer that combines cues to maximise accuracy and precision, since visual information tends to be both more accurate and more reliable than auditory information. Indeed, in line with ideal observer predictions, when the reliability of the visual stimulus is reduced (by adding blur) spatial judgments are biased towards the location of a disparate sound source (Alais & Burr, 2004).

Additionally, as predicted by sensory calibration, exposure to a consistent audio-visual spatial conflict leads to **the ventriloquist aftereffect**, in which localization of sound sources is shifted toward the visual position, correcting for the discrepancy in the adaptation period (Frissen, Vroomen, de Gelder, & Bertelson, 2005; Lewald, 2002a; Recanzone, 1998).

The Audio-Tactile Spatial Ventriloquist Effect

A spatial ventriloquist effect is observed with auditory and tactile stimuli too. Specifically, the perceived location of an auditory stimulus is shifted toward the location of a concurrent tactile stimulus (e.g. Occelli, Bruns, Zampini, & Roder, 2012; Renzi et al., 2013).

Temporal Ventriloquism

Whereas vision tends to be more accurate and reliable for spatial judgments, audition tends to dominate temporal judgments, causing ‘temporal ventriloquism’: When an auditory and a visual stimulus are presented in close temporal proximity, the perceived onset of the visual stimulus is influenced by the auditory cue (Bausenhart, de la Rosa, & Ulrich, 2014; Getzmann, 2007; Vroomen, de Gelder, & Vroomen, 2004). For example, Bausenhart et al. (2014) asked participants to judge the duration of either visual or auditory pulses, ignoring the simultaneously presented task-irrelevant visual or auditory modality. Despite this, the perceived duration of the visual (though also – albeit to a lesser extent – auditory) pulses was clearly biased toward the duration of intervals in the task-irrelevant modality (as predicted by Ernst’s 2006 coupling prior, see section 1.2.4).

The Sound-Induced Flash Illusion

Another example of auditory signals dominating visual perception is the sound-induced flash illusion, in which the presence of two auditory beeps presented simultaneously with a single flash, can result in observers reporting seeing two flashes (also referred to as the 'Shams fission illusion', see Shams, Kamitani, & Shimojo, 2000). Similarly, a single beep can result in the perception of a double flash as a single flash (Andersen, Tiippana, & Sams, 2004).

The Audio-Tactile Illusory Flash Effect

Similar to the sound-induced flash illusion described above, a single tactile stimulus presented with two successive sounds is frequently perceived as two tactile sensations (Hotting & Roder, 2004).

The Audio-Tactile Parchment Skin Illusion

When the sound generated by rubbing hands together is manipulated by either amplifying or reducing the high-frequency content of the sound, this changes an observer's perception of the experienced smoothness or dryness of the palm (Jousmaki & Hari, 1998). In particular, when either the proportion of high frequencies or the average level of the sound increases, the perceived roughness/moisture decreases, (and the perceived smoothness/dryness increases).

The McGurk Effect

The McGurk effect (McGurk & MacDonald, 1976) demonstrates the influence of lip movements on the perception of speech sounds. In particular, when certain syllables are presented with incongruent lip movements, this results in the percept of a different phoneme. For example, when the spoken syllable /ba/ is presented with lip movements representing /ga/, observers commonly report perceiving a third intermediate phoneme /da/. Similarly, the sound /pa/ tends to be perceived /ta/ when coupled with the visual lip movement for /ka/. This is consistent with observers combining visual (lip movements) with auditory information (speech sounds) to perceive the spoken syllable.

The Motion-Bounce (or Stream-Bounce) Illusion

When two identical visual targets move toward each other along the same direction and at the same speed, on meeting these can either be perceived as bouncing off or streaming through each other. However, Sekuler, Sekuler, and Lau (1997) demonstrated that when a brief sound is presented at (or just before) the moment that targets coincide, this biases perception toward bouncing.

1.2.6 Multisensory processing in the brain

As reviewed thus far, research studies assessing human behaviour have found that perceptual decisions (including errors) are well predicted by ideal observer models that combine multisensory cues optimally (to reduce reaction times, uncertainty or inaccuracies). Consequently, a question that has arisen is: where in the brain does this multisensory processing take place? Traditionally, it was assumed that individual senses were processed in independent channels, in unisensory areas, and only combined at later processing stages, in multisensory convergence zones in the brain, such as the superior colliculus. However, more recently, researchers have found that even traditional unisensory cortical areas (i.e. primary cortices) receive multisensory input (see reviews by Driver & Noesselt, 2008; Ghazanfar & Schroeder, 2006). Specifically, in addition to multisensory sites in association cortex (e.g. ventral intraparietal areas, ventral premotor cortex, posterior auditory association cortex, superior temporal polysensory areas; see Schroeder & Foxe, 2002), neuroimaging studies have shown that primary cortical areas also respond to inputs from more than one sensory modality (Martuzzi et al., 2007; Pekkola et al., 2005; Vetter, Smith, & Muckli, 2014). Moreover, evidence from neural recordings in animals suggests that such findings may reflect activity from multisensory neurons that respond to input from more than one-sensory modality, as opposed to activity from different-modality sensory-specific neurons (Fishman & Michael, 1973; Morrell, 1972). For example, Fishman and Michael (1973) made microelectrode recording from cells in cat visual cortex, and found 38% of neurons responded to both acoustic and visual stimuli. Furthermore, several studies have reported that multisensory co-stimulation can modulate activity in traditional unisensory cortical areas (Bizley, Nodal, Bajo, Nelken, & King, 2007; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007; Watkins, Shams, Tanaka, Haynes, & Rees, 2006). For example, Watkins et al. (2006) found that activity in human primary visual cortex (V1) was enhanced by concurrent auditory information that influenced visual perception (specifically, when the presence of two beeps resulted in a single flash being perceived as a double flash, known as the sound-induced flash illusion, see section 1.2.5). Although a neuron whose response to a unimodal stimulus is modulated by that of a different sensory modality is, arguably, not strictly multisensory (in that it does not respond to input from more than one modality), these findings indicate multisensory interactions in early processing stages.

Hence, increasingly research is finding evidence of multisensory activity in unisensory areas, leading some researchers to propose that perhaps all neocortex is essentially multisensory (Ghazanfar & Schroeder, 2006). Certainly, the brain does show regional preferences for one modality over another, however growing evidence suggests that these regions are modulated by inputs from other modalities too. A debate has arisen regarding the extent to which multisensory effects in primary cortices are simply the result of feedback connections from multisensory convergence sites elsewhere, as opposed to ‘feedforward’ activity driving later perception. Critically, reports of early-latency (< 100ms after stimulus onset) multisensory interactions, in animals where top-down modulations are blocked (Barth, Goldberg, Brett, & Di, 1995), and studies in humans showing improved perception of non-visual stimuli when coupled with occipital-TMS stimulation (Romei, Murray, Merabet, & Thut, 2007), imply at least some bottom-up multisensory connections (see reviews by De Meo, Murray, Clarke, & Matusz, 2015; Murray et al., 2015).

1.2.7 Three principles of neural multisensory integration

The most studied multisensory convergence zone in the brain is the superior colliculus, which receives inputs from primary auditory, somatosensory and visual areas. The superior colliculus plays an important role in directing behavioural responses, and in particular in controlling eye movements in primates. Notably, Stein and colleagues have conducted multiple studies investigating multisensory processing of neurons in the cat superior colliculus (e.g. Meredith & Stein, 1983, 1986; Stein, Meredith, Huneycutt, & McDade, 1989; Wallace & Stein, 1997; Xu, Yu, Stanford, Rowland, & Stein, 2015). These studies have been influential in shaping understanding of both the development of multisensory integration and defining three main principles under which multisensory neurons operate.

The Spatial Rule

A multisensory neuron in the superior colliculus has a receptive field for each of its sensory modalities, representing proximal regions of sensory space. When stimulated by two different sensory stimuli in close spatial proximity, the neuron's response is significantly greater than that evoked by the most effective of the two unimodal inputs; even exceeding the sum of the unimodal inputs (referred to as 'superadditivity'; Meredith & Stein, 1986, 1996; Wallace, Meredith, & Stein, 1998). In contrast, when the two unimodal stimuli do not coincide spatially, there is either no interaction, or the neuron's response is considerably depressed (Kadunce, Vaughan, Wallace, Benedek, & Stein, 1997). Thus, 'the spatial rule' of multisensory integration states that the neural response enhancement of multisensory stimulation is dependent on the spatial alignment of the individual sensory receptive fields. Accordingly, behavioural studies show that human multisensory combination breaks down when stimuli come from largely different locations (Gepshtein, Burge, Ernst, & Banks, 2005), which coincides with causal inference models in that the likelihood that two signals have the same cause correlates with the signals' spatial proximity. However, humans have shown integration behaviour over spatial disparities of up to 40 degrees (Harrington & Peck, 1998), and some evidence suggests that the spatial rule only applies when space is relevant to the perceptual task (see review by Spence, 2013).

The Temporal Window

Similarly, for multisensory integration to occur, two stimuli must be aligned in time, albeit not precisely: there is a 'temporal window' of integration (Meredith, Nemitz, & Stein, 1987) to account for the difference in speeds with which information from different modalities is processed. Sensory integration breaks down with signals that are temporally asynchronous (Bresciani et al., 2005), since the likelihood that two signals have the same cause correlates with the signals’ temporal proximity (as well as spatial proximity).

Inverse Effectiveness

According to inverse effectiveness, the magnitude of the multisensory enhancement is inversely related to the effectiveness (efficacy) of the unisensory stimuli (Wallace, Wilkinson, & Stein, 1996). However, more recently, Rowland, Perrault, Vaughan, and Stein (2015), contested this principle, arguing that the magnitude of multisensory enhancement is dependent on the reliability of the sensory estimates instead, which frequently covaries with efficacy. They repeatedly presented an alert cat with visual, auditory and visual-auditory cues, thereby reducing stimulus efficacy through habituation (but not decreasing the reliability of the sensory estimates). Importantly, both unisensory and multisensory superior colliculus neuron responses reduced in equal proportion.

Interestingly, multisensory neurons do not show these multisensory integration properties or abilities at birth, but instead these develop with multisensory experience (discussed later in section 1.4.1).

1.2.8 Development

As most objects and events in the physical environment stimulate multiple senses simultaneously, young infants must learn which stimulation patterns to combine (because they correspond to the same object/event) and which to differentiate (because they correspond to different objects/events). Two theories have emerged to account for this learning process: the integration view and the differentiation view. The integration view proposes that different sensory modalities are separate during the initial stages of postnatal development and the infant gradually learns to integrate these different senses through repeated experience with the environment (Piaget & Cook, 1952). The differentiation view, instead, argues that the different senses are initially unified and the infant gradually differentiates between increasingly finer levels of sensory stimulation, through experience with the environment (Gibson & Gibson, 1955). In line with this differentiation view, neural and behavioural evidence indicates that human infants are able to perceive properties conveyed by more than one sensory modality (e.g. spatial collocation, temporal synchrony) in the first several months following birth. Accordingly, some forms of multisensory integration are acquired early during the first year of life (see review by Lewkowicz, 2002), such as perception of the stream-bounce illusion (Scheier, Lewkowicz, & Shimojo, 2003), detection of asynchrony between speech and lip movements (Dodd, 1979), and reduced reaction times to bimodal (audio-visual) targets compared to unisensory targets (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006). Yet, in line with the integration view, over the first year of life, infants become more skilled at detecting intersensory relations that are modality-specific. For example, infants can learn to link the type of sound that a toy makes with the toy's colour. Hence, both integration and differentiation processes appear to be involved (for review see Lickliter & Bahrick, 2004). Importantly, however, research has shown that other multisensory abilities – most notably the combination of multisensory information to improve precision – do not develop until much later on in childhood (Gori, Del Viva, Sandini, & Burr, 2008; Nardini, Bedford, & Mareschal, 2010; Nardini, Jones, Bedford, & Braddick, 2008; Petrini, Remark, Smith, & Nardini, 2014). For example, the ability to integrate visual and haptic signals to reduce uncertainty in orientation and size judgments (Gori et al., 2008), does not develop until eight years of age.

It has been proposed that early differences in the development of multisensory processes may be explained by anatomical and physiological differences in sensory maturation rates, whereas later differences (occurring well after the maturation of individual senses) are due to the need for children to continually recalibrate their senses as they grow, due to, for example, changing body size. Accordingly, in multisensory tasks, children under eight years have shown a strong unisensory dominance, (i.e. they prefer to rely on a single sense, such as vision for audio-visual localization; Gori, Sandini, & Burr, 2012), and it has been suggested that this dominant sense, being more accurate, is used to calibrate the other sense for the task at hand (see review by Gori, 2015). However, children have also been shown to use suboptimal multisensory integration strategies, such as alternating between cues (Nardini et al., 2008) and failing to ignore irrelevant cross-modal information (Petrini, Jones, Smith, & Nardini, 2015). These findings indicate that learning to combine multisensory information optimally is a complex process, involving both sensory calibration and learning how to weight sensory information for improved precision.

As discussed earlier, an ideal observer maximizes both accuracy and precision. Since the role of calibration is to remove systematic biases, it makes sense that the most accurate (not the most precise) cue, for a given task, is used to calibrate the others (see e.g. Gori, Sandini, et al., 2012). An implication of cross-sensory calibration is that if the more accurate sensory modality for a specific task is unavailable during development, this will lead to impaired performance (relative to children for whom the more accurate cue is available) in the less accurate modality. Gori, Sandini, Martinoli, and Burr (2010) tested this by comparing the haptic orientation and size discrimination thresholds of congenitally visually impaired children (aged 5-19 years) with that of age-matched controls. Having already shown that vision calibrates touch for orientation judgments, whereas touch calibrates vision for size judgments (Gori et al., 2008), results confirmed predictions with visually impaired children showing significantly worse orientation, but not size, discrimination thresholds (see also Gori, Tinelli, Sandini, Cioni, & Burr, 2012).

Vision is the most accurate (and precise) sense for orientation judgments, and also other tasks involving spatial perception, such as localization of stimuli (see e.g. the ventriloquist effect, section 1.2.5). This is consistent with neurophysiological evidence in animals showing that visual cues play a fundamental role in calibrating the spatial maps of non-visual modalities in the superior colliculus (e.g. King & Carlile, 1993; Wallace, Perrault, Hairston, & Stein, 2004, see section 3.1). Interestingly, studies that have assessed the auditory localization abilities of congenitally blind individuals, have often reported superior auditory localization abilities (e.g. Ashmead et al., 1998; Lessard, Pare, Lepore, & Lassonde, 1998), and such compensatory behavioural changes have been linked to cortical reorganisation. The next section reviews studies documenting compensatory changes in non-visual processing following blindness.

1.2.9 Summary

Much research has found that human adults maximize the speed, precision and/or accuracy of their perceptual judgments, in line with ideal observer predictions, by combining (i) redundant information from multiple senses or (ii) sensory information with prior knowledge (see section 1.2.1-1.2.3). Multisensory perceptual advantages may be mediated by multisensory neurons, in primary cortical areas and/or multisensory convergence sites, which respond preferentially to spatially- and/or temporally- congruent multisensory inputs. Whilst some multisensory processes are acquired early in life, the ability to combine multisensory information to reduce perceptual uncertainty does not develop until late in childhood. This suggests that learning to combine information to maximize precision is a complex process that involves experience with events or objects that stimulate multiple senses. Visual (with non-visual) experience may be particularly important, since vision has been found to calibrate the spatial representations of non-visual senses.

1.3 Non-Visual Processing following Blindness

Blindness is the (temporary or permanent) complete or nearly complete loss of useful sight, and can be caused by damage to the eye, the optic nerve or the visual cortex. A distinction is frequently made between ‘early blind’ and ‘late blind’ individuals, where ‘early blind’ is used to refer to congenitally blind individuals or individual blinded before the age of three years, and late blind refers to individuals blinded later in life (after three years). The reason for this is that research has documented certain early (pre-three years) critical periods for the development of aspects of visual function (for more details see section 1.5.1). Visual impairment refers to partial vision loss that is not fixable by usual means such as glasses. There are many causes of visual impairment including problems in the eye (e.g. cataracts, glaucoma, macular degeneration) and the brain (e.g. stroke, prematurity, trauma; referred to as cortical visual impairment). In this next section, research studying the effect of blindness on non-visual processing is reviewed. Partial vision loss is discussed later (see section 1.6).

As vision plays an important role in calibrating the spatial maps of other sensory modalities (see sections 1.2.8 and 1.4.1), visual deprivation could be predicted to have a detrimental effect on spatial perception. However, extensive research has documented compensatory adjustments in residual senses following visual loss that often enable blind individuals to perform at least comparably to normally sighted individuals in some spatial tasks (see review by Collignon, Voss, Lassonde, & Lepore, 2009). These behavioural changes have been linked to neural reorganisation; it is proposed that altered sensory experience can cause the brain to reinforce existing neural connections or form new synapses, resulting in compensatory behaviour (see reviews by Bavelier & Neville, 2002; Merabet & Pascual-Leone, 2010; Noppeney, 2007). Such experience-dependent plasticity has been documented both within residual sensory regions and via the de-afferent visual cortex, particularly in young animals, but also in adults too (see review by Merabet & Pascual-Leone, 2010). The extent of the cortical reorganisation appears to depend on the onset, severity and duration of the sensory deprivation (see review by Lazzouni & Lepore, 2014). This section will review some of the key compensatory changes that have been documented following blindness.

1.3.1 Compensatory changes in auditory localization

Despite vision playing an important role in calibrating auditory space (see sections 1.2.8 and 1.4.1), many studies have shown comparable or enhanced auditory spatial processing abilities by early-blind individuals (Ashmead et al., 1998; Doucet et al., 2005; Lessard et al., 1998; Roder et al., 1999; Voss et al., 2004). Specifically, studies have found that, when asked to point or reach toward different azimuthal sound sources, early-onset blind children (Ashmead et al., 1998) and adults (Lessard et al., 1998) show similar or better binaural horizontal sound localization accuracy to sighted controls. Moreover, using a minimum audible angle task, Voss et al. (2004) showed that blind adults are able to map auditory space beyond their peri-personal environment, (where auditory representations could be calibrated using sensory-motor feedback instead of vision). Both late-onset and early-onset blind adults have shown improvements in accuracy (Fieger, Roder, Teder-Salejarvi, Hillyard, & Neville, 2006; Voss et al., 2004) or precision (visually-deprived ferrets and cats, King & Parsons, 1999; Rauschecker & Kniepert, 1994) relative to normally sighted controls in horizontal localization tasks; although some evidence suggests that these improvements are limited to the processing of sounds presented in peripheral space only (Fieger et al., 2006; King & Parsons, 1999; Rauschecker & Kniepert, 1994; Roder et al., 1999; Voss et al., 2004). Indeed, Lewald (2007) found that even normally sighted adults deprived of light for just ninety minutes showed improved accuracy (though not precision) in head pointing toward an auditory target (returning to pre-deprivation values after re-exposure to light), suggesting that compensatory behaviour can be rapidly initiated.

To localize a sound source, the human brain uses monaural spectral cues and interaural intensity and timing differences. Spectral cues refer to how the human outer ear (pinna and external ear) affect the perception of the sound, by filtering sounds based on their frequency and input direction, described by the head-related transfer function (HRTF). These are particularly relevant for determining the elevation of a sound source in the midline (where there are no differences in interaural cues) and resolving front/back confusions, whereas interaural differences are important for azimuthal localization. In addition to enhanced horizontal auditory localization, early-blind participants, unlike sighted controls, have been shown to accurately localize sounds monaurally, leading some researchers to propose that blind individuals may compensate by increased and more effective use of auditory spectral cues (Doucet et al., 2005; Lessard et al., 1998); though interestingly early-blind adults show impaired vertical sound localization (Lewald, 2002b; Zwiers, Van Opstal, & Cruysberg, 2001), which relies primarily on spectral cues (Carlile, Martin, & McAnally, 2005).

Hence, when pointing towards a sound source, or completing a minimum audible angle task, blind individuals show enhanced auditory localization in the azimuthal plane, albeit limited to sounds presented in peripheral space; however, when localizing sounds in vertical space, their performance is impaired relative to normally sighted controls. Thus, these findings suggest that vision is needed to calibrate auditory spatial maps for certain localization judgments, whereas, for other spatial decisions, sensory compensation can actually lead to advanced processing. Consequently, it is important to understand what drives compensatory behaviour, and what is distinctive about tasks that lead to enhanced, versus impaired, behavioural outcomes.

Compensatory differences in auditory processing may be linked to physiological changes in auditory processing structures, including multisensory areas and the primary auditory cortex (Elbert et al., 2002; Korte & Rauschecker, 1993; Petrus et al., 2014; Rauschecker & Harris, 1983). For example, cats deprived of binocular vision from birth show an increased number of audio-responsive neurons tuned to azimuthal space, in areas where different sensory modalities come together including the superior colliculus (Rauschecker & Harris, 1983) and the anterior ectosylvian cortex (Korte & Rauschecker, 1993). Strikingly, however, physiological changes have also been documented within *visual* cortex (Collignon, Davare, Olivier, & De Volder, 2009; Gougoux, Zatorre, Lassonde, Voss, & Lepore, 2005; Kujala, Alho, Paavilainen, Summala, & Naatanen, 1992; Leclerc, Saint-Amour, Lavoie, Lassonde, & Lepore, 2000; Poirier et al., 2006; Weeks et al., 2000). For example, Gougoux et al. (2005) asked early-blind and sighted participants to localize (binaurally or monaurally) 30ms broadband noise bursts presented from speakers mounted on a semicircular array, within a positron emission tomography (PET) scanner. Occipital cortex activation was found only in early-blind participants who showed superior sound localization performance during monaural testing (when one ear was plugged). This sub-group of early blind participants showed near-accurate monaural sound localization, whereas other early-blind and sighted participants were highly inaccurate. The degree of occipital cortex activation was strongly correlated with sound localization accuracy. Thus, the results suggest that the enhanced capacity of early-blind individuals to use monaural (spectral) cues is driven by computations within the occipital cortex. Collignon, Lassonde, Lepore, Bastien, and Veraart (2007) demonstrated the causality of occipital cortex activation on auditory localization further by using repetitive transcranial magnetic stimulation (rTMS): The application of rTMS to right dorsal extrastriate visual cortex, whilst participants discriminated the position of two sounds (presented from seven speakers on a semi-circular array), significantly disrupted the performance of early-blind, but not sighted, participants. Since studies have shown that the right dorsal extrastriate visual system is specialised for visuospatial processing (Ungerleider & Mishkin, 1982), the authors argued that there is ‘an anatomical functional correspondence’ between visual spatial processing and auditory spatial processing in sighted and blind individuals respectively. Indeed, other researchers have similarly found a functionally specific recruitment of visual cortex (reviews by Dormal & Collignon, 2011; Voss & Zatorre, 2012), indicating that the organisation (or ‘architecture’) of the brain is set, regardless of visual experience (Ricciardi, Bonino, Pellegrini, & Pietrini, 2013). Thus, compensatory behaviour can be driven by adaptive plasticity both within (intra-modal) and between (cross-modal) functionally relevant sensory areas.

Whilst evidence suggests that the brain can physically adapt to visual deprivation, contrasting research indicates that vision is fundamental in driving the maturation of the auditory spatial map, at least within the superior colliculus. The superior colliculus is a midbrain structure that receives multisensory input and is involved in the (often reflexive) orienting of the eyes and head toward nearby visual and auditory stimuli. It is hence involved in processing auditory spatial location; however note that the nature and development of the auditory and visual interactions in this midbrain structure may differ from those in cortical areas, given differences in the functionality of these different brain regions. Collicular maps of auditory space fail to emerge in guinea pigs deprived of vision early in life (Withington-Wray, Binns, & Keating, 1990), whereas in visually-deprived ferrets, auditory maps emerge but these do not align normally with the visual spatial map when vision is restored (King & Carlile, 1993). Moreover, when the spatial relationship between auditory and visual cues is systematically misaligned, a corresponding physiological shift in the representation of auditory space by collicular neurons has been documented (Feldman & Knudsen, 1997; King, Hutchings, Moore, & Blakemore, 1988; Knudsen & Brainard, 1991; Wallace & Stein, 2007), indicating that vision plays a key role in calibrating the auditory spatial map. Accordingly, as noted by King (2015), the plasticity observed in the superior colliculus is likely related to the development of the capability to integrate visual and auditory spatial information, as opposed to the ability to use hearing alone for spatial judgments (see e.g. Wallace, Perrault, et al., 2004). Consequently, there is a trade-off between the advantages of compensatory plasticity in auditory processing areas versus the disadvantage of not having vision to align visual and auditory space for multisensory processes. However, the question remains as to why vision is important for auditory vertical but not horizontal localization.

Gori and colleagues propose that blind individuals show impaired auditory localization performance on tasks that require a Euclidean (as opposed to simply topological) representation of auditory space (Finocchietti, Cappagli, & Gori, 2015; Gori, Sandini, Martinoli, & Burr, 2014). Gori et al. (2014) compared the auditory localization performance of congenitally blind and normally sighted adults using both a spatial bisection task, in which participants reported whether the second sound source was spatially closer to the first or third (final) sound source, and a minimum audible angle task, in which participants reported whether the first or second of two sounds was more rightward. Whilst similar minimum audible angle performances were observed for both groups, congenitally blind participants showed significantly impaired spatial bisection performance, (though no impairments were found in a temporal bisection task, suggesting that the deficit could not be accounted for by differences in memory). This finding was striking, being the first report of deficits in *horizontal* auditory spatial localization in the congenitally blind and standing in stark contrast to earlier findings showing comparable or enhanced auditory spatial processing abilities by early-blind individuals instead (Ashmead et al., 1998; Doucet et al., 2005; Lessard et al., 1998; Roder et al., 1999; Voss et al., 2004). The authors concluded that, whereas vision is not necessary for localizing a sound or indicating whether a sound is further right or left of another sound, vision is necessary for constructing a map of auditory space, and this map is required to complete the auditory spatial bisection task.

To test this further, Gori and colleagues predicted that, whilst early- and late-blind adults show superior performance in judging the direction of horizontal sound motion (Lewald, 2013), they would show impairments in reproducing the trajectory and final position of a sound motion (Finocchietti et al., 2015); because the latter task requires a Euclidean map relating the position of the sound in space and time (whereas horizontal direction discrimination can be evaluated by comparing sounds relatively). Participants (early- and late- blind, and sighted blindfolded adults) were sat opposite an experimenter and a graduated circular perimeter was mounted between them. On each trial, the experimenter moved the hand-held sound source from the centre of the circle towards one of eight random positions on the perimeter. Participants maintained their index finger at the centre of the circle, until the experimenter had finished, at which point they were asked to reproduce the complete trajectory of the audio motion. Early-blind, but not late-blind or sighted blindfolded, participants showed a clear deficit in encoding the sound motion in the lower side of the circular perimeter, however all participants were able to correctly judge the stimulus direction in the horizontal axis. Specifically, early-blind participants tended to compress the perceived location of lower sound targets (i.e. audio motion toward the lower half of the circle perimeter was perceived in higher space). Whilst the task involved processing the position of the sound in space and time, early-blind participants showed accurate performance in the horizontal plane, but not in the vertical plane. This finding is consistent with earlier research reporting that early blind individuals show impaired localization of static auditory targets in the vertical plane (Lewald, 2002b; Zwiers et al., 2001). To localize sounds in the vertical plane, humans rely on spectral cues to location (as opposed to interaural cues which are useful for horizontal localization). Consequently, the poor performance of blind individuals in the vertical (and not the horizontal) plane indicates the importance of vision for calibrating spectral cues, and forming a Euclidean representation of auditory space.

The ability to localize auditory targets in the physical environment with heightened accuracy and precision is clearly beneficial for visually deprived individuals, for example in aiding navigation. In particular, a specific skill developed by some blind individuals is the use of sound reflections to localize physical objects, known as ‘echolocation’. By creating sounds (most notably clicking noises with the mouth), individuals learn to identify the location and size of nearby objects, based on the sound reflections. Whilst sighted individuals can learn to use echolocation too, blind participants have been found to be more accurate at localizing objects based on echo cues (Dufour, Despres, & Candas, 2005), and, similarly to the auditory localization findings reported above, the processing of click-echoes has been shown to recruit visual (calcarine) cortex rather than auditory cortex (Thaler, Arnott, & Goodale, 2011). Whereas localizing the source of a sound would benefit from the suppression of sound reflections from other sources, echolocation would benefit from using these lagging reflections. Recently, Nilsson and Schenkman (2015) tested whether the improved sound localization ability of blind individuals (compared to sighted) is driven by differences in the processing of interaural cues (interaural level or time differences) or in the processing of lagging sounds. They measured blind and sighted participants’ discrimination thresholds for interaural level differences and interaural time differences present in single clicks (no lagging sounds), in the leading component of click pairs (involving suppression of lagging clicks), or in the lagging part of click pairs (involving use of lagging clicks). Blind listeners had greater interaural level difference and interaural time differences sensitivity than age-matched listeners. Furthermore, blind showed the greatest advantage for discriminating interaural level differences in lagging click pairs, suggesting an increased ability to discern interaural level differences in reflected sounds.

1.3.2 Parallels in visual localization following auditory loss

It is worth noting that, comparable to studies comparing the auditory processing abilities of visually deprived and sighted individuals, research has found both enhancements and impairments in the visual processing of deaf participants relative to hearing controls (reviews by Heimler, Weisz, & Collignon, 2014; Pavani & Bottari, 2012). In terms of enhancements, deaf individuals show improved detection and localization of visual stimuli, particularly peripherally-presented visual stimuli (Hong Lore & Song, 1991; Neville & Lawson, 1987; Proksch & Bavelier, 2002), which parallels the improved localization of peripheral auditory targets by the blind (Fieger et al., 2006; King & Parsons, 1999; Rauschecker & Kniepert, 1994; Roder et al., 1999; Voss et al., 2004). Such behavioural advantages may, similarly, be driven by neural plasticity, since bilateral deaf adults recruit auditory cortex when detecting static visual (particularly peripheral) targets (Karns, Dow, & Neville, 2012; Scott, Karns, Dow, Stevens, & Neville, 2014), and moving visual stimuli (Finney, Fine, & Dobkins, 2001; Vachon et al., 2013). Moreover, Lomber, Meredith, and Kral (2010) actually showed that temporarily deactivating posterior or dorsal auditory cortex in congenitally deaf cats eliminated their superior peripheral visual localization or visual motion detection abilities respectively; thereby providing a causal link between cross-modal reorganisation of auditory cortex and compensatory visual processing (but see Bottari, Caclin, Giard, & Pavani, 2011; Codina et al., 2011, suggesting improved visual localization/detection driven by changes within visual system instead). Although less documented than research in the blind, there is also some evidence for functionally specific cross-modal plasticity in deaf individuals, though largely from language studies reporting that sign language processing recruits the temporo-frontal network typically associated with spoken language processing (e.g. Emmorey, Mehta, & Grabowski, 2007; MacSweeney et al., 2002). In terms of impairments, whereas vision is fundamental for sensory spatial calibration (being more accurate and reliable for spatial judgments than the other senses), audition tends to dominate temporal judgments. Accordingly, deaf individuals show impairments (relative to hearing controls) in temporal tasks that involve the reproduction of a visual stimulus' duration (Kowalska & Szelag, 2006), or discriminating the duration of two tactile stimuli (Bolognini et al., 2012). Thus, akin to visual-deprivation, auditory-deprivation can lead to compensatory behaviours in some tasks, driven by experience-dependent plasticity, but also impairments in other tasks reflecting the importance of audition for temporal calibration.

1.3.3 Compensatory changes in navigation

To navigate effectively through the physical world, an individual must be able to form a spatial representation (‘cognitive map’) of the environment, update their position and orientation during travel, and plan routes subject to various constraints, including safety (i.e. avoiding obstacles). Two distinct mechanisms enable spatial updating and orientation: A landmark-based system that uses a physical landmark to allow an individual to fix their heading and position within the environment, and a path integration system that uses self-motion cues from visual (optic flow information), vestibular and proprioceptive sensory systems to calculate movements over time. Adults with healthy vision can improve their navigational performance by combining visual and non-visual cues (Bates & Wolbers, 2014; Kalia, Schrater, & Legge, 2013; Nardini et al., 2008), and as discussed above, vision (being the most accurate sense for spatial judgments) presumably plays an important role in calibrating non-visual navigational cues. However, early blind participants have been shown to perform similarly to sighted-blindfolded individuals in path reproduction (reproducing a walked path) and path completion (returning to the start position via the shortest possible route) tasks (Loomis et al., 1993), indicating that early blind individuals were able to learn the spatial correspondences of non-visual navigational cues, despite the absence of vision. In another study, early and late blind suggesting that the use of spatial navigational cues is not affected by prior visual experience. Moreover, Fortin et al. (2008) actually found that early and late blind individuals showed superior navigational skills compared to normally sighted adults on a route learning task, and significantly increased hippocampal volume. A key function of the hippocampus in humans is the representation of space and formation of a cognitive map (O'Keefe & Nadel, 1978). Increased hippocampal volume has been found in other individuals with expert spatial navigational skills. For example, Maguire, Woollett, and Spiers (2006) found a positive correlation between the number of years London taxi drivers had spent driving taxis and the grey matter density of their right posterior hippocampus. To compensate for not having vision to update spatial coordinates online, blind individuals may store large amounts of information regarding their environment (as taxi drivers do), and this may explain the increase in hippocampal volume observed.

1.3.4 Compensatory changes in tactile orientation sensitivity

Given that Gori et al. (2010) showed early visual loss can disrupt the calibration of haptic cues for orientation (section 1.2.8), it is interesting that other researchers have found heightened tactile orientation discrimination in the blind instead (Goldreich & Kanics, 2003; J. C. Stevens, Foulke, & Patterson, 1996; Van Boven, Hamilton, Kauffman, Keenan, & Pascual-Leone, 2000; Wong, Gnanakumaran, & Goldreich, 2011). For example, Wong et al. (2011) asked (early- and late-) blind and sighted participants to determine which of two sequentially presented gratings was horizontal in orientation, using either the index, middle or ring fingers of each hand, or two sides of their lower lip. Grating groove width was manipulated, and the grating width that could be reliably perceived with 76% probability (d') by each participant was taken as the participant's grating orientation threshold. Fingertip discrimination thresholds were significantly better for blind than sighted participants, but no difference in lip discrimination thresholds between participant groups was found. Furthermore, blind skilled Braille readers showed superior performance compared to blind non-Braille readers, when using their preferred reading index finger, suggesting that tactile experience drives this acuity enhancement. Accordingly, whilst researchers have reported selective activation by blind individuals of occipital cortex during Braille reading (Cohen et al., 1997; Sadato et al., 1996) and greater occipital activation in early- compared to late- blind individuals (Buchel, Price, Frackowiak, & Friston, 1998; Burton et al., 2002), Sathian and Zangaladze (2002) found that even normally sighted adults recruit visual cortex for orientation discrimination tasks. Specifically, using PET, they showed that sighted adults had greater regional cerebral blood flow in extrastriate visual cortex, when completing a grating orientation task, compared to a spatial frequency task (see also Zangaladze, Epstein, Grafton, & Sathian, 1999). Consequently, practice with tactile stimuli can lead to enhancements in orientation discrimination sensitivity, via the strengthening of (pre-visual deprivation) existing connections within occipital cortex. Hence, the question arises as to why blind participants showed impaired performance on the tactile orientation discrimination task developed by Gori et al. (2010), but not on other tasks.

A notable methodological difference between Gori and colleagues' (2010) task, compared to others that have reported improved orientation sensitivity

(Goldreich & Kanics, 2003; J. C. Stevens et al., 1996; Van Boven et al., 2000; Wong et al., 2011), is that the angle of difference in orientation was manipulated, instead of the grating groove width (see also similar method and result by Alary et al., 2009). Therefore, in Gori and colleagues' (and Alary et al., 2009) task, participants were asked to indicate which of two stimuli (a standard fixed at 45° or a comparison varying between 0°-90°) was more slanted, whereas in the other tasks, participants were asked to identify whether the first or second tactile stimulus contained the horizontally (or vertically) orientated grating. As a result, Gori and colleagues' task measures the ability to discriminate small differences in slant, whereas the other tasks measure the tactile acuity needed to recognise or identify a specific orientation pattern. Thus, it appears that visually deprived individuals show enhanced tactile acuity driven by experience-dependent plasticity, but impaired orientation sensitivity due to the important role that vision plays in calibrating touch for such judgments. More research will be needed to understand whether blind adults can show normal orientation sensitivity with practice, thereby compensating for the absence of vision, considered fundamental for calibration.

1.3.5 How is compensatory plasticity mediated?

Vision appears essential for calibrating certain non-visual cues on specific spatial tasks, however blind individuals show similar or improved performance relative to sighted controls on other tasks too. Compensatory behaviour following blindness has been linked to neural reorganisation, including notably the recruitment of primary visual cortex by non-visual modalities. However, it is not clear whether this plasticity is mediated via the reinforcement of existing connections – and multisensory connections may be particularly relevant here (see section 1.2.6) – or the formation of new synapses, driven by bottom-up or top-down mechanisms.

Evidence from animal studies indicates that the recruitment of primary visual cortex by residual senses could be mediated by (i) a reorganisation in subcortical activity that, for example, enables non-visual modalities to enter visual cortex through connections via the thalamus (e.g. Izraeli et al., 2002), and/or (ii) changes to cortico-cortical connectivity, for example direct projections from auditory cortex to primary visual cortex have been revealed in primates (Falchier et al., 2010) which may drive auditory recruitment of visual cortex following blindness (see review by Bavelier &

Neville, 2002). Within visual cortex, changes in connectivity may be the result of various mechanisms including local sprouting, unmasking of silent synapses and/or changes in existing connections. Of particular relevance to the latter, studies have reported auditory- and tactile- evoked activity in the neurons of the primary visual cortex of sighted animals and humans (e.g. Martuzzi et al., 2007; Spinelli, Starr, & Barrett, 1968) and evidence indicates that this reflects the activity of multisensory neurons (Fishman & Michael, 1973; Murray et al., 2015; see section 1.2.6). Hence, these multisensory neurons may be 'taken over' by non-visual inputs following visual deprivation. Indeed, animals reared in darkness show an increase in the number of neurons that respond to non-visual modalities in multisensory areas (Carriere et al., 2007; Hyvarinen, Carlson, & Hyvarinen, 1981; Wallace, Perrault, et al., 2004; see section 1.4.1), which is thought to reflect the 'takeover' of visual sections of multimodal areas by non-visual inputs, mediated by activity-based competition between different inputs. In line with this, where visual cortex is recruited by non-visual inputs, it often retains the same function, for example ventral stream areas recruited for non-visual identification tasks and dorsal stream areas for spatial localization tasks (Striem-Amit, Dakwar, Reich, & Amedi, 2012). Activity-based competition can drive plasticity during early development, but it is unclear whether this is still possible later in life. Indeed, compensatory plasticity has been found to vary depending on the onset (early vs. late) of the visual loss, with more profuse neural plasticity seen following early visual loss (e.g. review by Sathian, 2005).

It is not yet clear how the compensatory plasticity seen following visual deprivation is mediated, but this will likely depend on factors including the multimodal nature of certain regions, the onset and type of visual deprivation, the functionality of the region or compensatory behaviour, and cross-modal training.

1.3.6 Summary

Many studies have reported that blind individuals show behavioural and neural changes in non-visual (particularly auditory, tactile) processing (e.g. Buchel et al., 1998; Gougoux et al., 2005; Voss et al., 2004). These compensatory changes often enable blind individuals to perform at least comparably to normally sighted individuals in some perceptual tasks. However, vision does seem important for the acquisition of certain non-visual spatial processing capabilities, including auditory spatial bisection and tactile orientation sensitivity (though not navigational tasks such as path reproduction or path completion). The next section reviews whether vision is important for the acquisition of multisensory processing abilities that involve combining information from non-visual senses.

1.4 Multisensory Processing following Blindness

Many decisions that are made in daily life can benefit from the combination of multisensory information. In particular, humans can use redundant sensory information about a specific event to improve the speed, precision and/or accuracy of their behavioural response (section 1.2). Following blindness, individuals no longer have information from the visual sensory modality that they can combine with non-visual sensory information, but they could still in principle combine information from different non-visual modalities. However, visual experience may be essential for learning correspondences between non-visual sensory cues, and therefore for acquiring the capacity to integrate non-visual sensory information. Moreover, given compensatory neural changes (see section 1.3), it may be that multisensory neurons deprived of visual input become preferentially responsive to a specific non-visual input (i.e. become unisensory). Findings from animal and human research, (reviewed below), suggest that visual experience plays a fundamental role in the acquisition of multisensory interactions.

1.4.1 Reduced multisensory interactions in early blind

As noted previously (section 1.2.7), Stein and colleagues have conducted many studies investigating the multisensory processing of neurons in the cat superior colliculus, which have been influential in understanding both the process and the development of multisensory integration. In terms of the development of multisensory integration, studies have found that multisensory neurons in the cat superior colliculus develop gradually after birth (Stein, Labos, & Kruger, 1973; Wallace & Stein, 1997). Specifically, during the first days of postnatal life, all neuronal responses within the superior colliculus are unimodal, and when multisensory responses do appear, these do not show adult-like multisensory integration behaviour until several weeks later (Wallace & Stein, 1997). Whilst the superior colliculus of Rhesus monkeys already has many multisensory neurons at birth, these similarly do not show multisensory integration capabilities (Wallace & Stein, 2001; Wallace et al., 1996). These findings suggest that sensory experience is essential for the development of multisensory integration capabilities within the superior colliculus. In particular, some evidence indicates that visual deprivation in early life can result in permanently impaired multisensory integration capabilities (Carriere et al., 2007; Wallace, Perrault, et al., 2004).

Wallace, Perrault, et al. (2004) examined the sensory responses of superior colliculus neurons in adult cats that had been deprived of visual experience, (having been reared in darkness from birth). Compared to cats reared under normal lighting conditions, visually deprived cats showed a reduced incidence of neurons that responded specifically to vision, and an increase in auditory-specific and somatosensory-specific neurons. A similar incidence of neurons responded to stimulation from more than one sensory modality (albeit slightly reduced), and over 90% of these were visually responsive. However, the visual, auditory and somatosensory fields of these multisensory neurons, though topographically organised, were large, indicating the fundamental role of vision for their spatial calibration. Most importantly, although multisensory neurons responded robustly to each of their unisensory inputs when presented individually, their responses were not substantially enhanced when multiple spatially- and temporally- aligned stimuli were presented (unlike in control animals). Instead, multisensory responses were no different to unisensory responses, and this was evident for both non-visual (i.e. auditory-somatosensory) and visual multisensory neurons. Carriere et al. (2007) found similar neuronal properties in the AES (a multisensory cortical area that sends information to the superior colliculus) of dark-reared cats. Specifically, dark-reared and normally reared cats showed a similar incidence of both visually responsive (unisensory visual neurons and visually responsive multisensory neurons) and multisensory neurons. However, as found in the superior colliculus, a considerably reduced proportion of multisensory neurons showed response enhancements to multisensory stimulation. Interestingly, multisensory neurons in the AES tended to show response depression (i.e. smaller responses to multisensory than unisensory stimulation), and again this was evident in both visual and non-visual multisensory neurons. Thus, these findings suggest that visual experience is necessary for the development of multisensory integration within the superior colliculus, including non-visual auditory-tactile integration. Presumably, this reflects the importance of visual experience for learning that common events in the physical world occur in close spatial correspondence, and thereby driving the receptive fields of multisensory neurons to adjust accordingly.

Few studies have investigated whether the combination of auditory and tactile information is modified in blind humans, however those that have suggest that auditory-tactile interactions are reduced (see reviews by Hotting & Roder, 2009; Occelli, Spence, & Zampini, 2013). For example, Occelli et al. (2012) found congenitally blind adults were less susceptible to an audio-tactile ventriloquist effect (see section 1.2.5), than late blind and sighted adults. Participants were asked to report the perceived location of a sound (left, right or centre) presented at the same time as a tactile stimulus to the left or right hand. Late blind and sighted participants tended to perceive the auditory stimulus as being located toward the concurrent tactile stimulus more consistently than the congenitally blind group. Furthermore, whilst all participants showed a reduced audio-tactile ventriloquism effect when making judgments with hands crossed relative to the body midline, the reduction was significantly greater in the congenitally blind group. In a similar task, Collignon, Charbonneau, Lassonde, and Lepore (2009) asked early blind, late blind and (blindfolded) sighted controls to decide as quickly and as accurately as possible, whether auditory (100ms pink noise bursts), tactile (pulses delivered to middle fingers) or audio-tactile stimuli occurred on the left or the right. All groups showed better performance in the bimodal condition and bimodal reaction time reductions exceeded probability summation predictions (see section 1.2.1) for all groups when both hands were uncrossed too. However, whereas sighted and late blind also showed bimodal reaction reductions that exceeded probability summation predictions when hands were crossed, early blind participants did not. Thus, spatial audio-tactile interactions in early blind participants appear impaired, particularly when hands are crossed relative to the body midline. Collignon, Charbonneau, et al. (2009) suggested that this is due to early blind participants' dependence on a body-centred reference frame for encoding spatial events. Specifically, they argued that since audition is externally referenced and touch is body-centred, in order to combine auditory and tactile events into a common percept, in sighted and late blind participants tactile information is remapped to an externally defined reference frame. However, this remapping does not occur in early blind participants, and consequently the crossed posture results in a spatial conflict between auditory and tactile events, thereby preventing multisensory integration (see also Gori et al., 2014, section 1.3.1).

Hotting, Rosler, and Roder (2004) used electroencephalography to measure event-related potentials (ERPs) whilst congenitally blind and sighted participants responded to auditory or tactile stimuli. Participants were asked to attend to one modality (auditory or tactile) at one spatial position (left or right) and respond to deviant stimuli of that modality and position, (by pressing a foot switch), as accurately and as quickly as possible. There was no difference in error rates between blind and sighted groups, but reaction times to tactile stimuli were significantly faster for blind than sighted participants. In blind participants, somatosensory and auditory ERPs showed a more pronounced negativity to stimuli presented at the attended side, than the unattended side, when attending specifically to touch or audition respectively. Sighted participants' ERPs similarly showed more pronounced negativity to stimuli at the attended side (starting 80ms after stimulus onset), but this was irrespective of the stimulus being attended, thereby showing both an early unimodal and cross-modal spatial attention effect. However, at later processing stages (>200ms after stimulus onset) a cross-modal spatial attention effect (defined by a more pronounced positivity to stimuli at the attended side, irrespective of sensory modality attended) was observed in blind, but not sighted, participants. It was concluded that blind participants initially filtered information by modality only, (whereas sighted participants used both modalities), and at later stages suppressed task-irrelevant stimuli at the attended location, thereby showing reduced auditory-tactile interactions.

As vision is the most accurate sense for spatial perception and, therefore, plays an important role in calibrating the spatial maps of audition and touch, it is perhaps not surprising to find that early blind individuals show reduced auditory-tactile spatial interactions. In contrast, since audition tends to dominate temporal judgments (see section 1.2.5), it is reasonable to expect that the absence of vision would not influence audio-tactile temporal interactions. However, Hotting and Roder (2004) found that congenitally blind individuals were less susceptible to an auditory-tactile temporal illusion (see section 1.2.5) than sighted (seeing/blindfolded) individuals. Specifically, when a single tactile stimulus (a light touch from a metallic pin to the right index finger) was presented together with more than one task-irrelevant sound (a tone from two loudspeakers), all participants reported perceiving more than a single touch, however this illusion was significantly more pronounced in sighted than congenitally blind individuals. Visual loss may influence auditory-tactile temporal judgments, due to compensatory changes in residual senses and associated neural reorganisation. In particular, studies have shown that blind individuals have superior auditory (e.g. Muchnik, Efrati, Nemeth, Malin, & Hildesheimer, 1991; A. A. Stevens & Weaver, 2005) and tactile (e.g. Wan, Wood, Reutens, & Wilson, 2010) temporal perception, (and superior abilities have also been documented for some spatial tasks too, see section 2). Consequently, following visual loss, compensatory changes in auditory and tactile senses may affect combined audio-tactile processing in temporal, and also spatial, tasks.

1.4.2 Reduced multisensory interactions in late blind

Animal and human research indicates that experience of visual and non-visual multisensory events is necessary for the development of multisensory integration (see section 1.4.1), however another question is whether vision is necessary for the maintenance of multisensory integration capabilities. As discussed (see section 1.3), visual deprivation can lead to compensatory plasticity, which may have implications for multisensory interactions. Increasing research is documenting differences between early and late blind individuals in both compensatory behaviour and plasticity (e.g. Fieger et al., 2006; Sadato, Okada, Honda, & Yonekura, 2002; Tao et al., 2013; Voss, Gougoux, Zatorre, Lassonde, & Lepore, 2008; Voss, Pike, & Zatorre, 2014). Specifically, findings indicate that late blind individuals show reduced compensatory behaviour and recruitment of visual structures for non-visual tasks, compared with early blind individuals (e.g. Sadato et al., 2002; Voss et al., 2014), and at least some compensatory behaviours have been shown to be mediated by different mechanisms in late compared to early blind participants (e.g. Fieger et al., 2006; Tao et al., 2013). Therefore, if compensatory plasticity and/or lack of visual experience are causing the reduced multisensory interactions observed in the blind, it might be expected that late blind participants would show normal – or at least – less impaired audio-tactile interactions.

As reviewed above (section 1.4.1), Occelli et al. (2012) and Collignon, Charbonneau, et al. (2009) found that early blind but not late blind participants showed reduced audio-tactile interactions. Few researchers have studied audio-tactile interactions in blind humans, and even fewer have investigated differences between early and late blind participants. However at least one study has found that late visual loss can lead to changes in non-visual combination: Using the auditory-tactile parchment-skill illusion, Champoux et al. (2011) found that most early blind individuals showed no illusory change in tactile perception when the frequency of an auditory signal was modified, unlike sighted individuals who consistently reported that their palm skin was drier or moister according to variations in the audio sound intensity for certain frequencies (see section 1.2.5). Importantly, however, four of eight late blind individuals similarly showed no susceptibility to the illusion. Thus, these findings suggest that auditory and somatosensory interactions are also impaired in some late blind individuals, for some tasks. It is expected that late blind individuals would have developed multisensory integration capabilities via early visual experience (in accordance with animal studies, see section 1.4.1). Therefore, reduced audio-tactile interactions appear to suggest that vision may be necessary to maintain these abilities. However, although neurons in animals show significant response enhancements to multisensory stimuli early in development, certain multisensory processing abilities do not develop in humans until much later in childhood (see section 1.2.8). For example, children under eight years are unable to integrate multisensory cues to reduce uncertainty in perceptual decisions, preferring to either rely on one sensory cue or to alternate between cues (Gori, Sandini, et al., 2012; Nardini et al., 2008). Hence, it cannot be assumed that late blind individuals will have normally developed multisensory processing capabilities. Finally, to interact effectively with the environment, humans must maximise the accuracy of their behavioural responses. Doing so is complex, requiring continual recalibration (see section 1.2.3), and consequently, a lack of vision either early or late in development may have implications for this process, which may in turn affect multisensory interactions. More research is needed to understand how multisensory processing is affected following early and late visual loss, and if it is affected, why this is the case.

1.4.3 Summary

Findings from studies investigating multisensory processing in animals and humans suggest that vision is fundamental for acquiring normal multisensory interactions for certain tasks (see section 1.4.1). Animal studies have found that multisensory neurons in the superior colliculus, involved in orienting the eye and head to stimuli in the environment, do not develop normal multisensory integration capacities in the absence of vision. Human studies have found that early blind individuals show reduced auditory-tactile interactions, relative to sighted and late blind individuals on speeded reaction time tasks, or tasks testing their susceptibility to auditory-tactile illusions. Whilst late blind individuals perform similarly to sighted on some of these tasks, some evidence suggests that some late blind individuals might show impaired auditory-tactile interactions on some tasks too (see section 1.4.2). However, research on auditory-tactile interactions in blind humans is limited, and as yet, researchers have not assessed the ability of blind humans to combine auditory and tactile cues to improve the precision or accuracy of their perception in line with ideal observer models. Some multisensory integration capabilities are acquired early in life (including speeded reactions and susceptibility to certain illusions), and therefore these may not be affected in late blind individuals, whereas more complex multisensory interactions that develop with extensive visual and non-visual experience may be. Hence, here the impact of permanent visual deprivation on multisensory processing has been reviewed. The next section, instead, explores whether visual experience later in life, for example following treatment to restore the visual sense, can lead to normal visual, non-visual and multisensory processing.

1.5 Visual, Non-Visual and Multisensory Processing following Visual Treatment

The existing research reviewed thus far indicates that vision plays an important role in calibrating non-visual representations of space, and consequently visual deprivation can negatively impact the processing of non-visual and multisensory information. However, at least for some tasks, compensatory changes in non-visual modalities can enable blind individuals to perform comparably to typically sighted individuals. The neural reorganisation associated with compensatory behavioural changes in residual senses following visual loss may have implications for sensory restorative treatments (e.g. retinal prostheses). Hence, a key question is whether visual experience later in adulthood is sufficient to enable the development of normal visual, non-visual and multisensory processing in blind individuals.

1.5.1 Visual processing following visual treatment

Much research highlights the importance of uninterrupted early visual stimulation for the development of certain aspects of normal vision. In particular, visual stimulation is crucial during ‘critical periods’, defined as optimal temporal windows for the development of a particular sensory function. For example, patient M.M. and patient S.B. both lost their sight early in development, and received visual treatment as adults. Patient M.M. was blinded at 3.5 years, and received a corneal transplant treatment in his right eye at 43 years (Fine et al., 2003). Patient S.B. lost sight aged 10 months but received a corneal transplantation after 50 years of blindness (Gregory & Wallace, 1963). Both patients showed typical simple colour, form and motion processing, however more complex functions including complex form, object and face recognition were severely impaired, and visual acuity and contrast sensitivity remained severely compromised. Additionally, patient M.M. showed reduced fMRI BOLD responses to spatial frequency gratings in visual cortex. These results suggested that if visual stimulation were not present during the development period of certain visual functions (the critical period), such functions would never develop later, after treatment. Similar findings have been reported in patients treated for bilateral cataracts (e.g. Maurer, Mondloch, & Lewis, 2007; Putzar, Hotting, Rosler, & Roder, 2007). For example, Maurer et al. (2007) studied children born with dense central bilateral cataracts that were removed between one month and one year after birth. They found that whilst some visual abilities recovered completely, including for example sensitivity to high temporal frequencies and face detection, others, such as holistic face processing and sensitivity to high (but not low) spatial frequencies, showed severe lasting deficits. The inability to recover certain abilities, including sensitivity to high spatial frequencies and holistic face processing, was particularly interesting because these functions would normally manifest at a later period in infancy than the affected period (i.e. later than 1 year). These “sleeper effects” may reflect the need for vision earlier in infancy to set up, preserve and/or avoid inhibition/plasticity of the neural architecture needed for these visual functions.

In contrast, some studies have shown recovery of visual functions following visual deprivation during critical periods (Kalia et al., 2014; Ostrovsky, Andalman, & Sinha, 2006). For example, patient S.R.D. who was born blind and did not undergo surgery for the removal of dense congenital cataracts until age twelve years, was found to perform at a high level on form and face perception tests twenty years after surgery, despite compromised visual acuity (Ostrovsky et al., 2006). These results suggest that the visual system can retain considerable plasticity, allowing for the acquisition of visual functions following visual experience, despite visual deprivation during critical periods. However, some patients in these studies may have had residual visual functions beyond bare light perception prior to treatment (Kalia et al., 2014), and interestingly, patient S.R.D. did show some qualitative differences in her performance compared to normally sighted participants. For example, she relied on head orientation rather than eye position when making gaze direction judgments, indicating that she used different strategies to perform these visual tasks (Ostrovsky et al., 2006).

1.5.2 Non-visual processing following visual treatment

Whilst studies have explored the effect of visual restoration on visual treatment, less is known about the impact of visual restoration on non-visual processing. Much research has documented compensatory changes in non-visual processing following blindness (section 1.3). In particular, non-visual processing has been found to recruit typically visual processing areas, linked to superior performance on some non-visual tasks. Hence, for visual treatment to be successful, visual processing areas will need to learn to respond preferentially to visual, as opposed to non-visual, inputs. Recently, retinal prostheses have been developed that attempt to restore vision to patients blinded by retinal degenerative diseases by electrically stimulating retinal cells. Cunningham, Tjan, Bao, Falabella, and Weiland (2015) studied the effect of visual restoration on cross-modal responses in primary visual cortex (V1), in two late blind adults implanted with a retinal prosthesis. One participant, who had been implanted for six weeks, showed similar V1 responses to tactile stimulation as nine late blind participants (with only minimal light perception, who had not undergone prosthetic treatment). In contrast, the other participant had been implanted for 15 weeks and their V1 responses were comparable to those of nine typically sighted adults and nine visually impaired adults (with partial vision loss). Moreover, for both participants, increased V1 responses to tactile stimulation were found following a period of not using the retinal implant device. These results indicate that compensatory plasticity following visual loss can eventually be reversed by visual experience (see also Dormal et al., 2015), but also, strikingly, that compensatory plasticity in response to visual deprivation can occur reasonably quickly (see also Merabet et al., 2008).

Recently, Heimler et al. (2014) suggested that rehabilitation programs could use cross-modal training to drive plasticity following treatment, by pairing re-acquired with recruited inputs on multisensory tasks. For example, since blind adults show activation of occipital regions when using touch to recognise shapes (e.g. lateral occipital cortex, Amedi, Raz, Azulay, Malach, & Zohary, 2010), they propose that adults undergoing visual treatment should be encouraged to explore objects using both visual and tactile modalities together, because this may eventually drive the corresponding brain regions to respond preferentially to visual, instead of tactile, inputs (see also Isaiah, Vongpaisal, King, & Hartley, 2014).

1.5.3 Multisensory processing following visual treatment

As per visual processing in humans (section 1.5.1), animal studies investigating the acquisition of multisensory integration capabilities following visual deprivation, pointed toward a critical period for their development. For example, Royal, Krueger, Fister, and Wallace (2010) found that the superior colliculus neurons of cats deprived of vision from birth (by rearing in darkness) until adulthood, failed to develop normal spatiotemporal receptive fields (see also Carrasco & Pallas, 2006) and multisensory responses, suggesting that early visual experience is essential for the development of multisensory integration capabilities. However, Yu, Rowland, and Stein (2010) found that similarly reared cats were able to develop multisensory integration capabilities following exposure to spatially and temporally congruent visual and auditory stimuli. They suggested that exposure to spatiotemporally congruent stimuli elicits enhanced multisensory responses in neurons via Hebbian learning rules. Moreover, neurons accomplished multisensory sensitivity much more quickly than predicted based on their normal developmental chronology, perhaps partly due to the intense exposure the animals had to audio-visual events, but also due to the existence of a sufficiently mature underlying neural substrate (see e.g. Rowland, Jiang, & Stein, 2014).

These findings suggest that humans treated for early and late visual impairments should be able to acquire normal multisensory integration abilities with sufficient experience of multisensory stimuli. Roder and colleagues assessed multisensory processing in human adults (15-48 years) who had been deprived of pattern vision in the first months of life due to binocular congenital cataracts. In a reaction time task, patient and control groups showed similarly reduced response times to bimodal (auditory-tactile, auditory-visual and tactile-visual) stimuli than unimodal stimuli, (that exceeded race model predictions, see section 1.1; Putzar, Gondan, & Roder, 2012). However, reduced multisensory interactions were shown in other more complex tasks that involved the suppression of task-irrelevant tones or the combination of audio-visual cues to make language decisions (Putzar, Goerendt, Lange, Rosler, & Roder, 2007; Putzar, Hotting, & Roder, 2010). For example, patients asked to discriminate words in audio-, visual- or audio-visual format performed worse in audio-visual conditions than sighted participants (Putzar, Goerendt, et al., 2007). Differences in multisensory performance on different tasks may reflect differences in the development of certain multisensory functions (see section 1.2.8) and differences in the underlying circuitry involved. The ability to use multisensory stimuli to speed up responses may develop independently of sensory input, or be acquirable later in life, and not restricted to a critical period in infancy. In contrast, integration of more complex stimuli for making discriminatory decisions may depend on multisensory input in early years or substantial experience with specific cues for certain judgments.

Similarly to the findings in individuals treated for binocular congenital cataracts, Moro, Harris, and Steeves (2014) assessed the audio-visual localization performance of adults who had undergone monocular enucleation during childhood. The authors explained that individuals with one eye might show similar compensatory plasticity as that demonstrated following complete blindness, (since in a previous study they had shown superior accuracy in a binaural sound localization task; Hoover, Harris, & Steeves, 2012), which may affect their audio-visual combination capabilities. However, participants with one eye showed similar audio-only and visual-only discrimination precision as normally sighted controls, and combined auditory and visual cues to location in accordance with optimal predictions, (see section 1.2.2). Although these findings suggest that visual and multisensory processing is not affected in individuals with one eye, in a separate study the authors found that people with one eye showed a reduced McGurk effect (see section 1.2.5) compared to normally sighted controls (Moro & Steeves, 2015). Hence, as discussed above, differences in multisensory performance on different tasks may reflect differences in the critical and sensitive periods of development for certain multisensory functions.

1.5.4 Multisensory processing following auditory treatment

Much research has similarly studied multisensory capabilities in deaf individuals treated with a cochlear implant (a small device that can be surgically implanted into a person's cochlea and produces hearing sensations by electrically stimulating nerves inside the ear). Following deafness, as in blindness, compensatory plasticity has been shown, with visual and tactile stimulation activating auditory cortical regions (e.g. Finney et al., 2001; Schurmann, Caetano, Hlushchuk, Jousmaki, & Hari, 2006). Interestingly, many studies have shown that deaf individuals treated with cochlear implants are able to combine auditory information with visual lip movements to improve speech processing (Kaiser, Kirk, Lachs, & Pisoni, 2003; Moody-Antonio et al., 2005; Tremblay, Champoux, Lepore, & Theoret, 2010; Tyler et al., 1997). However, cross-modal plasticity has been found to influence the hearing ability of cochlear implant users (Buckley & Tobey, 2011; Doucet, Bergeron, Lassonde, Ferron, & Lepore, 2006; Lee et al., 2001), and correspondingly audio-visual interactions have been found to depend on the proficiency of the cochlear implant user (Champoux, Lepore, Gagne, & Theoret, 2009; Landry, Bacon, Leybaert, Gagne, & Champoux, 2012). In particular, whereas proficient cochlear implant users show normal audio-visual interactions, less proficient users show impairments in tasks that involve segregating auditory and visual information, argued to be due to a strong preference for visual cues and greater activation of auditory cortex by visual stimulation. Accordingly, studies using the McGurk effect (see section 1.2.5) have found multisensory perception in cochlear implant users to be dominated by vision (Desai, Stickney, & Zeng, 2008; Rouger, Fraysse, Deguine, & Barone, 2008; Schorr, Fox, van Wassenhove, & Knudsen, 2005), though increased reliance on vision could reflect greater uncertainty in auditory, as opposed to visual information, as predicted by Bayesian Decision Theory models (section 1.2.2).

Fewer studies have investigated audio-tactile (compared to audio-visual) interactions in deaf and cochlear implant users. Using the audio-tactile illusory flash effect (see section 1.2.5), Landry, Guillemot, and Champoux (2013) found that the tactile sensations of cochlear implant users (both early- and late- deaf) were not influenced by auditory information, unlike individuals with normal hearing. However, when using the parchment-skin illusion (see section 1.2.5), Landry, Guillemot, and Champoux (2014) found that individuals with little cochlear implant experience performed similarly to normal hearing controls, whereas experienced cochlear implant users showed a significantly greater illusory percept. Hence, in deaf individuals, both the extent of cross-modal plasticity and the amount of cochlear implant experience appear to influence multisensory interactions.

1.5.5 Summary

Existing research suggests that the success of restorative visual treatments for acquiring typical visual, non-visual and multisensory processing capabilities may depend on (i) the age at onset of visual deprivation and the duration of deprivation (which likely affect the extent of resulting cross-modal plasticity), as well as (ii) the age at visual treatment and the extent of visual experience following treatment (which, in turn, likely affect the extent of any ‘reversal’ in cross-modal plasticity). Some evidence suggests that there are critical periods for the development of certain visual processing capabilities (see section 1.5.1), and it is possible that, similarly, critical periods exist for the acquisition of certain non-visual and multisensory processes. However, some research suggests that exposure to non-visual and multisensory stimuli may be sufficient to reverse any non-visual cross-modal plasticity and develop multisensory integration capabilities (e.g. Yu, Stein, & Rowland, 2009).

1.6 The Impact of Partial Visual Loss

Notably, much of the research exploring cross-modal plasticity following visual loss (and reviewed up until now) has studied non-visual processing in blind individuals, for whom the extent of the visual loss is total or severe. Consequently, less is known about any changes to the non-visual or multisensory processing of individuals with partial or degenerative sight. One possibility is that even partial sight loss may lead to cross-modal reorganisation, which (as has been discussed) could have implications for non-visual and multisensory processing. This section will review existing research that has studied non-visual and multisensory processing following partial visual loss. Note that, as mentioned earlier (see section 1.3), there are many causes of visual impairment, including problems in the eye (e.g. cataracts, glaucoma, macular degeneration) and the brain (e.g. stoke, prematurity, trauma). More research is needed to understand how these different causes may impact any compensatory behaviour or plasticity.

1.6.1 Non-visual processing following partial visual loss

As mentioned above, whilst non-visual processing by blind individuals has received much research attention, less is known about the non-visual processing of individuals with partial sight. A recent study, however, suggests that gradual visual loss could lead to gradual cross-modal reorganisation in the brain. Specifically, Cunningham, Weiland, Bao, and Tjan (2011) found that blindfolded patients diagnosed with retinitis pigmentosa (a retinal degenerative eye disease) showed increased activation of visual cortex in response to tactile stimuli, compared to blindfolded sighted participants, and individuals with greater visual loss showed higher visual cortex activation. Interestingly, results also indicated that the specific location of the visual loss in the visual field correlated with the location of tactile-evoked responses in the visual cortex. Hence, patients with visual loss may not only have to account for changes in the reliability of their vision, but also changes in the reliability of non-visual cues, and moreover, such changes may be specific to where the impairments are in their visual field. An interesting implication is that such ‘gradual cross-modal plasticity’ may be promoted through cross-modal training, including for example the use of sensory substitution devices. Sensory substitution devices convert information from the substituted modality (typically vision) into another modality (typically touch or audition) that can then be interpreted. Paul Bach-y-Rita and colleagues (1969) introduced the idea that people deprived of one sense, such as sight, could regain access to that missing information if it were transformed into a format that another intact sense could process; stating “we see with our brain, not with our eyes”. Importantly, training with visual substitution devices has been shown to lead to non-visual evoked activation of the visual cortex (e.g. De Volder et al., 1999; Ortiz et al., 2011). It is not clear whether visual cortex activation reflects cross-modal recruitment or instead the use of mental (visual) imagery strategies (see review by Poirier, De Volder, & Scheiber, 2007), and further research is needed to differentiate these. However, if indeed sensory substitution devices are found to promote cross-modal reorganisation, non-visual processing abilities of individuals with partial vision loss may benefit from their use.

Research into the use of sensory substitution devices has typically focused on their ability to benefit perception in the absence of vision, studying either blind or blindfolded sighted participants. However, it would also be interesting to consider whether these devices could be used together with residual vision to improve the speed, precision or accuracy of perception. Specifically, as reviewed earlier (section 1.2), humans can combine redundant multisensory information to improve their perception. Often visual disease does not lead to total blindness, but instead can reduce the reliability of the visual sense non-uniformly across the visual field. Hence, combining this residual visual information with non-visual (auditory or tactile) information from a sensory substitution device could lead to perceptual benefits.

1.6.2 Multisensory processing following partial visual loss

Combining redundant information from multiple senses can lead to improvements in perception (section 1.2). Some studies have explored the use of multisensory processing in rehabilitation. For example, Keller and Lefin-Rank (2010) assigned patients with visual field defects to either an audio-visual or visual-only training program focused on improving visual search. Both groups showed improved visual search performance after eye movement training, but a greater improvement was seen in the group that had received audio-visual training, suggesting that the auditory sense could be used to train (or calibrate) the biased visual sense (see section 1.2.3). Other studies have noted the improvements in performance afforded by multisensory information, relative to unisensory information alone. For example, Frassinetti, Bolognini, Bottari, Bonora, and Ladavas (2005) asked patients with either a visual field deficit (hemianopia) or a visuospatial attentional deficit (e.g. neglect) to detect visual stimuli presented alone or together with an auditory stimulus. Despite being instructed to ignore the auditory stimulus, patients' visual detection was significantly improved by the presence of auditory stimuli, but only when the auditory stimulus was presented in a similar spatial position (within 16 degrees) as the visual target. Similarly, Targher, Occelli, and Zampini (2012) investigated whether auditory information could improve visual detection in patients with deteriorating vision (showing reductions in visual field and/or visual acuity). Participants fixated straight ahead, whilst visual (100ms flash from a green light emitting diode) and/or auditory (100ms white noise bursts) stimuli were presented across their visual field (0 ± 56 degrees). They were asked to detect the presence of visual stimuli only (ignoring auditory stimuli). Visual data was analysed to determine which parts of the visual field showed the most impairment. Results indicated that a simultaneous auditory stimulus (presented in the same location or displaced by up to 16 degrees), significantly improved the detection of visual stimuli in the most impaired visual field positions (compared to visual-only detection performance), but not the less impaired positions where performance for most participants was at ceiling (99% correct).

The results of these studies are in line with much research indicating that human adults can use multiple cues to improve their performance on multisensory tasks (see section 1.2.1 and 1.2.2). However, studies have not explored whether adults with low vision use vision in combination with redundant non-visual information to improve their bimodal performance. In order to combine visual and auditory information to improve bimodal precision, low vision adults must weight sensory cues according to their relative reliability (section 1.2.2). Therefore, low vision adults must account for any changes to the reliability of their vision, across their visual field. Whilst, normally sighted adults have been shown to account for immediate changes in the relative reliabilities of two cues (manipulated experimentally, see section 1.2.2), the effect of a gradual and long-term change on the reliability of a sensory cue is not clear. Some evidence investigating the influence of aging on navigation, suggests that gradual deterioration in the reliabilities of visual and vestibular cues, can lead to the suboptimal weighting of visual-vestibular cues (Bates & Wolbers, 2014).

1.6.3 Summary

Whilst much research has explored changes to non-visual processing following blindness (see section 1.3), less is known about compensatory changes in individuals with partial vision loss. Some initial evidence suggests that cross-modal plasticity may occur gradually as specific parts of the visual field become deprived of visual input, and might be promoted through cross-modal training, but further research is needed. Similarly, only a few studies have assessed how individuals with partial visual loss combine multisensory information to make perceptual decisions. Those that have indicate that multisensory information can be beneficial for perception, but it is unclear whether changes to the visual sense have impacted how different sensory information is relied upon in multisensory decisions.

1.7 Thesis Overview

The impact of temporary blindness, compensatory plasticity and partial visual loss on non-visual and multisensory processing is not yet well understood. Much research has studied how blind humans and animals process non-visual information, and findings have indicated that visual deprivation can lead to compensatory plasticity (see section 1.3). Yet, the mechanisms involved in mediating such plasticity are unclear. One possibility is that cross-modal reorganisation reflects changes in the activity of multisensory neurons in multisensory areas, which could impact how sensory information is combined. In line with this, animals deprived of vision show an increase in non-visually responsive neurons (relative to normally-sighted animals) in multisensory areas. Similarly, some evidence indicates reduced multisensory interactions in early-, and also some late-, blind individuals (see section 1.4). Regardless of the exact mechanisms involved, compensatory plasticity has clear implications for treatments that aim to restore vision by stimulating the visual system directly. Typically visually responsive areas that have become recruited by non-visual senses may need to re-learn to respond preferentially to visual inputs. Some evidence suggests that visual experience is sufficient to 'reverse' any compensatory plasticity, though, depending on the age and duration of visual loss, certain visual functions may never be re-acquired (see section 1.5). On the other hand, multisensory experience could be sufficient to lead to the normal development of multisensory processing abilities, though studies investigating multisensory interaction in humans following visual treatment have found mixed results (see section 1.5.3). Visual treatments, such as the retinal prosthesis, provide an opportunity to investigate how restored vision impacts perception on multisensory tasks. Since many daily tasks involve combining information from multiple modalities, assessing whether (and how) individuals use restored vision in combination with non-visual information is an important measure of the effectiveness of such treatments in improving perception. Moreover, failures to combine sensory information to improve perception in line with optimal observer models may be indicative of underlying changes within multisensory regions of the brain. To this end, Chapters 2 and 3 will study how blind individuals implanted with a retinal prosthesis use prosthetic vision in combination with non-visual information on multisensory tasks.

Whilst much research has studied the impact of total vision loss on perception, less is known about the effect of partial vision loss. Most cases of vision loss occur gradually and individuals often maintain some residual vision. Some studies have noted the benefits of presenting non-visual with visual information for improving the perception of humans with low vision (see section 1.6.2). However, surprisingly, the question of whether individuals account for changes in the relative reliability of visual versus non-visual senses, in their multisensory decisions, has not been assessed. One possibility is that changes to the visual sense impact the ability to combine multisensory information. This may be particularly relevant for spatial information, since vision is considered fundamental in aligning the spatial representations of non-visual modalities. Another possibility is that the ability to combine multisensory information is preserved, but the nervous system does not account optimally for long-term gradual changes to the reliability of the visual sense. Chapter 4 will assess how individuals experiencing gradual visual loss combine visual and non-visual cues to location.

Whilst the ability to combine visual and non-visual information can improve multisensory perceptual judgments, another important issue is how the nervous system calibrates sensory information to maximize accuracy. In normally sighted adults, the visual sense provides the most accurate and reliable spatial information, and therefore plays an important role in calibrating non-visual spatial representations. Consequently, visual loss could be expected to negatively impact the spatial localization of non-visual senses, yet research has found that (early and late) blind individuals often perform at least comparably to sighted individuals on horizontal localization tasks, and this has been linked to compensatory plasticity. However, less is known about the influence of partial visual loss on non-visual spatial localization. Some evidence suggests that reduced visual reliability could lead to difficulties in differentiating non-visual systematic spatial errors from visual random errors, and therefore slower calibration of non-visual space (see section 1.2.3). In addition to information from sensory representations, humans can rely on prior knowledge about their environment when making perceptual decisions. Interestingly some perceptual biases have been shown to reflect reliance on prior knowledge about the statistics of the natural environment (section 1.2.2). Chapter 5 studies whether prior knowledge could influence auditory localization: firstly, we assess whether a well-known auditory localization bias can be explained by reliance on prior knowledge that does not accurately represent the testing environment, and secondly whether visual feedback can be used to reduce the auditory localization bias, irrespective of the visual feedback reliability.

Chapter 2

The Speed and Precision of Multisensory Perception following Visual Treatment: Visual-Haptic Size Discrimination and Speeded Visual-Auditory Target Detection by Adults with a Retinal Implant

2.1 Introduction

2.1.1 Retinal Implants

Recently retinal implants (or ‘retinal prostheses’) are being developed that can restore vision to people that have been blinded by retinal degenerative diseases, by stimulating preserved cells within the retina (see Fig. 4). Currently, these devices are limited in the visual acuity that they can afford, providing only ‘ultra low vision’. Assessments are needed to measure the impact of these treatments on the quality of life of their users (see for e.g. Geruschat et al., 2015). Therefore, since many everyday activities are multisensory, it is informative to quantify any improvements afforded by retinal implants on multisensory tasks. In addition, by studying the multisensory processing of individuals receiving retinal implant treatment, it is possible to further understanding of any compensatory mechanisms triggered by late-onset visual loss and subsequent visual restoration, which could in turn inform future rehabilitation strategies. Chapters 2 & 3 will describe experiments conducted with patients that were blinded by a retinal degenerative disease (retinitis pigmentosa or choroideremia), before being implanted with the Argus II retinal prosthesis system. In these experiments, we assessed whether patients could use the visual input from the retinal implant to improve the speed, precision or accuracy of their multisensory perception.

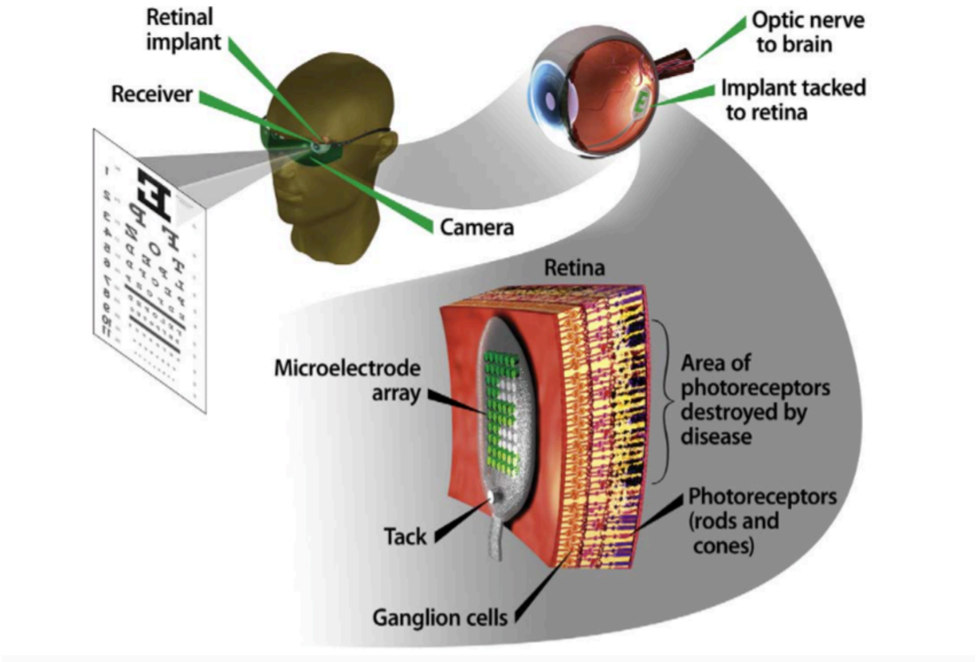


Figure 4: Diagram of retinal prosthesis Components.

External images are captured by a miniature camera and sent to a video processing unit that converts the data to an electronic signal, which is sent to a receiver and then to the microelectrode implant tacked to the retina. The retina is a tissue containing millions of photoreceptor cells that convert light into electrical signals, which are then sent via the optic nerve to the brain for interpreting the physical world. However, before reaching the optic nerve, these signals pass via several different cells including inner retinal (horizontal, bipolar and amacrine cells) and ganglion cells. In certain retinal degenerative diseases, including retinitis pigmentosa, despite considerable photoreceptor death, many of these other cells are preserved. Retinal prostheses stimulate these preserved retinal cells directly, thereby fulfilling the role of the lost photoreceptors. (Figure taken from Chader, Weiland, & Humayun, 2009).

The Argus II retinal prosthesis system (Second Sight Medical Products, Inc., Sylmar, CA) consists of a glasses-mounted miniature camera that sends live video data to an externally worn processing unit that transforms it into electrical stimulation patterns. These patterns are sent wirelessly to an implant on the retina, (a 6 x 10 electrodes epiretinal array secured over the fovea), directly stimulating preserved retinal cells (see Fig. 4). Using the Argus II system, patients who have been visually deprived for a number of years are once again able to receive visual input. However, prosthetic vision is different to native vision: Patients have to learn to interpret the pixelated phosphenes elicited by the implant, and, since the direction of 'gaze' is defined by head position (not by eye position) and the field of view (11 x 18 degrees) is limited, patients must learn to explore the environment by using head scanning movements.

The Argus II system received the CE mark in Europe in 2011 and FDA approval in the US in 2013, following an ongoing clinical trial involving thirty blind patients (28 diagnosed with retinitis pigmentosa, 1 choroideremia, 1 Leber congenital amaurosis) implanted at multiple sites worldwide. All patients had no measurable visual acuity prior to implantation (as assessed using a grating visual acuity test, with a 2.9 logMAR test floor, that involved differentiating the orientation of black and white bars of a range of widths), but all had some level of bare light perception prior to surgery (to ensure integrity of the pathway from the retina to the visual cortex). Following implantation, their visual acuity was assessed again, (using the same grating visual acuity test), and seven patients showed a measurable acuity below the 2.9 logMAR test limit with the prosthesis (Humayun et al., 2012). This corresponded to resolving the least coarsest grating tested (13 degrees), yet still worse than the theoretical resolution (4 degrees) achievable with the prosthesis (see Stronks & Dagnelie, 2014 for details).

The visual acuity afforded by this prosthesis was further assessed using two computer tasks that involved localizing a white (11 x 11 degrees) square (Ahuja et al., 2011) or identifying the direction of motion (7.9 degrees/sec – 31.6 degrees/sec) of a white bar (Dorn et al., 2013) on a computer screen. On these tasks most patients were found to perform more accurately and reliably with the prosthesis (‘system on’) than without the prosthesis (‘system off’), (though far less patients showed improvements in the direction-of-motion compared to square-localization task, suggesting difficulties perceiving moving stimuli). To assess the ‘real-world utility’ afforded by the prosthesis, two additional tests were used that involved (i) finding a (3 x 7 ft) door within a (20 x 20 ft) room and (ii) following a (6 in x 20 ft) white line on the floor. Again, patients performed more accurately with the prosthesis in both these tasks (Humayun et al., 2012). Although these latter tasks were designed to assess ‘real-world utility’, both involve (primarily) relying on vision alone to simply home towards a direct landmark (door or line). In most real-word tasks, however, it is often possible to use non-visual information too. It is unclear whether the vision afforded by the prosthesis would enable patients to improve their performance on tasks for which they could rely on other senses.

2.1.2 Speed and Precision of Multisensory Perception

Many of the perceptual decisions that humans make, such as crossing the road or making a cup of tea, are multisensory, in that information from multiple senses can be used. As reviewed (see sections 1.2.1-1.2.2), by combining (redundant) sensory information, human adults can improve the speed and/or reliability of their perception (e.g. Ernst & Banks, 2002; Miller, 1982; Nardini et al., 2008). Research has shown that visual experience is necessary to calibrate non-visual representations of space and acquire multisensory combination abilities (e.g. Gori, 2015; Wallace, Perrault, et al., 2004; see sections 1.2.8 & 1.4.1). Hence, it is possible that the late-onset visual loss experienced by retinal implant patients may have impacted multisensory processes that rely on continual recalibration and knowledge of changing relative sensory cue reliabilities. Additionally, following (typically early-, but also late-) visual loss, visual processing areas can become recruited by non-visual modalities (e.g. Burton et al., 2002; Collignon et al., 2013; Merabet et al., 2008; see section 1.3). The success of restorative visual treatments, including retinal implants, will therefore depend partly on the extent of any compensatory plasticity and the ability of visual and multisensory experience to ‘reverse’ any compensatory plasticity.

Behavioural studies that have assessed multisensory processing in human adults treated for early visual deprivation have found mixed results (see section 1.5.3). For example, whereas patients showed reduced reaction times to bimodal stimuli compared to unimodal stimuli (Putzar et al., 2012), they showed reduced multisensory interactions in more complex tasks that involved, for example, combining audio-visual cues to make language decisions (Putzar, Hotting, et al., 2007). Hence, it may be that the ability to use multisensory stimuli to quicken responses develops independently of sensory input or is easily acquirable later in life, whereas integration of more complex stimuli depends on multisensory input in early years or extensive experience later in life (see also section 1.2.8).

2.1.3 Experiment Aims

Here, we report the results of initial experiments conducted that studied the visual and multisensory processing of late-blind adults, deprived of vision for over 15 years before being implanted with the Argus II retinal prosthesis. The aims were to assess whether Argus II users were able to combine visual and non-visual information to (i) improve precision on a discrimination task and (ii) improve reaction times on a detection task, in line with the predictions of an ideal observer. We compared any predicted and measured improvements afforded by the prosthesis on these tasks to understand whether late-onset visual deprivation and subsequent treatment influences the ability to combine sensory information optimally. These initial experiments were limited by sample size, (given the limited number of patients that had been implanted with the Argus II prosthesis at Moorfields Eye Hospital), however they were nonetheless useful in understanding the constraints of prosthetic vision and designing future experiments. Note that, unfortunately, due to the limited sample size, it was not possible to assess whether the age of onset of disease or time since implantation affected sensory combination abilities.

2.2 General Method

2.2.1 Ethics Statement

Patients were recruited from the Moorfields Eye Hospital NHS Foundation Trust, London, UK, and the study had received ethical approval from the East Central London committee. Informed written consent, according to the Tenets of the Declaration of Helsinki, was obtained from all participants prior to participation.

2.2.2 Participants

Five adults aged 49-76 years (M = 64.2, SD = 10.5; 4 male) implanted with Second Sight's Argus II retinal prosthesis in their right eye, at Moorfields Eye Hospital, as part of the Argus II feasibility study, participated. All participants had been implanted in their right eye (as this was their worst eye) 3-6 years prior to testing. All had been diagnosed with a retinal degenerative disease (participants

002-005 retinitis pigmentosa, participant 001 choroideremia) prior to implant, and reported having been blind for 15-52 years. All patients had bare light perception (to ensure the optic nerve was functional), but visual acuity worse than 2.9logMAR prior to surgery. Following surgery, two scored reliably on a visual acuity test with the implant (2.8 & 2.9 logMAR). Patients had received visual rehabilitation training provided by Second Sight Inc. prior to this study that covered basic skills like head scanning.

Table 1: Clinical Details for Participants

ID	Age (yrs)	Sex	Diagnosis Age (yrs)	Years blind (pre-implant)	Post-surgery acuity (logMAR)	Functioning electrodes
001	70	M	46	51	>2.9	93%
002	49	F	11	15	>2.9	100%
003	59	M	7	36	>2.9	100%
004	67	M	28	25	2.9	88%
005	77	M	19	52	2.8	47%

2.3 Experiment 1: Visual-haptic task

2.3.1 Method

2.3.1.1 Apparatus & Stimuli

The visual and haptic stimuli used were nine white wooden balls that differed in diameter by 2 mm, ranging either from 41-57 mm (set 1, participants 001, 002, 003 & 004) or 49-65 mm (set 2, participant 005). Participant 005 used an overall bigger set of balls compared to the other participants, to assess whether absolute ball size affected discrimination performance (see Appendix A.2.1). These stimuli were chosen as they had been used to measure haptic size discrimination thresholds in a previous experiment with sighted children and adults (Petrini, Remark, et al., 2014). Each set consisted of one standard ball (set 1: 49mm, set 2: 57mm) and eight comparison balls. Balls were presented on a black rectangular foam surface. Matlab

(Version R2010a, The MathWorks Inc., Natick, Massachusetts, United States) was used to control stimuli presentation order and store participant responses.

2.3.1.2 Procedure

Participants were seated comfortably in front of a desk, covered with a black cloth, on which the black rectangular foam surface had been positioned. They were asked to place each hand on one side of the foam surface to familiarise themselves with the spatial position in which the balls would be presented. On each trial, the experimenter placed two balls (the standard and a comparison) on either side of the foam surface. Once both balls were in position, the participant was asked to indicate which of the two balls was bigger by using either unisensory (visual-only, haptic-only) or bimodal (visual-haptic) information. On visual-only trials, participants used head-scanning movements to align the Argus II system camera (mounted on the glasses) with the position of each ball, and then interpreted the pixelated phosphenes elicited by the implant for each ball. All participants wore an eye patch over the non-implanted (left) eye, to ensure that only their prosthetic vision was assessed. On haptic-only trials, participants tapped each ball once, using the flat palm of their dominant hand, whilst wearing a blindfold. On bimodal trials, participants used both visual and haptic information to decide which ball was bigger. The experiment consisted of fifteen blocks of sixteen trials: five vision-only blocks, five haptic-only blocks, and five visual-haptic blocks. Block order was pseudo-randomised so that each consecutive block involved a different (visual-only, haptic-only or visual-haptic) sensory cue. The position of the standard and comparison ball on the foam surface was counterbalanced across blocks. The experimenter recorded participants' responses and no feedback was provided. Before starting the experiment, participants completed a short practice task (see Appendix A.2.2 for details).

2.3.1.3 Data Analysis

The proportion of trials in which the comparison ball was perceived as bigger than the standard ball was plotted as a function of the size difference between the balls, for each sensory cue (visual, haptic, visual-haptic). Data were fitted with cumulative Gaussian functions, using psignifit 2.5.6 for Matlab (see http://bootstrap-software.org/psignifit/), a software package which implements the maximum-likelihood method described by Wichmann and Hill (2001a). The standard deviation (σ) and the mean (μ) of each function provided estimates of the cue's reliability and point of subjective equality (PSE), respectively. Functions were fitted to each individual participant's data. Unisensory variances were used to compute the estimate with the lowest possible variance, (assuming early independent noise, identically distributed likelihood functions and a uniform prior), known as the maximum likelihood estimate (MLE; see section 1.2.2, Eq. 6 & 7). Measured bimodal discrimination reliability was compared to the optimal predictions (MLEs) and measured unisensory (visual-only, auditory-only) discrimination reliabilities. Due to the small sample size ($N = 5$), the use of paired-sample *t*-tests to assess significance of planned comparisons is not advised (see e.g. review by de Winter, 2013). With such small samples it is not possible to assess whether the assumptions underlying the t-test are met, and the probability of Type I and Type II errors are high. Consequently, whilst the results of paired t-tests have been reported here, these should be interpreted with caution.

2.3.2 Results & Discussion

Figure 5 shows the unisensory (visual-only, haptic-only) and bimodal psychometric functions and corresponding reliabilities (σ) for the five adults tested. All participants showed poor visual-only reliability, with values exceeding the largest size difference tested (8 mm). Participant 005 performed at chance when using only vision to discriminate ball size (see Appendix A.2.4) and the psychometric fit to their visual-only data is a near-straight horizontal line indicating that they did not show better discrimination performance for larger ball size differences. Therefore, participant 005 was omitted from the group analysis, although doing so did not affect the overall result.

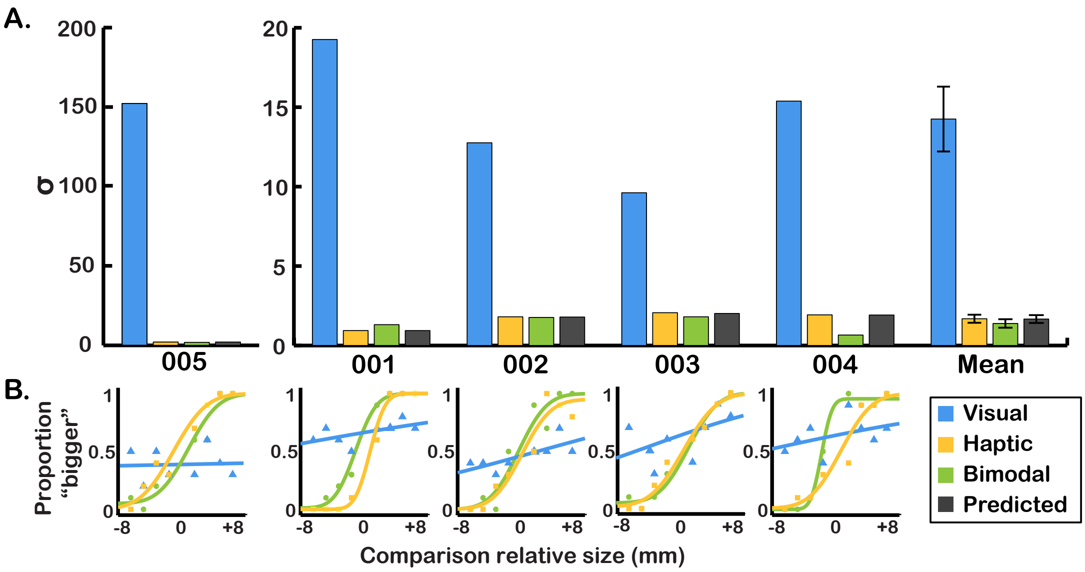


Figure 5: Unisensory (visual-only, haptic-only) and bimodal sigma (A) obtained from psychometric functions (B). Error bars represent the standard error of the mean.

As shown (Fig. 5), for all participants, visual-only discrimination performance was particularly impaired relative to discrimination performance in haptic-only and bimodal conditions (σ on average ~8.5 times greater for visual than haptic discrimination – excluding 005). Accordingly, visual-only judgments were significantly less reliable than haptic-only ($t_{[3]}$ = 5.55, p = 0.012) and bimodal judgments ($t_{[3]}$ = 5.84, p = 0.010). No significant difference was observed between haptic-only and bimodal discrimination reliability ($t_{[3]}$ = 0.85, p = 0.456) or predicted and measured bimodal discrimination reliability ($t_{[3]}$ = 0.80, p = 0.482; though note that due to the small sample size, statistical test results should be interpreted with caution, see section 2.2.3.3). The maximum likelihood estimate (MLE) combines unisensory estimates weighted by their relative reliability to produce the most reliable bimodal estimate. As visual information was highly unreliable relative to haptic information, for discriminating ball size, the MLE predicted that combining visual and haptic information would not measurably benefit bimodal performance. In line with this prediction, participants' bimodal judgments were not more reliable than when relying on haptic information alone. However, note that this result is also consistent with participants ignoring the visual information entirely, and using haptic information alone.

2.4 Experiment 2: Visual-auditory task

2.4.1 Method

2.4.1.1 Apparatus & Stimuli

Stimuli presentation was controlled using Matlab (Version R2014a, The MathWorks Inc., Natick, Massachusetts, United States) and the Psychophysics toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997), on an Apple MacBook Pro computer running OS X 10.9. The visual stimulus was a 1,000 msec presentation of a white screen (on an otherwise black screen) displayed on a 21-inch iiyama monitor (1280 x 800 px screen resolution, 60 Hz refresh rate), subtending a vertical visual angle of 34 degrees. The auditory stimulus was a 500 msec pure 1000 Hz tone, presented via two Logitech speakers positioned 15cm behind the monitor. The bimodal (audio-visual) stimulus was the presentation of both the visual and the auditory stimulus together. Following testing, an oscilloscope was used to obtain a measure of the onset asynchrony of the visual and auditory stimuli on bimodal trials: results indicated that auditory stimuli preceded visual stimuli by 14 msec.

2.4.1.2 Procedure

Participants were seated comfortably in front of a desk, covered with a black cloth, on which a keyboard, the iiyama monitor and the two Logitech speakers were positioned. They were asked to locate the space key on the keyboard, and maintain their index finger over this key throughout the experiment. A chin rest was used to fix their head position so that Argus II system camera (mounted on the glasses) was directly aligned with the monitor. As in Experiment 1, all participants wore an eye patch over the non-implanted (left) eye, to ensure that only their prosthetic vision was assessed. During the experiment, participants were asked to respond as quickly as possible to audio-only, visual-only or audio-visual stimuli, by pressing a space key as soon as any of these stimuli were detected. The experiment consisted of fifteen stimuli-specific blocks of 24 trials: 5 audio-only blocks, 5-visual only blocks, and 5 audio-visual blocks. Block order was pseudo-randomised so that each consecutive block involved a different (audio-only, visual-only or audio-visual) stimulus. Prior to the start of a new block, a distinctive auditory sound was played. Following stimulus presentation, participants were allowed up to 1,000 msecs to respond, (a total of 2,000 msecs from stimulus onset). If no response was registered during this period, the trial was then terminated and deemed a miss. The interval between a participant's response and the onset of the next stimulus was set to vary randomly between 1,000 and 1,400 msecs throughout the experiment. Before starting the experiment, participants were presented a visual-only, auditory-only and visual-auditory stimulus and asked to confirm whether they were able to perceive all three stimuli. A short practice, consisting of three blocks (1 audio-only, 1 visual-only, 1 audio-visual) of five trials, was completed to familiarize participants with the task.

2.4.1.3 Data Analysis

Reaction times to each different stimulus type (audio-only, visual-only or audio-visual) were recorded and the mean reaction time to each stimulus type was calculated. Percentages of misses were below 7% in all modalities for all individuals, and so were not further analysed. Where bimodal reaction times exceeded those of the best unisensory cue, redundancy gains were measured as the difference between the mean reaction times to the bimodal stimulus and the faster of the two unisensory stimuli. Faster responses to bimodal stimuli indicate that participants are using both visual and auditory signals together, either by processing each signal independently and responding to the signal that finished processing first (statistical facilitation), or by processing both signals pooled together (sensory integration). The maximum redundancy gain predicted by statistical facilitation, assuming statistically independent unisensory response latencies and no increase in noise in bimodal conditions, is given by the summed distributions of the unisensory stimuli (Miller's inequality, Miller, 1982; also referred to as the race model inequality). To assess whether participants could have improved reaction times to bimodal stimuli by processing both cues and responding to the faster of these ('the race model'), reaction time cumulative distribution functions (CDFs) were computed for unisensory stimuli, and the 'race model prediction' was computed as the sum of these unisensory CDFs. For participants that showed a bimodal reaction time advantage, the race model prediction was compared with bimodal cumulative reaction time distributions, to assess whether redundancy gains exceeded those predicted by statistical facilitation.

2.4.2 Results & Discussion

Figure 6 shows participants' mean reaction times to the unisensory (visual-only, auditory-only) and bimodal stimuli. When using the retinal prosthesis to detect light flashes (Fig. 6.A), all participants showed slower mean reaction times to visual-only stimuli than auditory-only and bimodal stimuli. Notably, participant 002's mean reaction time to visual-only stimuli was over four times slower than their mean reaction time to auditory-only stimuli. A bimodal reaction time advantage through statistical facilitation (the race model prediction) was predicted only for participants 003 and 004 (see Appendix A.2.5), whose reaction times to visual and auditory stimuli were the most closely matched in the group. Accordingly, only participants 003 and 004 showed faster reaction times to bimodal than unimodal stimuli, with redundancy gains of 54 and 34 msecs respectively.

Three participants also completed the experiment again without the prosthesis (system off), relying only on their residual vision (Fig. 6.B). Participant 005 was unable to perceive the visual stimulus without the prosthesis, and participant 003 opted not to participate. Performance was similar to that observed with the prosthesis for participant 001 and 002, in that responses to auditory stimuli were faster that responses to visual-only stimuli. However, participant 002 responded much quicker to visual-only stimuli without the prosthesis. Participant 004 showed faster reaction times to both visual and auditory stimuli but no longer showed a reaction time advantage to bimodal stimuli, (though mean bimodal reaction times were faster with the system off: 366 msecs vs. 413 msecs).

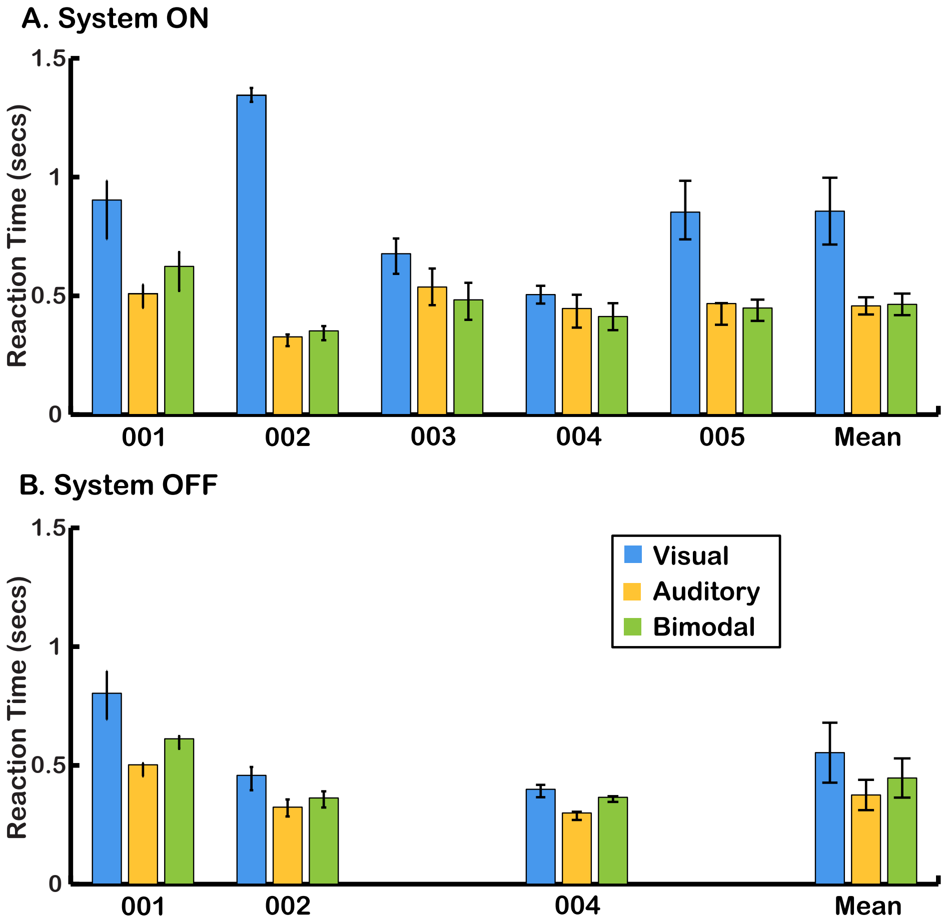


Figure 6: Unisensory (visual-only, auditory-only) and bimodal mean reaction times with the prosthesis (system on, A) or with residual vision (system off, B). Participant error bars represent the interquartile range (25 and 75 percentiles). Mean error bars represent the standard error. Participant 003 opted not to take part in the experiment with the system off. Participant 005 was unable to perceive the visual stimulus with the system off.

As participants 003 and 004 showed a reaction time advantage to bimodal stimuli, their data was further analysed to test for a race model violation. The race model inequality, (computed as the sum of the visual-only and auditory-only distributions), was violated by both participant 003 and participant 004 (see Fig. 7). The areas of violation, computed as the difference between the area under the bimodal reaction time CDF and the summed unisensory reaction time CDF, were 86 msecs and 75 msecs for participants 003 and 004 respectively. This violation area corresponds to 32% (participant 003) and 39% (participant 004) of the sum of the area underneath both curves. Results of a bootstrap analysis, in which the experiment was simulated 1000 times using reaction time values sampled from each participant's reaction time data (without replacement), indicated that a violation of the race model occurred on all 1000 simulated experiments for each participant. Violation area values obtained from the simulation ranged from between 19 and 197 msecs (8-53%) and 37 and 110 msecs (21–54%), for participants 003 and 004 respectively.

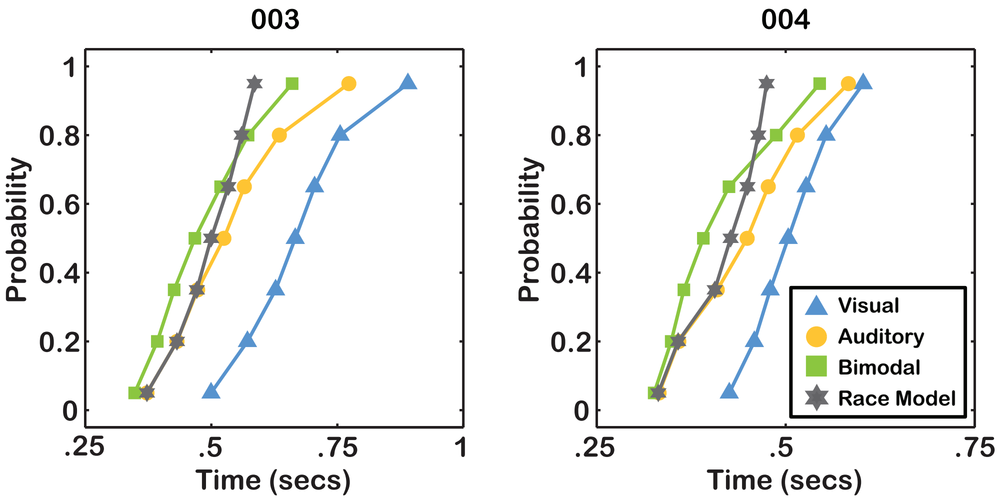


Figure 7: Reaction time cumulative distribution functions (CDFs) to unisensory and bimodal stimuli for participants 003 and 004. Bimodal reaction times for participants 003 and 005 exceeded race model predictions between the 5th and 65th percentiles.

As mentioned above, participants that showed a bimodal reaction time advantage also showed reaction times to unisensory stimuli that were most closely matched in this participant group. The other three participants (001, 002, and 005) showed mean reaction times to visual-only stimuli that were over 70% slower than their mean reaction times to auditory-only stimuli. These results suggest that for at least three participants tested there may have been a delay in processing the visual information, either reflecting a delay in transferring the information from the Argus II camera to the retinal implant, and/or a delay in interpreting the phosphenes elicited. No bimodal reaction time advantage was predicted by the race model, due to this large difference between reaction times to auditory and visual stimuli. Similarly, if stimuli were not perceived as occurring simultaneously, auditory and visual signals would not be pooled together and processed in combination (sensory integration). Therefore, asynchronies in visual and auditory stimuli perception may explain why participants 001, 002 and 005 did not show redundancy gains. To measure any perceived systematic delay between the visual and auditory stimuli, a third experiment was conducted that assessed each participant's sensitivity to visual-auditory asynchrony.

2.5 Experiment 3: Visual-auditory follow-up

2.5.1 Method

2.5.1.1 Apparatus, Stimuli & Procedure

As in Experiment 2, participants were seated comfortably in front of a desk, covered with a black cloth, on which the iiyama monitor and the two Logitech speakers were positioned. A chin rest was used to fix their head position so that Argus II system camera (mounted on the glasses) was directly aligned with the monitor, and all participants wore an eye patch over the non-implanted (left) eye, to ensure that only their prosthetic vision was assessed. On each trial, participants were presented an auditory and a visual stimulus together (audio-only and visual-only stimuli as described in Experiment 2), however the onset of each stimulus was manipulated, so that auditory-visual cue onsets were either synchronous (both auditory and visual stimuli presented together) or asynchronous (audio-leading by 333, 300 or 67 msecs, or visual-leading by 333, 300 or 67 msecs)*. They were then asked to make either a simultaneity judgment (SJ), by deciding whether the audio and visual stimuli occurred at the same time, or a temporal order judgment (TOJ), by deciding whether the audio or visual stimulus occurred first. The experiment consisted of twenty blocks of seven trials (one trial per cue-onset asynchrony level): ten blocks of simultaneity judgments and ten blocks of temporal order judgments. Block and cue-onset asynchrony were randomised. The experimenter recorded participants' responses and no feedback was provided. Before starting the experiment a short practice, of six trials (three SJ, three TOJ), was completed to familiarize participants with the task. *Note that, as in Experiment 2, results of a timing test indicated that auditory stimuli preceded visual stimuli by 14 msec.

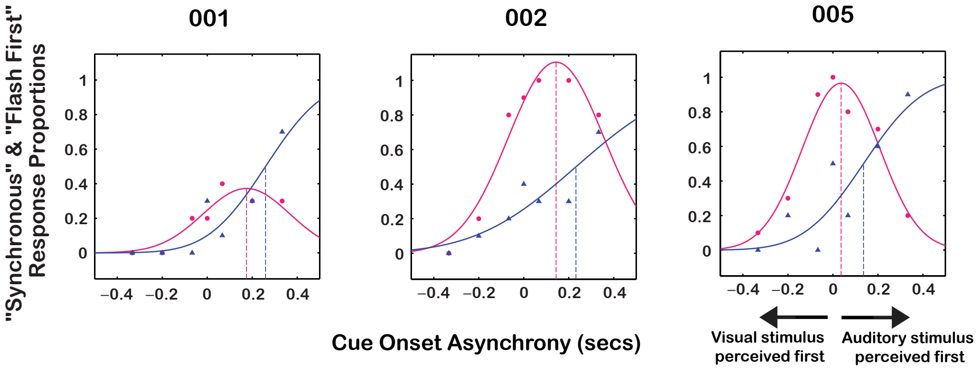
2.5.1.2 Data Analysis

For simultaneity judgments, the proportion of responses that were judged to be synchronous were plotted against the respective cue onset asynchronies, and Gaussian probability density functions (PDF) were fit to these data. For temporal order judgments, the proportion of responses in which the visual stimulus was judged to occur first was plotted against the respective cue onset asynchronies, and Gaussian cumulative density functions (CDF) were fit to these data. The point of subjective simultaneity (PSS) represents the cue onset asynchrony level at which visual and auditory stimuli are perceived to occur simultaneously. The PSS for simultaneity judgments was computed as the maximum of the SJ PDF, and the mean of the TOJ CDF. The temporal integration window (TIW) represents the range of onset times at which asynchronies (for SJ) or cue order (for TOJ) cannot be reliably perceived. PDF and CDF standard deviations were taken as a measure of the TIW for SJs and TOJs respectively.

2.5.2 Results & Discussion

Figure 8 shows each participant's SJ and TOJ responses. For each individual, differences in PSSs on SJ and TOJ tasks were observed. Such differences in PSSs across SJ and TOJ tasks have been frequently reported (see reviews by Keetels & Vroomen, 2012; van Eijk, Kohlrausch, Juola, & van de Par, 2008). It is thought that these may reflect differences in the assumptions observers make when completing simultaneity and temporal order judgment tasks. Specifically, in the temporal order judgment task, participants may assume that the stimuli are never simultaneous since only temporal order responses are given, whereas in the simultaneity judgment task participants may be inclined to assume that stimuli are simultaneous since the response must be either that synchrony was present or absent. Hence, SJ and TOJ tasks may measure different perceptual mechanisms (Love, Petrini, Cheng, & Pollick, 2013): SJ tasks measuring multisensory binding, and TOJ tasks measuring temporal discrimination instead.

A. Participants who showed no bimodal reaction time advantage:



B. Participants who showed a bimodal reaction time advantage:

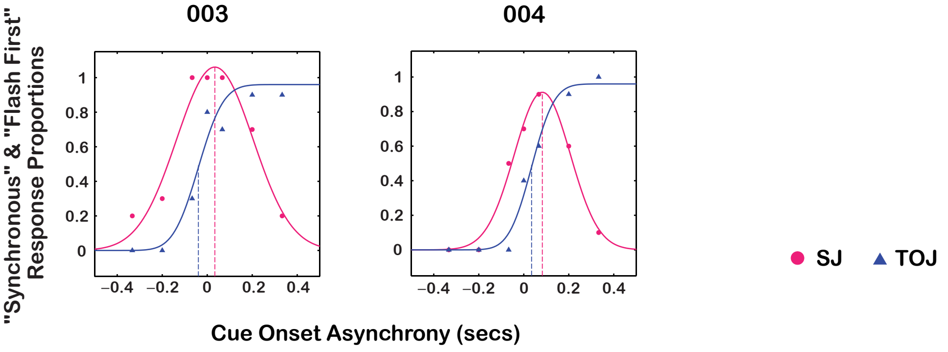


Figure 8: Gaussian probability and cumulative density functions fitted to synchrony (SJ) and temporal order judgments (TOJ). Dotted lines represent points of subjective simultaneity (PSS). Participants 003 and 004 showed a bimodal reaction time advantage in experiment 2.

Table 2: Points of Subjective Simultaneity (PSS) and Temporal Integration Windows (TIW) for Synchrony Judgments (SJ) and Temporal Order Judgments (TOJ) obtained from Probability and Cumulative Density Functions fitted to Participant Data. A timing test conducted post-experiment indicated that auditory stimuli had preceded visual stimuli by 14 msec. To account for this discrepancy, 14 msecs should be subtracted from the PSS values presented here.

	Synchrony Judgments (SJ)		**Temporal Order Judgments (TOJ)**	
	PSS (msecs)	**TIW (msecs)**	**PSS (msecs)**	**TIW (msecs)**
001	175	192	259	203
002	144	212	232	353
003	34	170	-39	88
004	83	124	35	80
005	36	173	136	208
Sighted*	**-50 to +150**	**188 (9.6)**	**-73 to +75**	**146 (13.4)**

* PSS values for sighted participants taken from review by van Eijk et al. (2008). Values represent the minimum and maximum PSSs reported for individual participants on 22 TOJ and 10 SJ tasks using flash-click stimuli. TIW values taken from Love et al. (2013). Values represent the mean (standard error) TIW for flash-beep SJ and TOJ judgments.

Despite differences in measured PSSs, for both simultaneity and temporal order judgments PSSs indicated that all participants perceived visual and auditory stimuli to be synchronous when visual stimuli preceded auditory stimuli by between 34 and 259 msecs (Table 2; 20 and 245 msecs if accounting for the 14 msec discrepancy measured during the timing test post-testing). Participant 003 was the only exception, for whom the measured TOJ PSS indicated that visual-auditory synchrony was maximal when auditory stimuli preceded visual stimuli by 39 msecs (53 msecs if accounting for the 14 msec discrepancy). Studies that have measured the PSSs of normally sighted adults have tended to find visual-leading PSSs in SJ tasks, but auditory-leading PSSs in TOJ tasks (Love et al., 2013; van Eijk et al., 2008). SJ and TOJ PSSs in normally sighted adults have been reported in the ranges of -50 and +150 msecs and -73 and +75 msecs respectively (see review by van Eijk et al., 2008). Measured TOJ PSSs for participants 001, 002 and 005 fell outside of the range reported for normally sighted adults. Specifically, visual stimuli were perceived as occurring before auditory stimuli, only when visual stimuli preceded auditory stimuli by at least 136 msecs. This suggests that the (physical or neural) processing of visual information by these participants may have been delayed by at least 61 (=136–75) msecs. Participants 003 and 004 (Fig. 8.B), who had shown a bimodal reaction time advantage in experiment 2, perceived maximal visual-auditory synchrony at smaller cue onset discrepancies that fell within the ranges reported in normally sighted adults.

To perceive visual-auditory events in the physical world as synchronous, the human brain must account for differences in the processing timing of visual and auditory stimuli. Light travels faster through air than sound (300,000,000 m/s vs. 300 m/s), but neural processing is typically slower for visual than auditory stimuli (approximately 50 msecs vs. 10 msecs, Keetels & Vroomen, 2012). Despite these physical and neural processing delays, humans tend to perceive synchrony for most visual-auditory events in the physical world. Moreover, the human brain allows for variation in such processing delays, since signals that are not temporally aligned may still be perceived as synchronous and processed in combination, provided that they fall within a temporal integration window (TIW). TOJ TIWs to beep-flash stimuli in normally sighted adults have been reported to be narrower than SJ TIWS (146 msecs vs. 188 msecs, Love et al., 2013). Here, participants 003 and 004 similarly showed narrower TIWs for TOJs than SJs, whereas participants 001, 002 and 005 showed wider TIWs for TOJs, that were at least 57 (=203–146) msecs wider than the mean TIW previously reported in normally sighted adults (Love et al., 2013).

In summary, compared to participants 003 and 004, (who showed faster reaction times to bimodal stimuli in experiment 2), participants 001, 002 and 005 perceived synchrony in temporal order judgments at larger visual-auditory discrepancies (larger PSSs), and were less sensitive to cue onset asynchronies for temporal order judgments (wider TIWs). PSSs and TIWs for TOJs by participants 003 and 004 were in line with those that have been previously reported for normally sighted adults (Love et al., 2013; van Eijk et al., 2008), whereas for participants 001, 002 and 005 both PSS and TIW values tended to be larger. The findings suggest that for three participants there may have been a delay in processing the visual information afforded by the retinal prosthesis, reflecting either a (physical) delay in transferring the information from the Argus II camera to the retinal implant, or a (neural) delay in interpreting the phosphenes elicited. Therefore, discrepancies in the perceived onset of visual and auditory stimuli may explain why participants 001, 002 and 003 did not show reduced reaction times to bimodal stimuli in Experiment 2.

2.6 General Discussion

The present experiments assessed whether five late-blind adults, implanted with the Argus II retinal prosthesis, were able to combine visual and non-visual information to improve either the precision or speed of their perceptual decisions. Experiment 1 tested whether Argus II users could combine prosthetic vision with haptic information to improve size discrimination judgments. Results indicated that for all five participants prosthetic vision was much less reliable than touch, and consequently, there was no predicted or measured benefit of using vision in combination with touch on this task. Experiment 2 assessed whether Argus II users could use prosthetic vision and auditory information to improve the speed of their behavioural responses to visual-auditory stimuli. Findings showed that for three participants there was no predicted benefit of processing visual and auditory stimuli in parallel, according to probability summation, due to a large delay reacting to visual-only compared to auditory-only stimuli. Two participants (who showed smaller delays to visual-only stimuli) were predicted to benefit from processing visual and auditory information in parallel, and did so. Importantly, however, their speed gains exceeded those predicted by probability summation (the race model), indicating co-active processing of visual and auditory information. Finally, to measure any perceived systematic delay between the visual and auditory stimuli (that might have explained the delayed reaction time to visual-only stimuli observed in Experiment 2), Experiment 3 assessed each participant's sensitivity to visual-auditory asynchrony. The two participants who had shown faster bimodal reaction times in Experiment 2, showed similar auditory-visual synchrony perception as normally sighted adults, whereas the other participants showed non-normal perceived synchrony of temporal order judgments. Unfortunately, as we were only able to test five patients implanted with the retinal prosthesis, it was not possible to assess how individual differences in factors such as age of blindness, duration of blindness and/or number of working electrodes in the implant, impacted the results of these experiments. This will be interesting for future research to address, as more patients become implanted with the prosthesis.

2.6.1 The Limitations of Restored Vision

The results of Experiment 1 showed that, unlike haptic information, the visual information afforded by the retinal prosthesis was insufficiently sensitive to inform reliable size discrimination judgments in this task. Consequently, the ability to combine visual and haptic information to improve the precision of perceptual judgments on this task was limited by the spatial resolution of the restored vision. Similarly, studies with deaf individuals treated with cochlear implants suggest that their ability to acquire normal audio-visual interactions depends on their cochlear implant proficiency (Champoux et al., 2009; Landry et al., 2012). It may be that with increased practice and/or future technological developments, the prosthesis will afford better spatial resolution, and users will, consequently, show perceptual benefits of using visual and non-visual information together.

In addition to limitations in spatial resolution, the results of Experiment 2 and 3 suggest that the prosthesis may be limited in temporal resolution too. Specifically, participants showed delayed reactions to visual stimuli, compared to auditory stimuli, and perceived visual and auditory stimuli to be maximally synchronous when visual stimuli preceded auditory stimuli by up to 259 msecs. Therefore, it may be that there is a physical delay in transferring the visual information captured by the Argus II camera to the retinal implant. Alternatively, however, the result may be driven by a neural delay in processing and interpreting the phosphenes that are elicited by the implant stimulation. Participants who also completed the reaction time task using residual (instead of prosthetic) vision showed faster reactions to stimuli perceived using residual vision than prosthetic vision, which also more closely matched their reaction times to auditory stimuli, indicating that there is not a delay in processing visual information per se.

2.6.2 The Effect of Late-Onset Visual Deprivation on Multisensory Processing

In Experiment 1, we found that the visual information afforded by the retinal implant was not sufficiently reliable to benefit size discrimination performance in Experiment 1. Consequently, it is not possible to infer from the results of this task whether, following a period of late-onset visual deprivation, human adults are still able to combine sensory information to improve the precision of discrimination judgments, as per normally sighted adults (e.g. Ernst & Banks, 2002). Importantly, however, visual information was not detrimental to the size discrimination performance of the participants assessed. This indicates that these participants were at least not over-weighting vision in their multisensory judgments, but, (in line with the predictions of an ideal observer), relying on the more reliable haptic information instead. This could, however, equally reflect a disregard of the visual information and a reliance on haptic information, without considering the relative reliability of both cues. Therefore, to assess whether participants were truly weighting visual and non-visual information according to their relative reliability, it would be necessary to measure how much participants relied on vision, during a task in which using vision would be beneficial.

In Experiment 2, two participants did show faster reaction times to visual and auditory information, indicating that they were using both senses to make their behavioural responses. Moreover, bimodal reaction times exceeded those predicted by probability summation, suggesting that both visual and auditory information were processed in combination, and thereby implying preserved or re-acquired multisensory processing abilities (see e.g. Miller, 1982; but see also Otto & Mamassian, 2012). It is not clear, however, whether this applies to all types of multisensory decisions. Whilst certain multisensory integration properties are acquired early in life (Lewkowicz, 2002; Wallace & Stein, 2001), others – including notably the combination of multisensory information to improve precision – do not develop until much later in childhood (Gori et al., 2008; Nardini et al., 2008) suggesting differences in the mechanisms involved. Additionally, research has found that human adults treated for early visual deprivation show speeded reaction times to multisensory stimuli (Putzar et al., 2012) despite impaired multisensory interactions on other more complex tasks (Putzar et al., 2012; Putzar et al., 2010; Putzar, Hotting, et al., 2007). Hence, it may be that faster multisensory perception is easily re-acquired following visual treatment, or is not actually affected by late-onset visual deprivation, whilst the precision or accuracy of multisensory judgments may instead be impaired. Importantly, it could alternatively be that certain neural pathways were not completely visually deprived, but instead stimulated by residual vision. Participants did have bare light perception remaining, and four of five participants were able to perceive the visual stimulus using their residual vision alone. Consequently, the multisensory reaction time advantage observed may reflect preserved multisensory function, following a period of some (albeit limited) visual input.

2.7 Conclusion

The results of these initial experiments highlighted that the visual information provided by the retinal implant is limited in spatial and temporal resolution. Consequently, prosthetic vision is not always sufficiently reliable to lead to either predicted or measured perceptual benefits, as was found in the experiments reported here. Specifically, there were no predicted or measured benefits (or disadvantages) of combining visual and haptic information for size discrimination. However, two participants were able to coactively process simple visual and auditory signals to improve the speed of their responses. Hence, for these two participants, at least some multisensory processes have been either preserved or re-acquired. Based on the results presented here, it is not possible to conclude whether late-onset visual deprivation or subsequent prosthetic treatment has impacted the speed, precision or accuracy of multisensory perception. More research is required, using tasks for which prosthetic vision (and not residual vision) is predicted to benefit perception.

Chapter 3

The Precision and Accuracy of Multisensory Perception following Visual Treatment: Visual and Non-Visual Navigation by Adults with a Retinal Implant

3.1 Introduction

The previous chapter studied whether late-blind adults with a retinal implant combine the visual information afforded by the implant with non-visual information to improve their precision or speed, (during a size discrimination or speeded detection task respectively). In these tasks, prosthetic vision was not sufficiently reliable to improve multisensory precision or speed. Navigation is a task for which even limited visual information about the location of a landmark could lead to improved orientation. The present chapter, therefore, assesses whether this limited visual information could be sufficiently informative to improve the precision or accuracy of navigation.

3.1.1 Combining Visual and Non-Visual Navigational Cues

During navigation, humans with healthy vision rely on both visual and non-visual sensory information to update their position and orientation within their environment. Like other mammals, humans can use both visual landmarks and idiothetic self-motion cues (including those from vestibular and proprioceptive sensory systems, as well as optic flow information) to track their own movements over time (Burgess, 2008; Gallistel, 1990; Mittelstaedt & Mittelstaedt, 1980; Morris, 1981). Thus, to navigate effectively, humans often rely on cooperation between visual and non-visual senses. Moreover, research has shown that adults with healthy vision can improve their navigational performance by combining visual and non-visual cues (Bates & Wolbers, 2014; Kalia et al., 2013; Nardini et al., 2008; Tcheang, Bulthoff, & Burgess, 2011). In many of these studies, improvements in navigation precision were well predicted by an ideal (Bayesian) observer model that averages visual and non-visual sensory estimates, weighted by their reliability (see section 1.2.2).

3.1.2 Navigation without Vision and with a Retinal Implant

Following visual loss, individuals must rely solely on non-visual sensory information and, consequently, this may impact their navigation. Although blind individuals are no longer able to use visual information in combination with non-visual information, they could, in principle, combine information from different non-visual modalities to reduce their sensory estimate uncertainty (Petrini, Remark, et al., 2014; but see section 1.4). Moreover, research has found that blind individuals show enhanced non-visual processing relative to sighted individuals on certain tasks (see section 1.3), including navigation (see section 1.3.3). Hence, blind individuals may be able to compensate partly for their loss of vision however, since vision provides the most accurate and reliable spatial information (see section 1.2.8), it plays an important role in forming a spatial representation ('cognitive map') of the environment.

Mobility aids for low vision, such as walking canes and guide dogs, can detect obstacles and changes in elevation, enabling safe route planning. Retinal implants are less efficient than existing mobility aids at detecting obstacles, however they provide users with some – albeit limited – visual information about their immediate environment that they could use to form a cognitive map. Accordingly, prosthetic vision could provide additional orientation information, allowing individuals to update their position in space relative to their locomotion. However, the ability of users to use this prosthetic visual information may be limited by (i) any cross-modal plasticity, where visual processing areas have been recruited by non-visual processes (see section 1.3), and importantly by (ii) the visual resolution and field of view of the retinal implant system (see section 2.1.1).

The Argus II retinal prosthesis system provides users with 'ultra low vision', and consequently interpreting this visual signal can be challenging (see section 2.1.1). Users must use continual head-scanning movements to explore their environment, and due to the limited resolution, do no have sufficient information to accurately perceive depth or distance. However, even weak visual information could be sufficient to improve navigational accuracy or precision. For example, navigators could fix their position relative to the direction indicated by a visual landmark, thereby improving their orientation, despite not having any information about the distance of the landmark.

3.1.3 Experiment Aims

The present study examined whether late-blind patients, implanted with the Argus II retinal prosthesis, could use this new visual signal together with non-visual information, to improve their performance on two well-known navigation tasks: a path reproduction and a triangle completion task. In both these tasks, visual information (about an indirect landmark) and non-visual (i.e. vestibular and proprioceptive) self-motion information were potentially useful for improving performance. We assessed whether patients improved their navigational precision or accuracy when given visual and non-visual self-motion cues together, compared to when using non-visual information alone.

3.2 Method

3.2.1 Ethics Statement

Patients were recruited from the Moorfields Eye Hospital NHS Foundation Trust, London, UK, and the study had received ethical approval from the East Central London committee. Ethical approval for conducting the study with control adults was received from the research ethics board of University College London. Informed written consent, according to the Tenets of the Declaration of Helsinki, was obtained from all participants prior to participation.

3.2.2 Participants

Four patients (3 male; aged 49-77 yrs; M = 66.0 yrs) implanted with the Argus II prosthesis in their right eye, at Moorfields Eye Hospital, 4-7 years prior to testing, participated. All had been diagnosed with a retinal degenerative disease (3 retinitis pigmentosa, 1 choroideremia) prior to implant, and reported having been blind for 15-52 years. All patients had some level of bare light perception, but no measurable visual acuity (> logMAR 2.9 in both eyes). Patients had received visual rehabilitation training provided by Second Sight Inc. that covered basic skills like head scanning. (Note, all four patients had participated in the tasks described in Chapter 2: IDs 001, 002, 004 & 005. As patient 003 opted not to participate, here patients 004 & 005 are represented by IDs 003 & 004 respectively). Six young adults (aged 23-29 years; mean age 25.7 years) and 5 age-matched adults (aged 54-74 years; mean age 63.0 years), all with normal or corrected vision, also participated in this study.

3.2.3 Procedure

Participants were asked to complete two tasks: path reproduction and triangle completion. Both tasks were conducted in a darkened room (6.5m x 7.75m), with black walls and black carpet, and involved using a single landmark: an illuminated white square paper shade floor lamp (0.23m x 0.23m x 1.52m, 200cd/m^2 lamp against 0.04cd/m^2 walls & carpet).

3.2.3.1 Path Reproduction Task

Participants were led to a start position and advised that the experimenter would guide them along a path which they would then be asked to reproduce as accurately as possible. The path comprised of an initial 2.5m leg, a 75° rotation, and a final 2m leg. The landmark was positioned midway along the second leg. This meant that it could potentially provide information about the correct initial heading, the distance after which to turn, and the correct turning angle, (see Fig. 9.A).

3.2.3.2 Triangle Completion Task

As for the path reproduction task, participants were led to a start position and guided by the experimenter along an outbound path, comprising of an initial 2.5m leg, a 75° rotation, and a final 2m leg. However, in this task, participants were asked to return directly to the start position as accurately as possible on reaching the end of the outbound path, thereby completing a walked triangle (see Fig. 9.B). The landmark's position was the same as for the path reproduction task. It could potentially provide information about the correct return turning angle.

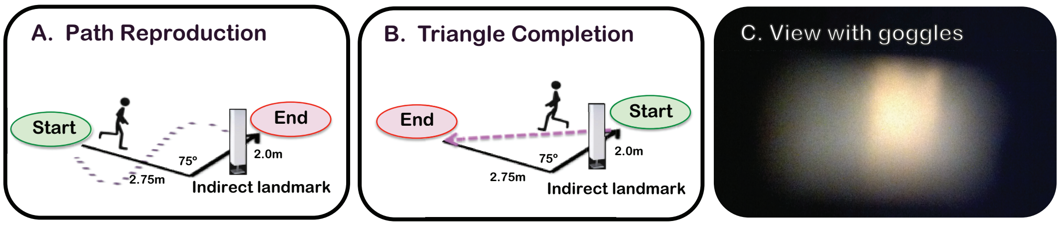


Figure 9: Schematic of path reproduction (A) and triangle completion (B) tasks. View of the landmark through the goggles worn by normally sighted participants (C).

Participants were guided along the black path by the experimenter, and then: (i) For path reproduction, guided back to the start position and asked to reproduce the path as accurately as possible. (ii) For triangle completion, asked to return to the start position as accurately as possible.

Patients were asked to complete both tasks using (i) the retinal prosthesis (i.e. system on) and (ii) no vision (i.e. blindfolded and landmark light off). Three patients able to locate the landmark using their residual vision also completed the task with the system off (see Appendix A.3.1). Control (normally sighted) participants were asked to wear goggles that restricted their field of view (11 x 18 degrees) and visual resolution (using blur, 1.6 logMAR as assessed using a logMAR chart), and similarly completed both tasks using (i) restricted vision and (ii) no vision (i.e. blindfolded and landmark light off). (They also complete the task wearing goggles that restricted their field of view, but not visual resolution, see Appendix A.3.2).

Control participants were used to establish whether similarly restricted vision (in terms of field of view and resolution) could provide useful information about the visual landmark to individuals who would usually rely on vision for navigation. This would allow us to exclude the possibility that any failure of patients to use the landmark simply reflects its limited field of view and spatial resolution for these tasks. Additionally, we were able to compare control participant and patient performance when navigating without vision, so as to assess whether patients' long-term visual deprivation may have led them to develop improved non-visual navigation skills.

To ensure patients were able to localize the visual landmark, all were initially asked to walk directly toward the landmark from ten different locations within the room (differing in distance and angle from the landmark) with the system on. All four patients were able to complete this task from all ten locations. They then completed the path reproduction task first, followed by the triangle completion task. Condition order within tasks was random, and each participant completed two blocks of five trials per condition.

3.2.4 Data Acquisition & Analysis

Participants' positions were tracked using an optical tracking system (8 Vicon MX13 cameras) through the monitoring of five helmet-mounted reflective markers. Position coordinates were saved using Vizard (Version 4.0; Santa Barbara, CA: WorldViz LLC.) and analysed using Matlab (Version R2014a, The MathWorks, Inc., Natick, Massachusetts, United States) and the Psychophysics toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). A bivariate normal distribution was fitted to each participant's end positions (i.e. where participants decided to stop), to estimate the x mean, y mean, x variance, y variance and x-y covariance for each condition. The FASTMCD algorithm (Rousseeuw & Driessen, 1999), as implemented in the Libra toolbox for Matlab (Verboven & Hubert, 2005) was used to estimate these values robustly, with the assumption of 1% aberrant (outlier) values (i.e. a value of 0.99 for the alpha parameter). The sum of the variance in x and y directions was used to obtain a single measure of total variable error, reflecting the uncertainty (or imprecision) of spatial estimates. Variable error is expected to reduce when more precise information is available, or when information from multiple sources is averaged (Cheng, Shettleworth, Huttenlocher, & Rieser, 2007; Ernst, 2006b). Additionally, a measure of constant error was calculated as the distance between the correct end location and the participant's end position. Constant error reflects a systematic bias (or inaccuracy) in spatial estimation, and is expected to reduce when a less biased information source is available (see also section 1.2.3).

Paired samples t-tests were run on control data to test for significant reductions in errors when using vision compared to when navigating without vision. Where vision was found to improve performance on navigational tasks, improvements were quantified as follows: The improvement in precision when using vision (restricted/prosthesis), was calculated as the difference in variable error when navigating with vision, compared to when navigating without vision. The improvement in accuracy when using vision (restricted/prosthesis) was calculated as the difference in constant error with vision, compared to when navigating without vision.

Given the small number of patients tested, *p* values from inferential statistical tests have not been reported. Instead, each patient's performance was compared to the 95% confidence intervals calculated from participants with normal vision. Patient data falling outside the confidence limits indicated that the difference is unlikely to have resulted from measurement error alone.

3.3 Results

3.3.1 Differences in error between young and age-matched controls

A one-way ANOVA indicated no statistically significant differences in variable errors or constant errors between young and age-matched participants in the path reproduction or triangle completion tasks, (see Table 3). Consequently, control data was pooled together for further analysis.

Table 3: Results of a one-way ANOVA comparing Variable and Constant Errors between Young and Age-Matched Participants

	Variable Error		**Constant Error**	
	Path Reproduction	Triangle Completion	Path Reproduction	Triangle Completion
Vision	$F_{[1,9]} = 0.003$, $p = 0.958$	$F_{[1,9]} = 1.070$, $p = 0.328$	$F_{[1,9]} = 2.076$, $p = 0.183$	$F_{[1,9]} = 0.020$, $p = 0.890$
No Vision	$F_{[1,9]} = 2.199$, $p = 0.172$	$F_{[1,9]} = 0.429$, $p = 0.529$	$F_{[1,9]} = 0.548$, $p = 0.478$	$F_{[1,9]} = 3.500$, $p = 0.094$

3.3.2 Variable Error

3.3.2.1 Path Reproduction

A paired samples t-test indicated that control participants had significantly higher variable errors without vision than with vision ($t_{[10]} = 3.806$, $p = 0.003$). Based on these control data, 95% of normally sighted participants would be expected to show reductions in error of 0.105m – 0.402m. Patient data fell outside of these confidence intervals, and three of four showed better performance *without* vision. In addition, all four patients' variable errors without vision were less than the lower limit of the 95% confidence intervals of normally sighted participants (ID 001-004: 0.115m, 0.031m, 0.084m, 0.117m compared to 95% CI: 0.178m – 0.483m; see Fig. 10.A).

3.3.2.2 Triangle Completion

A paired samples t-test indicated that control participants had significantly higher variable errors without vision than with vision ($t_{[10]} = 2.780$, $p = 0.020$). Based on these control data, 95% of normally sighted participants would be expected to show reductions in error of 0.038m – 0.346m. Two patients showed data that fell outside of these confidence intervals, showing better performance *without* vision. Again, all four patients' variable errors without vision were less than the lower limit of the 95% confidence intervals of normally sighted participants (ID 001-004: 0.153m, 0.047m, 0.294m, 0.225m compared to 95% CI: 0.299m – 0.670m; see Fig. 10.A).

3.3.3 Constant Error

3.3.3.1 Path Reproduction

A paired samples t-test indicated that control participants had significantly higher constant errors without vision than with vision ($t_{[10]} = 2.537$, $p = 0.030$). Based on this control data, 95% of normally sighted participants would be expected to show reductions in error of 0.027m – 0.415m. Patient data fell outside of these confidence intervals, and all patients showed better performance *without* vision. Three of four patients' constant errors without vision were less than the lower limit of the 95% confidence intervals of normally sighted participants (ID 001, 002, 004: 0.178m, 0.133m, 0.335m compared to 95% CI: 0.391m – 0.861m; see Fig. 10.B).

3.3.3.2 Triangle Completion

A paired samples t-test indicated that control participants showed no significant differences in constant errors without vision than with vision ($t_{[10]} = 1.0316$, $p = 0.3266$). Two of four patients' constant errors without vision were less than the lower limit of the 95% confidence intervals of normally sighted participants (ID 003, 004: 0.169m, 0.327m compared to 95% CI: 0.463m – 1.693m; see Fig. 10.B).

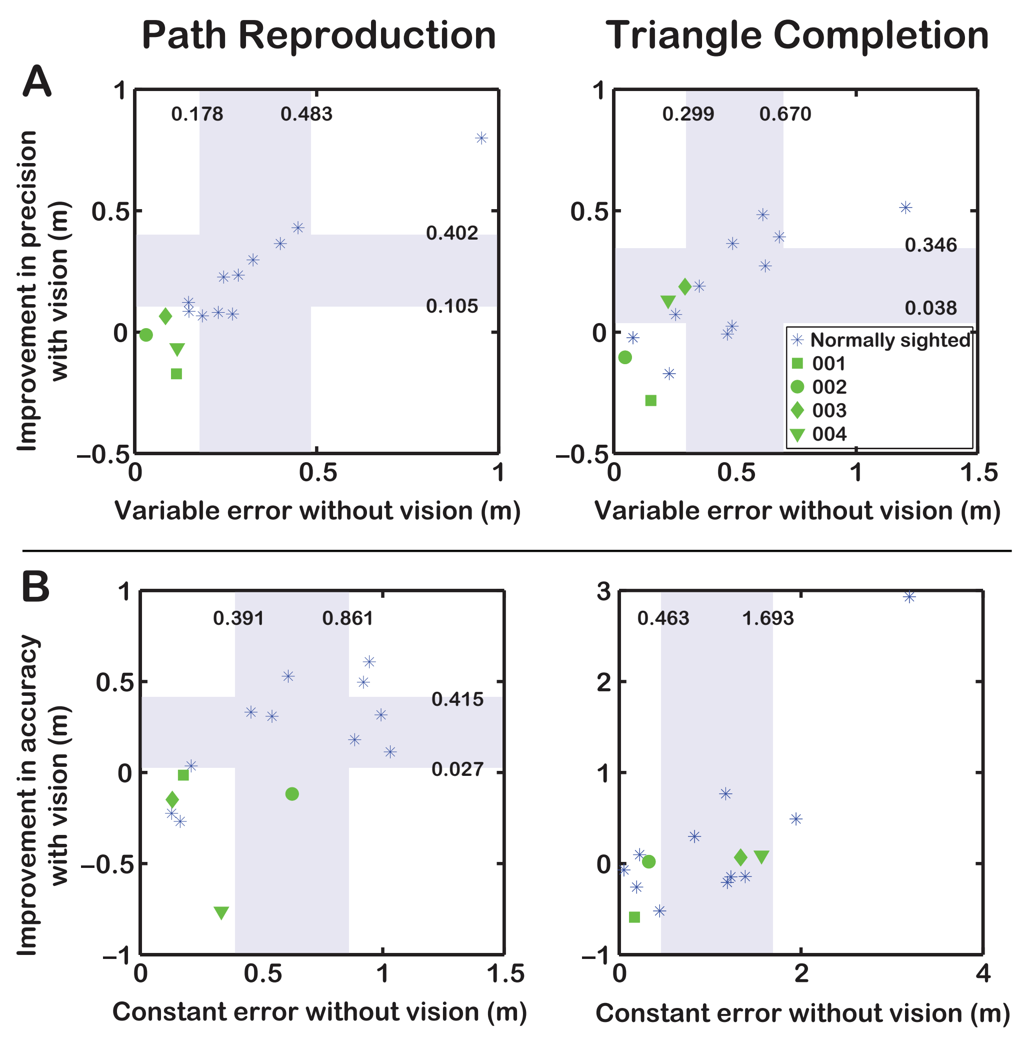


Figure 10: Graph showing improvement in variable error (A) or constant error (B) when using vision against errors when navigating without vision. Shading indicates the 95% confidence intervals computed from the control data.

Path Reproduction: Patients did not show similar improvements in precision or accuracy when navigating with the prosthesis as controls. All had lower variable errors without vision, and three had lower constant errors without vision, compared to controls.
Triangle Completion: Two of four patients showed similar improvements in precision when using vision as controls. All patients showed lower variable errors without vision compared to controls. Two patients had lower constant errors without vision compared to controls.

A Spearman's correlation was used to test whether any inability to improve performance when using vision could be due to floor effects (i.e. participants having already very precise and accurate non-visual performance). There were significant positive associations between no vision errors and the percentage improvement in errors with vision, for both tasks, indicating that participants with greater no vision errors showed greater improvements with vision (Variable errors: path reproduction $r_{s[13]} = 0.70$, $p = 0.005$, triangle completion: $r_{s[13]} = 0.58$, $p = 0.027$; Constant errors:

path reproduction $r_{s[13]} = 0.55$, $p = 0.038$, triangle completion $r_{s[13]} = 0.52$, $p = 0.049$). This shows that more precise/accurate non-visual navigators had limited potential to improve with vision.

3.3.4 Learning Effect

As the same path was repeated, to understand whether participants learnt over the course of the experiment, linear regression analyses were run to assess (i) the effect of trial number on constant error and (ii) the effect of block number on variable error. Blocks rather than trials were used for variable error because variance can only be calculated over a set of trials. Trial number did not statistically significantly predict constant error in either path reproduction or triangle completion tasks for control or patient participants (path reproduction: controls: $F_{[1,28]} = 0.140$, $p = 0.712$; patients: $F_{[1,28]} = 3.258$, $p = 0.082$; triangle completion: controls: $F_{[1,28]} = 0.194$, $p = 0.663$; patients: $F_{[1,28]} = 0.594$, $p = 0.447$). Similarly, block number did not statistically significantly predict variable error in either task for either group (path reproduction: controls: $F_{[1,4]} = 0.006$, $p = 0.941$; patients: $F_{[1,4]} = 2.040$, $p = 0.227$; triangle completion: controls: $F_{[1,4]} = 1.070$, $p = 0.360$; patients: $F_{[1,4]} = 0.927$, $p = 0.390$). No significant effect of trial number on constant error, or block number on variable error, within conditions was found (see Appendix A.3.4). These results indicate that participants did not learn over the course of this experiment. If participants had shown learning, it would have been necessary to consider the effect of task order and any differing learning rates between groups, when comparing their navigational errors.

3.4 Discussion

This study assessed whether patients implanted with the Argus II retinal prosthesis would use this new visual signal to improve navigational precision, by using spatial information provided by an indirect visual landmark (an illuminated floor lamp) as well as non-visual self-motion cues. Low resolution, restricted field of view vision was sufficiently informative to lead to improvements in navigational precision in normally sighted participants, in both a path reproduction and triangle completion task. In a multisensory cue integration framework (Cheng et al., 2007; Ernst, 2006b), this shows that the visual cue was at least as useful (reliable), for normally sighted adults, as non-visual self-motion information. Two patients implanted with the Argus II prosthesis showed similar improvements in precision (to normally sighted adults) on a triangle completion task. In contrast, three patients on the path reproduction task and two patients on the triangle completion task showed reduced precision when navigating with the Argus II prosthesis. This result is consistent with patients either (i) using the (less reliable) visual cue only (ii) switching between visual and non-visual cues or (iii) combining visual and non-visual information, but not weighting these according to their reliability.

Argus II patients' inability to use the visual landmark to improve their navigational precision is partly due to dissimilarities between the new vision afforded by the prostheses and native vision (even when restricted by goggles): Firstly, the field of view and resolution of the Argus II were more limited for some patients than that approximated by the control-worn goggles. Both field of view and resolution are dependent on the number of functioning electrodes in the implant, and only one patient had all electrodes working (this varied among the four patients, from 47-100%; see section 2.2.2). Secondly, the vision afforded by the prostheses consists of pixelated phosphenes and thus is qualitatively different to the blurred vision experienced by control participants. Thirdly, because the phosphenes elicited by the device have variable persistence independent of the stimulus (Perez Fornos et al., 2012), using this visual information sometimes requires memory and/or continual head scanning to elicit further phosphenes. Despite differences between native and prosthetic vision, all four patients were able to perceive the landmark, as shown by their ability to walk directly toward it from various locations within the room. However, obtaining accurate estimates of angles and distances using prosthetic vision was challenging for them. Patients reported using their head scanning movements together with the percepts elicited to estimate angles and distances. For example, they computed the landmark's distance by considering the amount they had to move their head to detect its edges (small head scanning movements being sufficient from far distances, larger head scanning movements needed if near to the landmark). These computations are effortful and subject to inaccuracies. Indeed, in the path reproduction task, controls showed significantly lower constant errors when navigating with vision, compared to without, but patients showed higher mean constant errors with the prosthesis instead, indicating that the prosthesis tended to bias their navigation estimates.

In addition, to limitations in the signal afforded by the prosthesis, differences between groups in non-visual processing may also partly account for the differential results between groups and across conditions. Specifically, patients showed lower variable errors when navigating without vision, compared to controls, in both path reproduction and triangle completion tasks. Results showed that the degree of improvement in error when navigating with vision was positively associated with the magnitude of error when navigating without vision. Thus, participants who were already very precise non-visual navigators were limited in their potential to improve their performance with vision. Path reproduction can be done by accurate encoding of distances and turns via self-motion (Petrini, Caradonna, Foster, Burgess, & Nardini, 2014), whilst in triangle completion participants must compute angles and distances so as to decide on a *new* (previously not walked) home-bound path (Tcheang et al., 2011). Triangle completion can be done by self-motion alone (Loomis et al., 1993), but visual landmarks can usefully reduce errors, and predominate over self-motion in healthy adults (Foo, Warren, Duchon, & Tarr, 2005). Correspondingly, using the prosthesis did not improve navigational precision in the path reproduction task, where patients could rely on their enhanced non-visual sensory information instead. Three of four patients actually showed increased imprecision when navigating with vision on this task. However, in the triangle completion task for which vision is more relevant and patients showed higher non-visual errors (compared to in the path reproduction task), patients showed differing results, with two of four showing similar improvements in precision as controls when using vision.

Patients' improved non-visual navigation, demonstrated by their reduced variable errors compared to controls when navigating without vision, is likely to be due to their long non-visual experience: All four patients in the study had been blind for at least 15 years (and up to 52 years) prior to implant, and during this period had developed strategies for independent navigation, including the use of other mobility aids e.g. canes and guide dogs. Consequently these patients had become accustomed to relying on non-visual information for navigation. This expertise may include specific behavioural strategies and/or cortical reorganisation: Previous research studying navigation by the blind has found mixed results, with some reporting superior non-visual navigation by the blind (Fortin et al., 2008; Gagnon, Kupers, Schneider, & Ptito, 2010) and others finding no differences between blind and normally sighted (Klatzky et al., 1990; Loomis, Klatzky, & Golledge, 2001). Fortin et al. (2008) found that superior navigational skills in early and late blind individuals correlated with increased hippocampal volume, which – given their inability to rely on vision to update spatial coordinates online as they navigate – may reflect increased storage of spatial information. Neuroimaging studies have shown occipital cortex recruitment by the blind for non-visual tasks, even for tasks in which patients do not show superior non-visual performance (e.g. Lewald, 2013; Renier et al., 2010; Weaver & Stevens, 2007; see section 1.3). It may be that further reorganisation is possible when visual information is partially restored, but this will likely depend on practice with the new restored visual signal. Although all four patients completed training covering how to use the device on receiving the implant, all four patients in this study reported that they tended not to use the prosthesis for navigation in everyday life, having learnt to rely on non-visual navigational strategies, (all four patients had been blind for at least 15 years). Whilst in this study, we found that neither controls nor patients learnt over the course of the experiment, it would be interesting to assess whether with more trials and feedback, patients could improve their navigational precision, as they learnt to compute distance and angle information with the new visual signal. Similarly, patients who have been visually deprived for a shorter amount of time may show less-developed non-visual sensory skills, increased reliance on the sensory signal afforded by the prosthesis in daily activities, and consequently possibly improved performance on multisensory tasks when the prosthetic visual signal is available.

In summary, the visual information afforded by the retinal prosthesis did not consistently improve performance on both tasks, partly because the visual signal is different to native vision and involves using effortful strategies to estimate angles and distances, but also partly because patients were expert non-visual navigators. It is possible that patients could be shown to benefit from the new visual signal (i) following increased use with the implant on navigational tasks and (ii) on more complex paths, with more turns and greater distances, for which non-visual information may be considerably less reliable.

3.5 Conclusion

Patients implanted with the Argus II retinal prosthesis were not able to improve their navigational uncertainty in the path reproduction task by using prosthetic vision (whereas normally sighted controls, wearing goggles that limited their vision, did). However, in the triangle completion task, two patients showed a similar reduction in navigational uncertainty when using prosthetic vision, as normally sighted controls did (wearing the googles). Furthermore, all patients showed greater precision than controls in both tasks when navigating without vision. Therefore, the differential results between patients and control participants may be partly accounted for by differences in (i) the reliability of the visual signal afforded by the prosthesis and the control-worn goggles, and (ii) the reliability of non-visual processing between groups, due to the duration of visual deprivation, practice and/or sensory reorganisation. The results indicate that the patients have compensated for not having vision for navigation by developing precise non-visual spatial estimates of their environment. However, at least on some (more complex) tasks, some patients are able to improve their navigation precision by using visual information provided by the prosthesis.

Chapter 4

The Precision of Multisensory Perception following Visual Loss: Audio-Visual Localization by Adults with Progressive Retinal Disease

4.1 Introduction

Much research has reported changes in non-visual processing following blindness (see section 1.3), which could have implications for treatments that aim to restore the visual sense. Consequently chapters 2 and 3 assessed whether retinal implant treatment affects the multisensory perception of late-blind adults. However, less is known about the effect of partial vision loss on non-visual and multisensory processing, even though most individuals with vision loss often maintain some residual vision. The present chapter assesses how individuals experiencing gradual vision loss combine residual visual and non-visual information.

4.1.1 Combining Cues as Relative Reliability Changes

Research has found that human adults with normal sight can combine sensory estimates to reduce uncertainty in the manner of an optimal decision-maker, (e.g. Alais & Burr, 2004; Ernst & Banks, 2002; Gepshtein & Banks, 2003; Helbig & Ernst, 2007; see section 1.2.1). For example, Alais and Burr (2004) asked human adults to localize briefly-presented visual Gaussian blobs and/or auditory clicks presented in central space (±20°). Results showed that human adults minimized the uncertainty of their bimodal location estimates, indicating that they were combining visual and auditory location estimates optimally. Moreover, as the reliability of the visual cue decreased (when the stimulus was made more blurred), participants increased the weight that they assigned to the auditory information, demonstrating that they were weighting cues according to their relative reliability.

Whilst researchers have shown that adults are able to re-weight signals as their relative reliability changes from one trial to the next (e.g. Alais & Burr, 2004;

Ernst & Banks, 2002), less is known about how human adults adapt to the gradual changes in sensory reliability that occur during ageing or disease. Studies of development and ageing indicate sub-optimal cue weighting in young children (Gori et al., 2008; Nardini, Begus, & Mareschal, 2013) and in older adults (Bates & Wolbers, 2014). For example, in a navigation task, Bates and Wolbers (2014) found that older adults weighted vision less (and non-visual, e.g. vestibular information, more) than predicted by the relative reliabilities of the cues, whereas, consistent with earlier research (Nardini et al., 2008), younger adults showed optimal cue combination. In development and ageing the relative reliabilities of different senses are often changing, and participants might use sub-optimal weights because they have not fully taken these changes into account. However, why might participants reweight cues as their reliability changes from trial to trial (e.g. Alais & Burr, 2004; Ernst & Banks, 2002), yet fail to account for longer-term changes?

How the nervous system accounts for uncertainty is not yet clear (Ma, Beck, Latham, & Pouget, 2006; Ohshiro, Angelaki, & DeAngelis, 2011), but an interesting possibility raised by these results is that longer-term changes in sensory reliability are dealt with differently to short-term trial-to-trial changes. For example, there could be a general reliability setting for a particular sensory cue (e.g. a visual cue to location; Alais & Burr, 2004) that is immediately modulated by the specific sensory information on a particular trial, but whose overall setting is more difficult to change. However, in development and ageing there is also the possibility that the cue combination process itself is immature or deficient (e.g. Dekker et al., 2015), and consequently age-related changes in reliability do not offer a clear way to address this question.

4.1.2 Changing Visual Reliability due to Degenerative Disease

Retinal degenerative diseases, including retinitis pigmentosa and macular dystrophy, lead to progressive visual deterioration that is often, at least initially, limited to certain parts of the visual field. Disease manifestation and progression differs substantially between patients, however retinitis pigmentosa often begins with rod cell death, causing difficulty to see in the dark and a loss of peripheral vision, gradually leading to central vision loss later in life, whereas macular dystrophy is characterized by a reduction in central vision that does not usually affect peripheral vision. Consequently, in such cases, the nervous system must account for both deteriorations in visual reliability and changes in visual reliability across the visual field, when combining visual with non-visual sensory information.

Changes in the relative reliability of visual and non-visual cues may be further complicated by compensatory changes in residual senses. For example, as shown earlier, both visual and auditory senses can provide information about the location of an event or object, and therefore by combining these cues, observers could reduce uncertainty in their localization judgments. However, research has found that (particularly early- but also late- onset) blind humans show improved auditory localization accuracy and reliability on certain tasks (e.g. Voss et al., 2004; see section 1.3.1). Whilst the effect of partial vision loss on residual senses is less clear, some findings suggest blind individuals with residual vision show changes in non-visual processing too (Cunningham et al., 2011; Lessard et al., 1998), in which case individuals would need to account for changes in the reliability of both visual and non-visual senses.

4.1.3 Changing Visual Reliability Across the Visual Field

Even in normally sighted adults, the reliability of vision changes across the visual field, visual precision decreasing with eccentricity due to a reduction in the density of cone photoreceptors (Dacey & Petersen, 1992). Research has not assessed whether normally sighted human adults weight vision optimally in peripheral (> 20 degrees) as well as central space. However, Charbonneau, Veronneau, Boudrias-Fournier, Lepore, and Collignon (2013) found that the visual capture of spatially misaligned auditory information in human adults declines with eccentricity, suggesting that adults do reduce their reliance on vision in audio-visual peripheral spatial decisions.

Interestingly, auditory localization thresholds also deteriorate with eccentricity, and so individuals with normal sight and hearing show both greater minimum auditory angles (Mills, 1958; Perrott, 1984) and minimum visual angles (Perrott, Costantino, & Cisneros, 1993) in peripheral than central locations. Consequently, whilst the relative reliability of visual and auditory cues may change across the visual field (depending on the stimuli to be localized), increased eccentricity generally has a deleterious effect on the reliability of both cues.

4.1.3 Experiment Aims

Here we assessed whether adults with progressive visual loss weight and combine visual and auditory cues to location optimally, i.e. in line with MLE predictions. Normally sighted adults and those diagnosed with a retinal degenerative disease causing either primarily central or peripheral visual loss were asked to localize stimuli using vision and/or hearing. Measured visual weights and measured bimodal estimates were compared to MLE predictions. Localization performance was assessed in both central and peripheral space. This allowed us to ask: (i) Do adults with normal vision combine audio-visual cues to location optimally in the periphery, as well as in the centre (Alais & Burr, 2004), accounting for any changes in the relative reliability of both cues? (ii) Do patients who are losing vision combine audio-visual cues to location optimally, accounting for any changes in the relative reliability of both cues, caused by central / peripheral localization differences as well as their own visual field loss?

4.2 Methods

4.2.1 Ethics Statement

Patients were recruited from Moorfields Eye Hospital NHS Foundation Trust, London, UK, and normally sighted adults were recruited through the UCL psychology online subject pool. The study received approval from the London Hampstead research ethics committee. Informed written consent, according to the Tenets of the Declaration of Helsinki, was obtained from all participants prior to participation.

4.2.2 Participants

Participants were twelve adults with central vision loss (7 male, M= 49.2 yrs, SD = 11.5 yrs), ten adults with peripheral vision loss (7 male, M = 40.9 yrs, SD = 10.4 yrs; see Table 4), and twelve age-matched normally sighted adults (6 male, M = 48.5 yrs, SD = 16.0 yrs). Participants were identified as having either primarily central or peripheral vision loss by their clinician, based on their diagnosis, clinical findings and results of investigations (retinal imaging and visual field testing), on attending an appointment at Moorfields Eye Hospital. Note that participants diagnosed with peripheral vision loss had progressive retinal conditions that affect peripheral vision in the first instance with central visual loss later in the disease process. However, at the time of this study, their peripheral vision was most severely affected, and their central visual fields (up to 18 degrees) were relatively preserved. Participants identified as having central vision loss had retinal conditions that affected the cells in their macular (central vision) only (isolated macular dystrophy). All normally sighted adults showed visual acuities of between -0.18 and 0.16 logMAR, as assessed using a logMAR chart. All participants reported having normal hearing.

Table 4: Details of all Patients with Central or Peripheral Vision Loss who participated.

ID	**Visual Disease**	**Gender**	**Age**	**Visual Acuity**	
				Right	**Left**
Patients with Central Vision Loss					
01	Stargardt disease	F	59	2/60	3/60
02	Stargardt disease	F	39	6/60	6/12
03	Stargardt disease	F	51	6/5	6/5
04	Macular dystrophy	M	51	6/18	6/9
05	Stargardt disease	M	50	1/60	1/24
06	Stargardt disease	M	62	6/5	6/18
07	Stargardt disease	F	51	6/36	6/36
08	Stargardt disease	F	59	6/5	6/5
09	Stargardt disease	M	60	3/60	6/5
10	Stargardt disease	M	43	6/60	6/36
11	Macular dystrophy	M	21	6/36	6/36
12	Stargardt disease	M	44	6/5	6/6
Patients with Peripheral Vision Loss					
01	Retinitis pigmentosa	M	48	6/9	6/12
02	Retinitis pigmentosa	F	41	6/60	6/36
03	Retinitis pigmentosa	M	28	6/5	6/5
04	Retinitis pigmentosa	M	32	6/9	6/12
05	Rod Cone Dystrophy	M	40	6/12	6/9
06	Retinitis pigmentosa	F	55	4/60	6/9
07	Retinitis pigmentosa	F	35	6/5	6/6
08	Retinitis pigmentosa	M	35	6/5	6/5
09	Retinitis pigmentosa	M	60	6/9	6/24
10	Retinitis pigmentosa	M	35	6/12	6/9

4.2.3 Apparatus & Stimuli

Stimuli were presented using 122 light-emitting diode pixels (Adafruit 12mm diffused flat digital RGB LED pixels; see Jones, Garcia, & Nardini, 2015) and 9 speakers (50mm x 90mm Visaton speaker SC 5.9), mounted on a 2.5m semi-circular ring (circle radius: 2.87m), spanning -15 to +30 degrees (see Figure 11). A further 2 light-emitting diode pixels (LEDs) and 1 speaker were mounted on the wall, 20 degrees left of the ring, acting as a fixation during peripheral stimuli presentation. Stimulus presentation was controlled using Matlab (Version R2014a, The MathWorks Inc., Natick, Massachusetts, United States) and the Psychophysics toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997), on a Windows 7 computer. An Arduino Uno microcontroller (SmartProjects, Strambino, Italy) was used to interface between the control computer and the LED pixels. The Matlab PsychPortAudio ASIO interface controlled audio presentation via a Focusrite Saffire PRO 40 sound card and audio signals were amplified using Lypin Hi-Fi 2.1 stereo amps. The sampling rate was 44.1kHz and speakers were equalized for intensity using a sound level meter.

Figure 11: The ring of LEDs and speakers.

On each presentation a flash of lights from a subset of LEDs (outlined in purple) and/or a noise from a speaker (outlined in blue) was presented. Participants maintained their head position fixed at straight ahead, using a chin rest (outlined in red), and entered responses using the keyboard (outlined in green).

All 122 LEDs were powered to show white light (2223 cd/m^2) constantly throughout the duration of the experiment. The visual stimulus was a 25 msec flash of white light from 50 adjacent LEDs, (spaced 0.5° apart, spanning 25°). The luminance of the visual stimulus was increased for peripheral (3055 cd/m^2) compared to central (2639 cd/m^2) space, to account for the approximate doubling of Differential Luminance Sensitivity (DLS) from 36° to 1° (Brenton & Phelps, 1986). The luminance of the visual stimulus was also increased for patients, where necessary, to increase the reliability of the visual stimulus. Audio stimuli were 100 msec (25 ms rise and 25 ms fall time) band-pass-filtered noise bursts (tenth octave centred on 1000Hz) presented at 50 dB SPL (± 1 dB), hidden in continuously played background pink noise presented at 20 dB SPL.

4.2.4 Procedure

Participants were asked to localize visual (light flash) and auditory (noise burst) stimuli presented separately or together, in a dimly lit, quiet room. Each trial began with the presentation of a fixation cue at 0 degrees (i.e. straight ahead), consisting of a red 400 msec light flash from two LEDs (13600 cd/m^2) and a simultaneous 400 msec 500 Hz (50 dB SPL) tone played from the corresponding speaker. Participants were asked to maintain their eye gaze in this direction throughout the whole experiment, and a chin-rest was used to fix their head position. Following the fixation cue, two sets of stimuli were presented successively: a standard (central: 1°, peripheral: 36°, right of fixation) and one of eight comparison stimuli (0-17° right of the standard). The order of the standard and comparison presentation was counterbalanced. Participants were asked to indicate whether the first or second stimulus was further to their right using a key press.

Blocks consisted of audio-only, vision-only or bimodal (audio-visual) stimuli. Where visual and auditory stimuli were presented together, stimuli were either presented in congruent locations (no-conflict), or the visual stimulus was displaced leftward (central: by 3°, peripheral: by 4°) compared to the auditory stimulus (conflict). The conflict trials were used to measure cue weighting.

The experiment was divided into two parts, one part consisting of localization in central space (central condition), the other of localization in peripheral space (peripheral condition). The order of these was counterbalanced across participants.

Note that the set-up in central and peripheral conditions was exactly the same, except that participants were rotated leftwards by 35 degrees in the peripheral condition.

Prior to commencing the test blocks in each part, participants completed two practice blocks (32 trials each), one with each of the unimodal stimuli used in the experiment. During testing, they completed 24 test blocks (6 audio-only, 6 vision-only, 12 audio-visual) of 64 trials, at each location (central and peripheral). Each block included 8 trials at each of the following comparison angles: 1°, 2°, 3°, 4°, 6°, 9°, 13°, and 18°. Equal numbers of conflict and no-conflict trials were randomly interleaved within audio-visual blocks. Thus, there were equal numbers of trials that were audio-only, visual-only, audio-visual (consistent) and audio-visual (conflict). There were 48 trials per comparison distance for each of these conditions.

4.2.5 Data Analysis

The proportion of trials in which the second stimulus was perceived as being right of the first was plotted against the size of the displacement between the two stimuli, for each cue (audio-only, vision-only, audio and vision: no conflict and conflict), and for each location (central, peripheral). Data were fitted with cumulative Gaussian functions, using psignifit toolbox version 2.5.6 for Matlab (see http://bootstrap-software.org/psignifit/), a software package which implements the maximum-likelihood method described by Wichmann and Hill (2001a). The standard deviation (σ) and the mean (μ) of each function provided, respectively, estimates of the cue's reliability (precision) and point of subjective equality (PSE). Functions were fitted to each individual participant's data.

Assuming early noise, independent and identically distributed likelihood functions, and a uniform prior, the ideal bimodal estimate is given by the average of the single cues weighted by their respective reliabilities (σ^2), known as the maximum likelihood estimate (MLE; see section 1.2.2, Eq. 6 & 7). Participants' unimodal variances were used to compute the MLE prediction, and measured bimodal variances were compared to this prediction.

The PSE describes the point at which participants were equally likely to perceive the comparison stimulus as left or right of the standard (see section 1.2.2). To assess whether participants weighted cues optimally during their localization estimates, no-conflict and conflict PSEs were used to compute the actual weighting given to vision in bimodal trials (Eq. 9), and this was compared with the predicted optimal visual weight (Eq. 9).

$$\hat{w}_V = \frac{PSE_{Conflict} - PSE_{No\ Conflict}}{Visual\ Displacement} \tag{9}$$

Thus, a difference in conflict and no conflict PSEs equal to the size of the visual displacement would indicate that participants relied entirely on visual information in their bimodal localization judgments, whereas no difference in PSEs would indicate that participants relied entirely on auditory information.

4.3 Results

Five patients with peripheral vision loss did not complete the peripheral condition, as they were unable to perceive the visual targets presented in peripheral space. Therefore, the results of all ten patients in the central localization task, and the results of just five patients in the peripheral localization task, are reported here.

4.3.1 Uncertainty

We first analysed standard deviations (σ) of fitted functions, a measure of uncertainty – higher values of σ indicate greater uncertainty (lower precision) of perceptual estimates. Figure 12 shows the mean uncertainty for the single cue (audition-only, vision-only) and bimodal conditions, and the ideal (MLE) predictions, for each group, in central and peripheral conditions. The results for all three participant groups in the central localization task (Fig. 12, top) indicated lower mean uncertainty for bimodal relative to unisensory judgments, although bimodal uncertainty was not significantly different to the best unimodal cue (see Table 5). Regression analyses of individual participants' bimodal reliabilities as compared with their individual MLE predictions show that the MLE model significantly predicted individual participants' bimodal reliabilities in all three groups (Fig. 13A & Table 6).

In peripheral space (Fig. 12, bottom), mean unisensory localization uncertainty was better matched for normally sighted participants, and their mean bimodal uncertainty was significantly lower compared to that of the best unimodal cue, again in line with MLE predictions (see Tables 5 & 6, Fig. 13B). In contrast, in the patient group with peripheral loss, mean visual uncertainty was much greater than auditory uncertainty. Although mean bimodal uncertainty was slightly reduced, this was not significantly different to the best unimodal (auditory) cue, but the results were in line with MLE predictions (see Tables 5 & 6, Fig. 13B). Patients with central vision loss also showed reduced mean bimodal uncertainty, although again this was not significantly reduced relative to the best unimodal (visual) cue. Although individual bimodal performance was well predicted by the MLE (Table 6, Fig. 13B), mean uncertainty was significantly higher than predicted ($t_{[11]} = 2.61$, $p = 0.024$; this is also clear from the position of the "central" group's regression line in Fig. 13B, which is higher than the identity line). Overall, this suggests that in the periphery, the central vision loss group behaved partly in line with MLE and obtained some cue combination benefits, but obtained a systematically lower benefit than the ideal observer prediction. An unexpected result was that in the periphery, the central vision loss group also showed very high auditory uncertainty compared with controls (Fig. 12, bottom). Auditory cue reliabilities are compared across groups in more detail below.

4.3.2 Cue weighting

Next, we analysed cue weighting. Figure 13C-D plots individual measured vision weights against individual optimal (MLE) visual weight predictions at central (left) and peripheral (right) locations. Significant linear relationships between measured and predicted vision weights were found for normally sighted adults in both central and peripheral space ($p <= 0.01$; see Table 6). A similar relationship was shown in the periphery by patients with peripheral vision loss ($p = 0.075$), but this was not statistically significant, very likely due to the small sample size ($n = 5$). In contrast, visual weights by patients with central loss showed little or no relationship with predicted weights in either central or peripheral space ($p > 0.6$; Fig. 13C-D, Table 6). This anomaly, compared with the other two groups, is likely to be related to the result that this group alone showed significantly worse-than-optimal (MLE) threshold reductions via cue combination (Fig. 13 and Table 6).

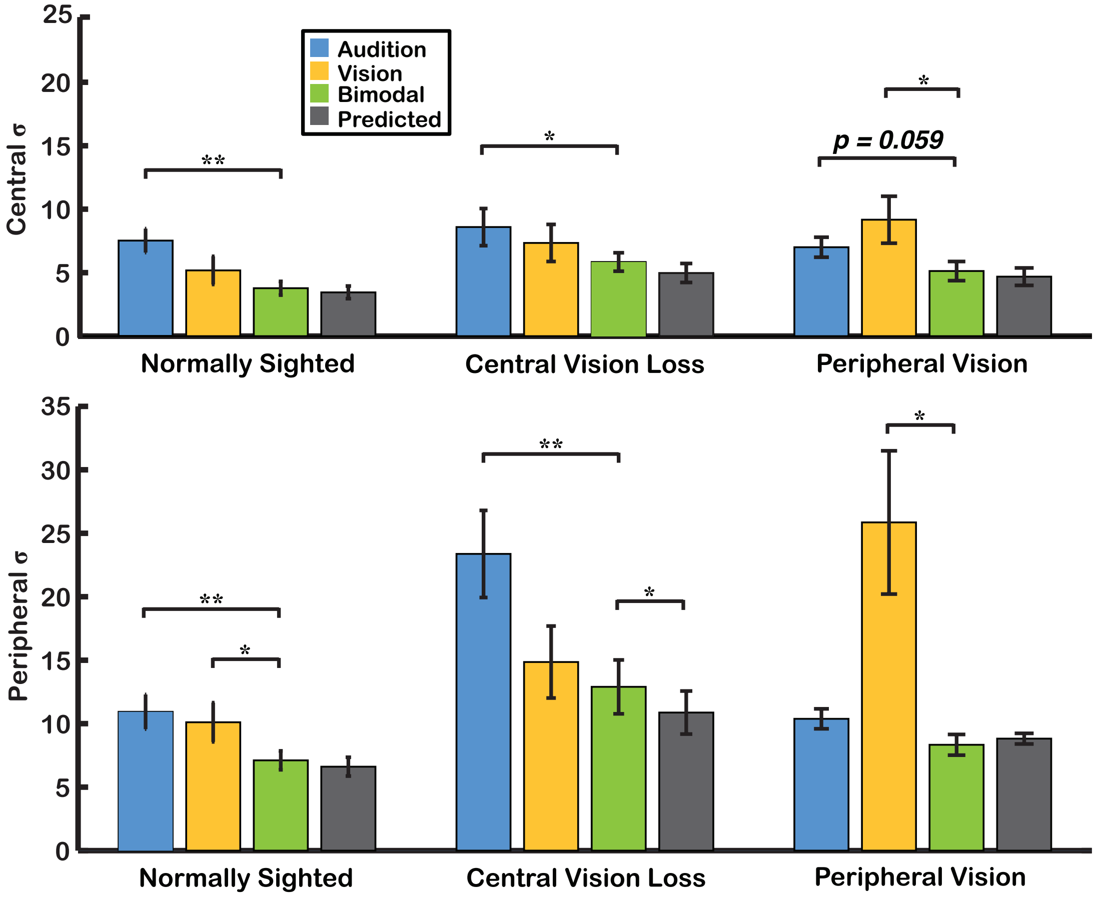


Figure 12: Visual, auditory, bimodal and predicted localization uncertainty, in central (upper panel) and peripheral (lower panel) space, for participants with normal sight, central vision loss or peripheral vision loss. The bars show the standard error of the mean, (note that this is different to the standard error of the difference, compared in paired t-tests). Bimodal uncertainty was compared with each single cue's uncertainty, and also with the ideal (MLE) prediction. (* indicates $p < .05$; ** indicates $p < .01$).

Table 5: Results of Paired Sample t-Tests comparing Unimodal (Visual-only, Auditory-only) and Bimodal Uncertainty. Shaded cells indicate the best unimodal cue. (* indicates $p < .05$; ** indicates $p < .01$).

		Normally Sighted	Central Vision Loss	Peripheral Vision Loss
Central	Vision	$t_{[11]} = 1.85$, $p = 0.091$	$t_{[11]} = 1.29$, $p = 0.225$	$t_{[9]} = 2.85$, $p = 0.019$ *
	Audition	$t_{[11]} = 3.21$, $p = 0.008$ **	$t_{[11]} = 2.49$, $p = 0.030$ *	$t_{[9]} = 2.17$, $p = 0.059$
	Prediction	$t_{[11]} = 1.82$, $p = 0.096$	$t_{[11]} = 2.01$, $p = 0.070$	$t_{[9]} = 0.94$, $p = 0.371$
Peripheral	Vision	$t_{[11]} = 2.25$, $p = 0.046$ *	$t_{[11]} = 0.80$, $p = 0.438$	$t_{[4]} = 3.44$, $p = 0.026$ *
	Audition	$t_{[11]} = 3.29$, $p = 0.007$ **	$t_{[11]} = 4.69$, $p < 0.001$ **	$t_{[4]} = 1.31$, $p = 0.261$
	Prediction	$t_{[11]} = 0.95$, $p = 0.361$	$t_{[11]} = 2.61$, $p = 0.024$ *	$t_{[4]} = 0.67$, $p = 0.538$

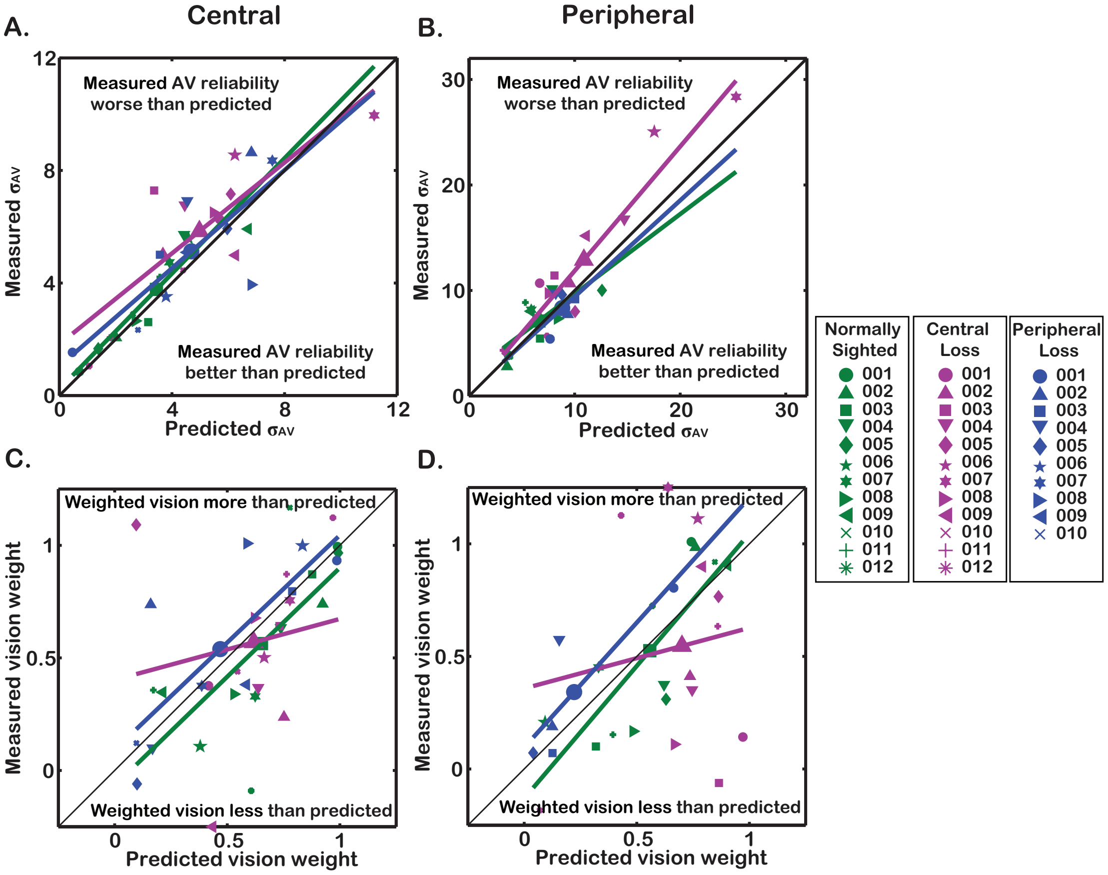


Figure 13: Top Panel - Predicted and measured audio-visual (AV) reliabilities in central (A) and peripheral (B) space. Lower Panel - Predicted and measured vision weights in central (C) and peripheral (D) space. Group means depicted by larger symbols.

Table 6: Results of Linear Regression Analyses comparing Predicted and Measured Reliabilities and Vision Weights. (* indicates $p < .05$; ** indicates $p < .01$).

		Normally Sighted	Central Vision Loss	Peripheral Vision Loss
Central	Thresholds	$F_{[2,10]} = 83.7$, $p < 0.01$ ** $R^2 = 0.89$, $\beta_1 = 1.03$	$F_{[2,10]} = 22.0$, $p < 0.01$ ** $R^2 = 0.81$, $\beta_1 = 0.81$	$F_{[2,8]} = 13.6$, $p < 0.01$ ** $R^2 = 0.63$, $\beta_1 = 0.87$
	Weights	$F_{[2,10]} = 9.98$, $p = 0.01$ * $R^2 = 0.5$, $\beta_1 = 0.97$	$F_{[2,10]} = 0.26$, $p = 0.62$ $R^2 = 0.03$, $\beta_1 = 0.27$	$F_{[2,8]} = 12.9$, $p < 0.01$ ** $R^2 = 0.62$, $\beta_1 = 0.95$
Peripheral	Thresholds	$F_{[2,10]} = 13.0$, $p < 0.01$ ** $R^2 = 0.57$, $\beta_1 = 0.76$	$F_{[2,10]} = 78.6$, $p < 0.01$ ** $R^2 = 0.89$, $\beta_1 = 1.18$	$F_{[2,3]} = 0.84$, $p = 0.43$ $R^2 = 0.22$, $\beta_1 = 0.91$
	Weights	$F_{[2,10]} = 17.3$, $p < 0.01$ ** $R^2 = 0.63$, $\beta_1 = 1.18$	$F_{[2,10]} = 0.18$, $p = 0.68$ $R^2 = 0.02$, $\beta_1 = 0.27$	$F_{[2,3]} = 7.22$, $p = 0.08$ $R^2 = 0.71$, $\beta_1 = 1.11$

We expected that performing the task in the periphery as compared with the centre would alter the relative reliabilities of vision and audition and so call for reweighting. However, as the cue reliabilities in Fig. 12 show, differences in auditory and visual thresholds in the centre were often quite similar to differences in the periphery (blue vs. yellow bars, top vs. bottom plots). The largest change in reliabilities was seen in the peripheral patient group, although there the number completing peripheral testing is small (N=5). We asked, first, whether differences in *predicted* visual weights between central and peripheral space were significant in any group. As Table 7 shows, these differences were not significant for any group. As we did not see statistically significant reweighting predicted even for ideal observers in this experiment, it is perhaps not surprising that we also did not see significant differences in *measured* central versus peripheral vision weights (Table 7).

Table 7: Results of Paired Sample t-Tests comparing Predicted and Measured Vision Weights between Central and Peripheral Space.

	Normally Sighted	Central Vision Loss	Peripheral Vision Loss
Predicted	$t_{[11]} = 1.02$, $p = 0.33$	$t_{[11]} = 0.77$, $p = 0.46$	$t_{[4]} = 1.76$, $p = 0.15$
Measured	$t_{[11]} = 0.29$, $p = 0.78$	$t_{[11]} = 0.15$, $p = 0.89$	$t_{[4]} = 0.73$, $p = 0.51$

4.3.3 Control Experiment

In a further experiment, conducted with 12 younger normally sighted participants only, we checked central and peripheral reliabilities and weighting whilst also manipulating the uncertainty of the visual cue, by asking participants to localize both a more-reliable visual stimulus (the visual stimulus described and used here) and a less-reliable visual stimulus (see Appendix A.4.1 for details). We found that predicted weights, and measured weights, did vary from centre to periphery, for the more reliable visual stimulus (used in the present experiment); see S1 for details. Predicted weights, and measured weights, also varied with the uncertainty of the visual cue - as expected in the centre from previous work (Alais & Burr, 2004) – in both the centre and the periphery. From this we can conclude that humans can and do reweight vision and audition in central and peripheral space, but that the stimuli used in the main experiment did not change the relative reliabilities of these cues in centre versus periphery enough to call for significant reweighting.

4.3.4 Comparison of Auditory Thresholds

A repeated measures ANOVA with location (central, peripheral) as the within-subjects factor, and participant group (normally sighted, central vision loss, peripheral vision loss) as the between-subjects factor showed a significant effect of location on auditory uncertainty ($F_{[1,26]} = 30.8$, $p < 0.001$), with greater uncertainty in the periphery. Furthermore, there was a significant interaction between group and location on auditory uncertainty ($F_{[2,26]} = 9.76$, $p = 0.001$). As Figure 12 shows, this is driven by the unusually high auditory uncertainty of patients with central vision loss in the periphery. Post-hoc t-tests (p values corrected for multiple comparisons) showed that patients with central vision loss showed significantly reduced higher auditory localization uncertainty relative to normally sighted controls ($t_{[22]} = 3.37$, $p = 0.008$), but due to the small sample size (n=5), not patients with peripheral vision loss ($t_{[15]} = 2.39$, $p = 0.12$). No differences in auditory localization in central space between patients with central vision loss and other participants were found (normally sighted controls: $t_{[22]} = 0.61$, $p = 0.55$; patients with peripheral vision loss: $t_{[20]} = 0.90$, $p = 0.38$). These results indicate that the cue weighting demands patients with central vision loss had to manage involved not only accounting for their loss of vision, as we expected, but also for a loss in auditory localization ability.

4.3.5 Summary

In both central and peripheral space, both controls and patients with peripheral vision loss showed bimodal uncertainty that did not significantly differ from ideal observer predictions (Fig. 12 & 13). Although bimodal uncertainty was not always significantly reduced relative to the best single cue, individual participants' bimodal uncertainties were well predicted by their individual MLEs (Fig. 13A-B), as were individual cue weights (Fig. 13C-D). Localization of the stimuli used did not require (or show) significant re-weighting by individuals across central versus peripheral space, although such re-weighting was demonstrated by controls in a separate experiment with different stimuli (S1). Patients with central vision loss showed a different pattern of results: (1) In the periphery, bimodal uncertainty was significantly worse than ideal observer (MLE) predictions (Fig. 12), and (2) in both the centre and the periphery, unlike other groups, individual measured vision weights did not match individual predictions based on cue reliability (Fig. 13C-D). Interestingly, (3) this group also showed unexpectedly high auditory uncertainty in the periphery, indicating that they needed to account not only for their vision loss but also a loss in auditory localization ability.

4.4 Discussion

This study aimed to understand whether adults diagnosed with progressive visual loss are able to account for the long-term changes to the reliability of their vision. Results showed that normally sighted adults combined visual and auditory location cues near-optimally in both central and peripheral space – they weighted cues according to their relative reliability to minimize uncertainty in their bimodal estimate. Similarly, patients with progressive visual loss that primarily affected their peripheral vision also showed reductions in bimodal uncertainty and visual weights that did not significantly differ from the predictions of the optimal MLE model. In contrast, patients with central vision loss showed significantly worse bimodal localization than MLE predictions in the periphery. In line with this, they did not weight the cues optimally in either central or peripheral space; measured vision weights showed no relation to predictions. The results suggest that human adults are able to combine multisensory cues in a way that compensates for some types of long-term progressive sensory changes, but not others.

Adults in the peripheral vision loss group, like normally sighted adults, weighted vision in line with reliability-weighted cue averaging predictions in their central localization judgments (and though limited by sample number, results are consistent with optimal combination during peripheral localization too). This suggests that, as well as rapidly re-weighting sensory cues as their relative reliability is manipulated from trial to trial (e.g. Alais & Burr, 2004; Ernst & Banks, 2002; see also Supplementary Information), the nervous system can also account near-optimally for some longer-term changes to sensory reliability following sensory loss.

The central visual loss group, however, showed a markedly different pattern of results – failure to weight by reliability and a failure to meet MLE uncertainty reduction predictions. This group did not show a systematic tendency to either over-weight or under-weight the visual cue, but instead, individual subjects' measured visual weights showed no relationship with their own optimal reliability-based predictions. The group's bimodal localization estimates did not have significantly lower uncertainty than those using their most reliable unisensory cue (vision). One account of this result would be if the group relied only on single cues (i.e. only on vision). However, since measured weights did not show a complete reliance on either vision or audition, the findings suggest that patients with central vision loss combined visual and auditory information, but using sub-optimal weights i.e. weights that did not properly account for each individual's relative cue reliabilities.

Overall, the results show two seemingly similar patient groups, one succeeding and one failing at combining cues according to the MLE rule. Why might the group with central loss, in particular, have failed to weight cues by reliability and so obtain optimal (MLE predicted) uncertainty reduction? An interesting result is that this group also showed strikingly elevated *auditory* localization uncertainty in the periphery, (see similar finding in congenitally blind adults with residual vision by Lessard et al., 1998). It was anticipated that differences across groups would reflect changes to one sense (vision), and that the task for patients, in terms of cue combination, would be to account for progressive changes in this one sense. Instead, the results suggest that the central group had to contend with changes to two senses – potentially a more challenging task for maintaining optimal cue weights than a change only to one sense. This increased difficulty of dealing with changes in both senses could have contributed to this group's difficulties with maintaining correct cue weighting.

We had not expected differences in auditory localization between these different participant groups. Consequently, one possibility is that the impaired auditory localization of participants with central vision loss is linked in some way to the deterioration of their vision. Future research is needed to address whether this is the case. However, irrespective of why participants with central vision loss showed greater auditory localization uncertainty, the question remains as to why they did not account for the relative reliability of their vision and audition when combining these cues.

It is frequently reported that participants with central vision loss learn to rely on eccentric viewing, developing a preferred retinal locus (PRL) that avoids the area of central vision loss (Crossland, Engel, & Legge, 2011). Accordingly, the central vision loss patients may have been learning a different correspondence between the auditory, head-centred, spatial map and the visual, eye-centred, representation of space, (as has been demonstrated in animals following a misalignment of visual-auditory cues, e.g. Feldman & Knudsen, 1997; Wallace & Stein, 2007). Patients in the process of learning this new mapping may have perceived a discrepancy in the spatial location of the target via vision versus audition, at least at some of the comparison positions. They may have fixated the required visual targets centrally, which would change the audio-visual mapping from a usual mapping they may have been learning to use during eccentric fixation. Alternatively, they may have fixated the targets eccentrically, but have still been in the process of learning a new audio-visual mapping for eccentric fixation. Either way, on some trials, some patients may not have combined cues in line with reliability-based MLE predictions due to a perceived spatial disparity following changes to their PRL.

Ideal observer models have been developed for tasks in which cues are systematically biased and/or spatially inconsistent (e.g. Burge et al., 2008; Kording et al., 2007), however the present study did not measure subjective biases or discrepancies across visual versus auditory cues. We propose that subjective misalignment of cues due to changes in fixation behaviour could contribute to apparent failures of cue combination in the central vision loss group, but further research is needed to test this interpretation directly. The perceptual uncertainty we measured may be a combination of uncertainty and of effects due to cues sometimes being perceived as systematically biased or not coming from the same source. This would add noise to measures of uncertainty and of cue weighting, and to measures of optimally predicted cue weighting, which depends on measured uncertainty.

In the main experiment, all participant groups showed visual and auditory discrimination thresholds that deteriorated from central to peripheral space. However, the relative reliability of both cues did not change significantly; participants did not have to adjust their relative reliance on visual versus auditory cues between central and peripheral locations and, accordingly, participants showed similar cue weighting across locations. Consequently, it is not clear whether patients with progressive visual loss account for differences in the relative reliability of visual and auditory cues across their visual field, as our control experiment (see Appendix A.4.1) showed that normally sighted adults do. Follow-up tests using different stimuli that are better suited to finding such differences are needed to establish this.

In summary, the results indicate that human adults can optimally account for (at least some) long-term progressive changes to the reliability of their vision and so combine such sensory information in a near-optimally weighted average during multisensory decisions. However, one group of patients – those with central vision loss - did not weight sensory information in line with MLE predictions based on cue reliability. Importantly, the progressive visual change appeared to influence both the reliability of vision and audition. The seemingly sub-optimal behaviour of patients with central vision loss could have been due to difficulties with accounting for long-term changes to the relative reliabilities of both cues. However, the apparent changes in the reliability of audition suggest another explanation. Changes in the spatial correspondence between audition and vision due to the development of eccentric fixation strategies may have led to subjective perceptual mismatches between vision and audition. Whether such mismatches are present – and whether they are dealt with in line with ideal observer principles (e.g. Burge et al., 2008; Kording et al., 2007) – are questions for future research. It is possible that developing eccentric fixation to deal with central vision loss may come at the (possibly temporary) cost to combining visual and auditory cues for localization.

4.5 Conclusion

Can humans account for progressive visual loss in line with MLE principles during multisensory cue combination? To our knowledge, here we describe the first data to address this question. We found one patient group that followed MLE principles, and one that did not. We suggest that the latter group may have experienced changes to cross-modal mapping not captured by the basic MLE model. If so, then it is possible in theory that the latter group's behaviour would also be near-optimal, if issues due to remapping could be taken into account – although the measures we collected do not allow us to test that here. This interpretation suggests that in most cases of visual loss, humans should be able to account for changes in the relative reliability of vision in line with MLE principles; however, further studies with other groups and modalities are clearly needed. The results highlight the need to consider possible changes in cross-modal mapping, as well as in unimodal reliability, following sensory loss.

Chapter 5

The Effect of Prior Knowledge and Visual Feedback Reliability on Auditory Localization Accuracy

5.1 Introduction

In Chapter 4, we found that gradual changes to the visual sense affecting central vision could also lead to changes in auditory localization. We proposed that this could reflect a misalignment between visual and auditory cues following central vision loss, indicating that visual and/or auditory cues to location could be biased. Accordingly, in order to combine visual and auditory cues to location optimally, patients would need to re-align these, however as vision deteriorates it can become increasingly unreliable which could impact this recalibration. In the present chapter we study a commonly reported auditory localization bias in normally sighted adults, and assess whether the reliability of visual feedback provided during training impacts if/how they overcome this auditory bias. Understanding why certain biases exist and how these can be accounted for will also be informative for understanding sensory misalignments in patients following changes to the visual sense, (including following central vision loss as in chapter 4, or following retinal implant treatment as in chapters 2 and 3).

5.1.1 Prior Knowledge to reduce Perceptual Uncertainty

Humans can rely on prior knowledge about their environment, in addition to information from sensory representations, when making perceptual decisions. There is evidence that humans combine prior knowledge with sensory information optimally, in much the same way that they combine information from different sensory modalities, to minimize perceptual uncertainty (e.g. Kording & Wolpert, 2004; see section 1.2.2). Specifically, the optimal (Bayesian) observer changes their relative reliance on sensory estimates and prior knowledge as the relative reliabilities (signal-to-noise ratios) of these changes (see section 1.2.2). Interestingly, some previously unexplained systematic errors (biases) in perception have been proposed to reflect this Bayesian reliance on prior knowledge, since biases increase as sensory uncertainty increases. For example, humans tend to underestimate the speed of moving objects. This bias in object speed has been explained by use of a “slow motion prior” reflecting that objects in natural visual scenes are most likely to be static (Stocker & Simoncelli, 2006; Welchman et al., 2008). In line with Bayesian statistics, as the contrast of the visual motion stimulus decreases (sensory uncertainty increases), observers rely increasingly on this slow motion prior, leading to greater underestimations of object speed (Stocker & Simoncelli, 2006; Thompson, 1982). This chapter reports the results of an experiment in which we firstly tested whether a frequently reported auditory localization bias also increases as sensory uncertainty increases, reflecting reliance on a Bayesian prior for sound location.

5.1.2 Biases in Auditory Localization

Psychophysical studies have frequently reported that, when asked to align a visual stimulus with a sound source, human adults show systematic errors in their judgments. Many studies report overestimations of the azimuth of a sound source (Dobreva, O'Neill, & Paige, 2011; Lewald & Ehrenstein, 1998), though underestimations have been documented too (Miyauchi, Kang, Iwaya, & Suzuki, 2014). For example, Lewald and Ehrenstein (1998) asked participants (with normal sight and hearing) to adjust the position of a laser spot toward the perceived location of a band-pass filtered noise. Results indicated a general tendency to overestimate auditory eccentricity, with greater overestimations for more eccentric sound sources (reaching up to 10.4 degrees overestimation at 22 degrees eccentricity). In addition to eccentricity (Lewald & Ehrenstein, 1998; Makous & Middlebrooks, 1990), the magnitude of auditory localization errors has been found to vary according to the elevation (Carlile, Leong, & Hyams, 1997) frequency (Lewald & Ehrenstein, 1998) or bandwidth (R. A. Butler, 1986) of the auditory stimulus, participant eye (Lewald, 1998; Lewald & Getzmann, 2006) or head (Lewald, Dorrscheidt, & Ehrenstein, 2000) position, participant age (Dobreva et al., 2011) and method of response (Lewald et al., 2000).

Differences in patterns of systematic error across response methods, or different eye and head positions, suggest that biases in somatosensory or visual modalities may contribute to the auditory localization bias. For example, participants asked to point toward a transient visual stimulus have been found to overestimate the target's position when fixating straight ahead, but to underestimate it when eye movements toward the target are allowed (Morgan, 1978; but see also Sheth & Shimojo, 2001). However, regardless of the direction, size, or cause of the visual-auditory misalignments measured, the question arises as to why any perceived systematic discrepancy exists. Visual and auditory space have different frames of reference: visual space is initially eye-centred, based on direct projections to the retina, whereas auditory space is initially head-centred, computed from binaural differences and spectral cues. Consequently, one proposed explanation is that auditory-visual spatial mismatches arise due to shortcomings in accounting for the position of the eyes relative to the head when relating auditory to visual representations of location (Cui, O'Neill, & Paige, 2010; Razavi, O'Neill, & Paige, 2007). However, systematic errors have been documented even when the eyes remain stationary at straight ahead with respect to the head (Cui et al., 2010; Lewald & Ehrenstein, 1998; Miyauchi et al., 2014) and, moreover, humans continuously receive visual and auditory feedback from their environment, which should enable them to detect and correct for misalignments in cross-modal spatial representations. For example, wearing prism glasses causes the visual field to shift, altering the visual-motor mapping, but after a brief period of exposure, human adults quickly adapt to the visual displacement (von Helmholtz, 1993). Hence, it is unclear why humans would not learn to similarly adapt or 'recalibrate' their visual-auditory mapping, so as to reduce any spatial inconsistencies. One possibility, that we explore here, is that humans are relying on prior knowledge about the statistics of sound locations within the environment.

5.1.3 Visual Feedback to Improve Accuracy

In normally sighted adults, the visual sense provides the most accurate and reliable spatial information, and therefore plays an important role in calibrating non-visual spatial representations (see sections 1.2.8 & 1.4.1). Accordingly, previous research has found that sound localization biases introduced experimentally by manipulating auditory cues can be reduced following training with visual feedback (Shinn-Cunningham, 2000). To account for a bias, the perceptual system needs to identify the cause of the error as systematic as opposed to random (reflecting uncertainty in the sensory representation or prior). In line with this, adults have been found to more quickly reduce systematic errors in their motor reaching responses when the position of visual feedback was more certain (Burge et al., 2008). Consequently, changes to the visual sense that impact visual reliability could lead to slower calibration of non-visual senses, reflecting difficulties in differentiating non-visual systematic spatial errors from visual random errors. Here we tested whether the reliability of visual feedback (manipulated experimentally) during a training phase would impact the extent of any subsequent reduction in auditory localization bias, in normally sighted adults. If participants use visual feedback to adjust systematic errors (in the auditory prior or sensory representation) we would expect participants trained with more reliable visual feedback to show greater improvements in accuracy. Alternatively, improvements in auditory localization could reflect reduced reliance on prior knowledge due to changes in the uncertainty of the prior or sensory representation, irrespective of visual feedback.

5.1.4 Experiment Aims

The present experiment aimed to understand how prior knowledge and visual feedback influence the accuracy of non-visual perception. Specifically, here we tested whether (1) auditory localization biases indicate reliance on a “prior” for sound location, and whether (2) the reliability of accurate visual feedback during training influences any reduction in auditory localization bias. The experiment was conducted with normally sighted adults, and so the reliability of the visual sense was manipulated experimentally here.

5.2 Methods

5.2.1 Participants

24 adults aged 18 to 24 years (6 male, M = 20.5yrs, SD = 1.9yrs) with normal vision and normal hearing participated. Participants were recruited through the UCL psychology online subject pool. The study received approval from the London Hampstead research ethics committee. Informed written consent, according to the Tenets of the Declaration of Helsinki, was obtained from all participants prior to participation.

5.2.2 Apparatus & Stimuli

As per Chapter 4 (see 4.6.1.2): stimuli were presented using nine speakers (50mm x 90mm Visaton speakers SC 5.9) and up to 122 light-emitting diode pixels (Adafruit 12mm diffused flat digital RGB LED pixels), mounted on a 2.5m semi-circular ring (circle radius: 2.87m), spanning -15 to +30 degrees (see 4.6.1.2, Fig. 11). A further speaker was mounted on the wall, 20 degrees left of the ring. Stimulus presentation was controlled using Matlab (Version R2014a, The MathWorks Inc., Natick, Massachusetts, United States) and the Psychophysics toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997), on a Windows 7 computer. The Matlab PsychPortAudio ASIO interface controlled audio presentation via a Focusrite Saffire PRO 40 (Focusrite plc., UK) sound card and audio signals were amplified using Lvpin Hi-Fi 2.1 (Lvpin Technology Co. Ltd, China) stereo amps. The sampling rate was 44.1kHz and speakers were equalized for intensity using a sound level meter. An Arduino Uno microcontroller (SmartProjects, Strambino, Italy) was used to interface between the control computer and the LED pixels (see Jones et al., 2015).

Responses were made by rotating a dial (Griffin Technology PowerMate NA USB Controller) to control, (via Matlab), which LED pixel was illuminated. Eye position was monitored using a Tobii X120 (Tobii AB) eye tracker. An acoustically transparent curtain was arranged in front of the speakers.

Auditory stimuli were 100 msec (including 25ms rise/fall time) band-pass-filtered noise bursts (tenth octave centred on 1000Hz) presented at 50 dB SPL, (from speakers positioned at 0, 2, 6, 9, or 13 degrees, right of straight-ahead). These were hidden in background pink noise presented at 10 dB SPL (“more reliable" stimulus, which we denote *A1*) or 30 dB SPL (“less reliable”, stimulus, *A2*), (-35, 0, 1, 2, 3, 4, 6, 9, 13 and 18 degrees; mean position = 2°**;** mean position excluding speaker at -35° = 6°). Visual and visual feedback stimuli were 25 msec flashes of white light (4620 cd/m^2) from either 45 (“more reliable”, *V1, VF1*) or 5 (“less reliable”, *V2, VF2*) LEDs, randomly sampled (on each trial, without replacement) based on a truncated normal distribution ranging from ± 25 LEDs (with mean = 0, corresponding to the centre of the 50 LEDs, and standard deviation of 12 LEDs).

5.2.3 Procedure

The experiment was divided into four tasks, split over two days (see Table 8). During each task, participants were asked to maintain their eye gaze at a fixation cue, consisting of two LEDs emitting red light (1300 cd/m^2) presented at 0 degrees. A chin-rest was used to fix participants’ head position, and an eye-tracker was used to monitor eye position (a quick eye-tracking calibration task was completed before commencing the experiment). Participants initialized experimental tasks by pressing a keyboard key, and, provided that the eye-tracker detected that participants were fixating in the correct position, the trial would commence.

Table 8: Summary of the Experiment Phases, Tasks and (Within- & Between- Subject) Variables.

Day	Phase	Localization Task	Within-Subject	Between-Subject
1	Before training	Auditory	Auditory reliability (A1, A2)	-
		Visual	Visual reliability (V1, V2)	-
2	Training	Auditory + visual feedback	None – all A2	Visual feedback (VF1, VF2)
	After training	Auditory	Auditory reliability (A1, A2)	(None, but analysed by VF)

5.2.3.1 Auditory and visual localization before training

On the first day, participants completed an auditory localization task and a visual localization task. On each trial of the auditory localization task, a brief noise burst was played at one of five speaker positions (0, 2, 6, 9, or 13 degrees relative to straight-ahead), whilst on each trial of the visual localization task, a brief flash of light was presented from a sample of LEDs, the mean of which was centred at one of these same five positions (0, 2, 6, 9, 13, degrees). On each trial, following the stimulus presentation, two randomly selected adjacent LEDs (width spanning 1 degree) lit up. Participants were asked to move these two lights, by rotating the dial, toward the perceived source of the noise burst or light flash, maintaining their eyes fixed on the central fixation cue. Once participants were satisfied that the LEDs were aligned with the sound or flash location, they pressed a keyboard key to store their response and this immediately commenced the next trial. Each task consisted of four blocks of fifty trials (10 trials per location tested): two blocks with less reliable stimuli and two blocks with more reliable stimuli. Block order and stimulus location presentation was random. Prior to commencing the test blocks, participants completed a short practice block comprised of five trials (1 per location tested) with the more reliable stimulus.

5.2.3.2 Training

On the second day, participants completed an auditory localization task with visual feedback. As in the initial auditory localization task, on each trial, a brief noise burst was played at one of the five speaker positions, following which participants adjusted the position of two white LEDs until they were aligned with the perceived sound source. However, on pressing a keyboard key to store their response, participants were presented a brief flash of light from a sample of LEDs whose mean position was centred at the veridical sound source location. Again, the task consisted of four blocks of fifty trials each. The auditory stimulus was the same across all blocks, corresponding to the “less reliable” (A2) stimulus in the previous task. Visual feedback stimuli were also the same across all blocks for each participant, but varied across participants (see Table 8): twelve participants were presented with a more reliable visual cue as feedback (VF1), and twelve were presented with a less reliable visual cue as feedback (VF2). The properties of these were the same as of V1 and V2 respectively during the visual localization task.

5.2.3.3 Auditory localization after training

After training, the initial auditory localization task was repeated exactly as before, with two auditory reliability levels (A1, A2) and without any feedback.

5.2.4 Data Analysis

Trials during which the mean and/or standard deviation of eye coordinate position exceeded 2 degrees from the fixation target were excluded (<2% trials for any participant). Due to the size of the experimental set-up, participants' localisation estimates were restricted to a maximum of 30 degrees. To account for this, truncated normal distributions (truncation point at 31 degrees) were fitted to each participant's localisation estimates, at each location (0, 2, 6, 9, 13 degrees), for each stimulus (A1, A2, V1, V2). The mean and standard deviation of these distributions provided measures of each participant's localisation bias (participant mean estimate – correct location) and variability, respectively. Biases at each position were then averaged across locations for a measure of mean bias. The standard deviation of deviances at each position were also averaged for a measure of mean variability (uncertainty).

5.3 Results

5.3.1 Bias and variability before training

First we considered whether biases in auditory localization might be explained by the existence of a Bayesian prior. According to Bayesian statistics, the influence of prior knowledge should increase when sensory information is less reliable. To test this, we measured participants' auditory and visual localization biases and variability, before any training, and assessed whether biases in auditory localization increased as the signal-to-noise ratio (reliability) of the auditory stimulus decreased.

Participants localized "more reliable" (A1, V1) and "less reliable" (A2, V2) auditory and visual stimuli. To manipulate the reliability of visual and auditory localization, the background noise level was increased and the number of visible LEDs was reduced. Figure 14 shows mean auditory and visual localization variability for each of the locations at which stimuli were presented (A) and across all locations (B). As intended, increasing the background noise level significantly increased the variability (i.e. reduced the reliability) of auditory localization ($\sigma_{A2} > \sigma_{A1}$, $t_{[23]} = 5.88$, $p < 0.001$). Similarly, reducing the number of visible LEDs significantly increased the variability of visual localization ($\sigma_{V2} > \sigma_{V1}$, $t_{[23]} = 7.94$, $p < 0.001$). Hence, our cue manipulations succeeded in varying sensory uncertainty.

Figure 14C shows mean auditory and visual localisation biases at each of the locations at which stimuli were presented. Participants tended to overestimate the eccentricity of auditory stimuli at each of the locations tested, especially the “less reliable” stimulus A2. Estimates of visual stimuli were more accurate. Biases are not explained by participants’ responding towards the mean location of the target stimulus set (depicted in Fig. 14C by the grey dashed line) or the mean location of the background speaker set (depicted in Fig. 14C by the black dashed line).

If participants combined sensory evidence with prior knowledge in the manner of an ideal (Bayesian) observer, they would weight prior knowledge more as sensory evidence becomes less reliable (more uncertain). Hence, if biases in localization reflect (at least partly) the use of a prior, biases would be expected to increase as the reliability of sensory information decreases. As Figure 14D shows, auditory localization biases were significantly greater for the less reliable auditory cue than more reliable auditory cue ($\mu_{A2} > \mu_{A1}$, $t_{[23]} = 7.56$, $p < 0.001$). Mean biases (across locations) were also significantly greater for the less reliable visual cue than more reliable visual cue ($\mu_{V2} > \mu_{V1}$, $t_{[23]} = 2.83$, $p = 0.009$), however, participants did not consistently overestimate the azimuth of less/more reliable visual stimuli across all locations tested (see Fig. 14.C).

5.3.2 Effects of training with feedback

Second, to test whether sound localization biases can be reduced with experience, we compared sound localization during and after completion of auditory localization with visual feedback, and compared biases and variability to before-training results. There are at least four ways in which training could improve localization accuracy (reduce localization bias): (i) Sensory reliability could improve (a narrower sensory likelihood function), leading to less reliance on a prior; (ii) Assuming a biased sensory representation, sensory accuracy could improve (likelihood function shifted) – this account requires no influence of a prior; (iii) The accuracy of the prior could improve (prior distribution shifted); (iv) Prior uncertainty could increase (a wider prior distribution).

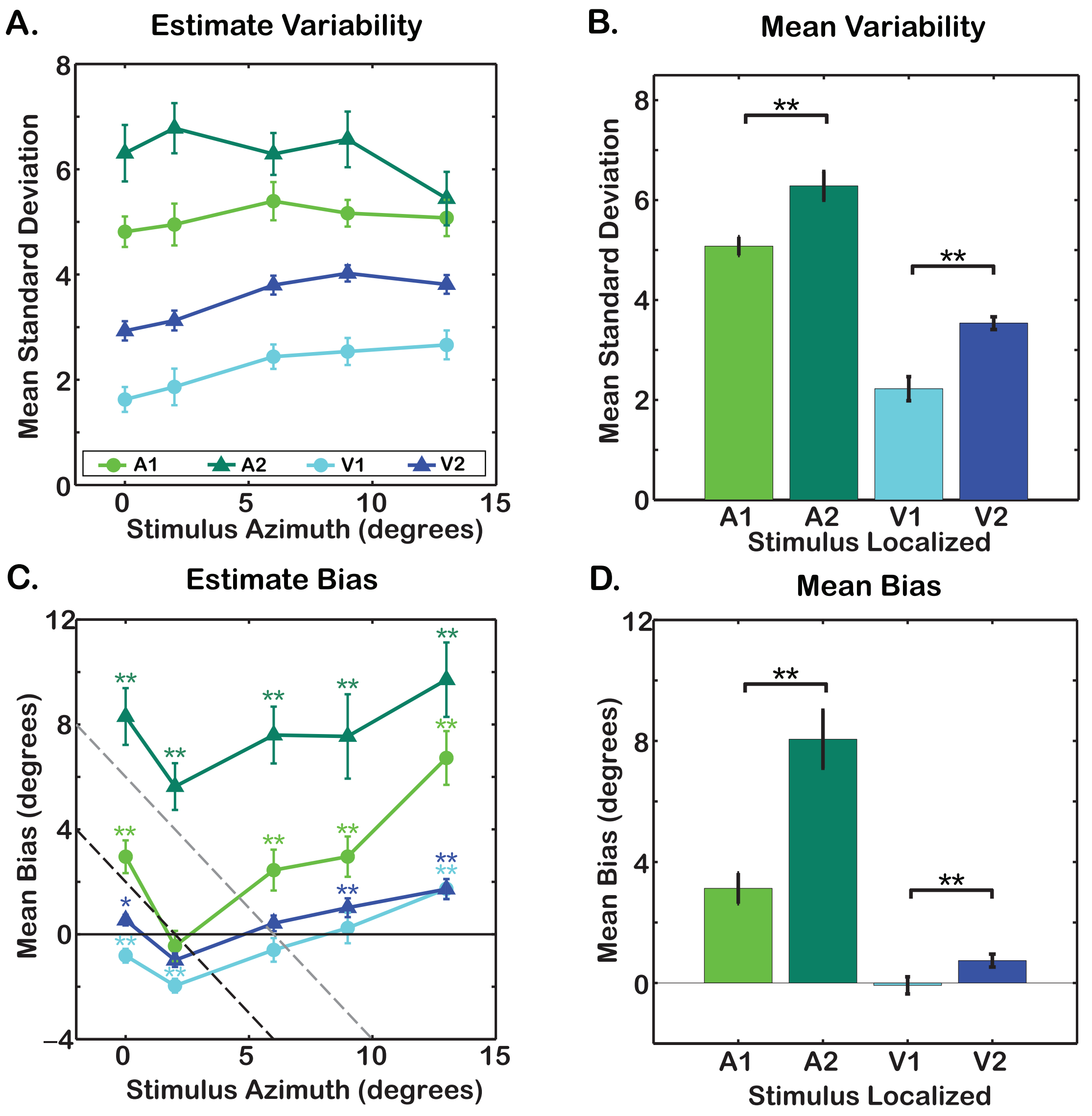


Figure 14: Bias and Variability in Localization of Auditory and Visual Stimuli Before-Training.

A. Mean variability at each location for each stimulus. B. Mean variability at each location for each stimulus. * 95% CI excludes 0; ** 99% CI excludes 0. C. Mean bias at each location for each stimulus. ** means differ significantly on paired t-test with $p<0.01$. D. Mean bias for each stimulus across all locations tested. Grey dotted line in C indicates the line predicted by responding according to the mean of the target stimulus set. Black dotted line in C indicates the line predicted by responding according to the mean of the speakers presenting background noise. Error bars represent the standard error of the mean.

Before-, during-, and after- training, all participants were asked to localize the less reliable auditory stimulus (A2); see Table 8. During training, half the participants received more reliable visual feedback (VF1), and half received less reliable feedback (VF2). The reliability of the visual feedback was manipulated to assess whether this affected the degree of any improvement in sound localization accuracy. Errors in localization could be systematic (reflecting a bias), random (due to sensory uncertainty), or both. An ideal system would adapt more quickly to erroneous feedback that is more likely to reflect systematic than measurement errors(Burge et al., 2008). If visual feedback were used to adjust inaccuracies in either the sensory representation or prior, we might expect a more reliable visual feedback cue to lead to a greater reduction in bias. Repeated-measures ANOVAs with phase (before-, during-, after- training) as the within-subjects factor and visual feedback reliability (VF1, VF2) as the between-subjects factor were run to assess whether phase or visual feedback reliability had an impact on A2-localization bias or variability.

5.3.2.1 Bias with the trained auditory stimulus

As bias data did not meet the assumption of sphericity ($\chi^2_{[2]} = 12.72$, p = 0.002), Greenhouse-Geisser corrections are reported. Phase had a significant effect on bias ($F_{[1.38,30.26]} = 12.44$, $p < 0.001$), while neither feedback reliability ($F_{[1,22]} = 0.02$, $p = 0.896$) nor the interaction ($F_{[1.38,30.26]} = 0.97$, $p = 0.360$) did. Figure 15A shows the mean bias for each of the experimental phases. Post hoc comparisons with a Bonferonni adjustment showed that bias was significantly reduced in training ($t_{[23]} = 4.02$, *Bonferonni-corrected* $p = 0.002$) and after-training ($t_{[23]} = 3.26$, *Bonferonni-corrected* $p = 0.011$) phases, compared to the before-training phase. There was no significant difference in bias between the after-training and during-training phase ($t_{[23]} = 1.89$, *Bonferonni-corrected* $p = 0.229$).

5.3.2.2 Variability with the trained auditory stimulus

Phase also had a significant effect on variability ($F_{[2,44]} = 17.06$, $p < 0.001$). Feedback reliability did not have a significant effect ($F_{[1,22]} = 0.26$, $p = 0.614$), although there was a significant phase by feedback reliability interaction ($F_{[2,44]} = 5.65$, $p = 0.007$): Participants trained with less reliable visual feedback, showed significantly reduced variability during-training than before-training, whereas participants trained with more reliable visual feedback did not (see Appendix A.5.1).

Figure 15B shows the mean variability for each of the experimental phases. Post hoc comparisons showed that variability was significantly reduced in training ($t_{[23]}$ = 4.46, *Bonferonni-corrected* $p < 0.001$) and after-training ($t_{[23]}$ = 4.17, *Bonferonni-corrected* $p = 0.001$) phases, compared to the pre-training phase. There was no significant change in variability between during-training and after-training phases ($t_{[23]}$ = 0.74, *Bonferonni-corrected* $p > 0.999$).

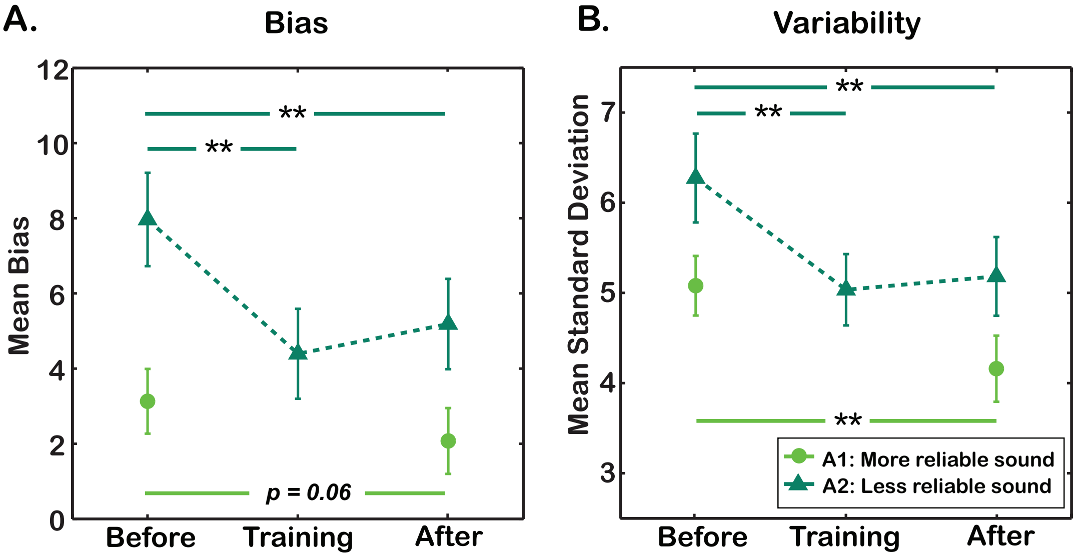


Figure 15: Mean Bias (A) and Mean Variability (B) for the Different Experimental Phases and Auditory Stimuli. Bars represent standard error of the mean. Paired-sample t-test results: * p<0.05; ** p<0.01.

5.3.2.3 Bias and variability with the untrained auditory stimulus

Before- and after- training, all participants were also asked to localize the "more reliable" auditory stimulus (A1). Bias and variability of these judgments were analysed using repeated-measures ANOVAs with phase (before- and after-training) as the within-subjects factor and visual feedback reliability (VF1, VF2) as the between-subject factor. Variability ($F_{[1,22]}$ = 16.91, $p < 0.001$) but not bias ($F_{[1,22]}$ = 3.89, $p = 0.061$) was significantly reduced in the after-training compared to before-training phase. There was no effect of visual feedback reliability on either bias ($F_{[1,22]}$ = 0.04, $p = 0.836$) or variability ($F_{[1,22]}$ = 0.09, $p = 0.762$), and no significant interactions (bias, $F_{[1,22]}$ = 0.09, $p = 0.769$; variability, $F_{[1,22]}$ = 0.54, $p = 0.471$). A1 bias and variability before- and after-training are plotted in Figure 15.

5.4 Discussion

This study aimed to assess (1) whether systematic biases in auditory localization can be explained by reliance on a prior and (2) whether the reliability of visual feedback during training impacts any reduction in auditory localization bias. We found that when the variability (uncertainty) of auditory localization estimates increased, biases also increased. This is consistent with observers relying on prior knowledge to reduce the uncertainty in their auditory location estimates. We also saw significant effects of training: Both biases and variability declined in during- and after- training phases compared to before- training, irrespective of visual feedback. The improvements that were obtained after training are consistent with observers either relying less on prior knowledge or modifying their prior in response to visual feedback.

Participants showed a tendency to overestimate the eccentricity of a sound source (in before-, during- and after- training phases) that could not be explained by their simply responding toward the mean of the stimulus set (as has been reported for other judgments (e.g. Jazayeri & Shadlen, 2010; Sciutti, Burr, Saracco, Sandini, & Gori, 2014). This systematic overestimation of sound location eccentricity is consistent with many previously reported findings (Cui et al., 2010; Lewald & Ehrenstein, 1998; Odegaard, Wozny, & Shams, 2015; Razavi et al., 2007), although underestimations have also been reported (Miyauchi et al., 2014). Previous studies have explained biases in terms of factors such as distorted spatial working memory, leading to biases for transient as opposed to on-going auditory targets (Dobreva, O'Neill, & Paige, 2012; Lewald & Ehrenstein, 2001), or errors in accounting for the position of the eyes relative to the head, causing mismatches between eye-centred visual space and head-centred auditory space (Razavi et al., 2007). Differences in the persistence and magnitude of biases have also been documented across age groups (Cui et al., 2010) and different auditory frequencies or bandwidths (R. A. Butler, 1986; Lewald & Ehrenstein, 1998). We found that participants made systematic transient sound localization errors, even though we controlled and monitored eye position (as did Lewald & Ehrenstein, 1998; Miyauchi et al., 2014). However, moreover, we found that biases increased with increased auditory localization uncertainty, suggesting that participants were combining sensory information with prior knowledge to estimate the sound source.

Why might a sound localization prior that is generally biased toward the periphery (but biased centrally in certain tasks, Miyauchi et al., 2014; or for certain individuals, Odegaard et al., 2015) exist? The statistics of humans' auditory environments do not intuitively suggest a non-uniform prior for azimuth. Specifically, whilst it makes sense that most visual patterns might move relatively slowly (Stocker & Simoncelli, 2006), it is not clear why the statistics of sound azimuths around the head should be non-uniform, since people move around freely and sounds can come from anywhere. However, natural auditory statistics could, in principle, be collected to address this question empirically. For example, Parise, Knorre, and Ernst (2014) recorded natural sounds in the environment and found evidence for the existence of a natural frequency-elevation mapping, in which high frequency sounds tended to originate from elevated sources. They also found consistent frequency-dependent biases in horizontal sound localization (see also e.g. Lewald & Ehrenstein, 1998), which may account for differences in the direction and magnitude of biases across different experiments. Alternatively, it is possible that mechanisms other than Bayesian use of priors can explain increasing biases under uncertainty. If no evidence is found that natural sound azimuths are non-uniform, this may motivate a search of such mechanisms, and a possible re-assessment of the argument that increasing biases under uncertainty indicate a role for prior knowledge (e.g. Senna et al., 2015; Stocker & Simoncelli, 2006). There are also two further interesting theoretical possibilities: first, that for some reason humans employ a prior that does not accurately reflect natural statistics; and second, that it is loss functions (reflecting differing perceived costs of mislocalizing an auditory stimulus at different azimuths) rather than priors that are non-uniform.

Interestingly, we also found that reducing the reliability of visual localization led to an overall overestimation in the eccentricity of the visual target, however, relative to auditory biases, visual biases were small (< 2°). Previous studies have found evidence of the use of a common prior for specific visual and auditory judgments. For example, humans systematically underestimate the speed of moving objects and moving sounds, in line with their reliance on a common prior that objects in the environment are more likely to be static (Senna et al., 2015; Stocker & Simoncelli, 2006; Welchman et al., 2008). Here, however, unlike for auditory localization, overestimations of visual target eccentricity were not observed at all azimuths tested, and for the more reliable visual stimulus, there was actually a tendency to underestimate the eccentricity of targets presented at central positions. Previous studies have also reported biases in visual target localization, and these have been shown to vary according to the position of eye gaze (Morgan, 1978), the presence of other visual targets (Kerzel, 2002; Musseler, van der Heijden, Mahmud, Deubel, & Ertsey, 1999), spatial attention (Fortenbaugh & Robertson, 2011) and retention interval (Sheth & Shimojo, 2001). When, as in our study, participants had to maintain fixation and head position at straight ahead, both consistent (2-4°) overestimations across eccentricities (Bock, 1986; Morgan, 1978) and underestimations that increase (up to 4°) with eccentricity have been found (Sheth & Shimojo, 2001). There is a cortical magnification of central visual space, since the fovea is represented by a higher number of neurons than the periphery. It is possible that visual biases reflect combined effects of (i) sensory likelihoods biased towards central space because of incomplete accounting for this magnification and (ii) a prior biased towards the periphery. The influences of these may vary by task, stimulus, and location. We found a visual bias whose direction and magnitude differed at different eccentricities and for the different visual cues. The two visual stimuli chosen to manipulate visual localization reliability also differed in the extent of the visual field that they covered: the “more reliable” stimulus had a mean width of 24 degrees, whereas the “less reliable” stimulus had a mean width of 13 degrees. Since factors including the spatial distance between elements on a display (Musseler et al., 1999) have been found to influence visual biases, it may be that differences in the width and spacing of visual stimuli may account for the differences in the visual bias we observed.

To perceive objects and events in the environment accurately, and to adapt to bodily and sensory changes during development and ageing, humans must keep sensory estimates calibrated. Errors or perceived mismatches (e.g. in prism adaptation, von Helmholtz, 1993) provide the feedback for this kind of learning. One question raised in the Introduction was why auditory localization biases should be so prevalent, given lifelong opportunities to correct errors. A BDT explanation, which we investigated, proposes that such “errors” might reflect use of a prior. In the second part of our study we also conducted an initial test to assess whether the reliability of visual feedback during training influences any reduction in auditory localization biases. Both auditory localization biases and variability, for the trained (less reliable) and untrained (more reliable) auditory stimulus, were reduced in after-training compared to before-training phases, irrespective of the reliability of visual feedback. These results could be explained in several ways within a BDT framework, via changes to the sensory representation (likelihood) and/or changes to the prior distribution. Specifically, reduced variability in the sensory representation, or increased variability in the prior distribution, would both predict reduced reliance on the prior, and consequently a reduction in both variance and bias, as observed. Moreover, the sensory likelihood and/or prior distribution may have shifted toward the centre following training, leading to an additional reduction in bias (overestimation of eccentricity). Since behavioural paradigms are unable to disambiguate changes in the sensory likelihood versus the prior, future research using alternative approaches (e.g. using neural recordings, Gold, Law, Connolly, & Bennur, 2008) is needed to identify exactly how improvements in accuracy are mediated.

During training, participants could have attributed localization errors signalled by visual feedback to systematic error (bias) – due to a mismatch in auditory and visual spatial representations – or random – due to visual or auditory sensory uncertainty. Consequently, we had expected that observers provided with more reliable visual feedback would be more likely to attribute errors to a mismatch in the auditory and visual mapping, as opposed to random sensory noise (e.g. Burge et al., 2008), and would therefore show faster learning, reflected by greater improvements in accuracy. However, visual feedback reliability did not influence learning, which suggests, although does not show conclusively, that improvements in bias were driven by improvements in auditory localization reliability, rather than modifications to the prior or to the accuracy of the auditory representation. Specifically, the nervous system should only respond to an error by modifying a prior or representation if the error is systematic and cannot be explained simply by sensory uncertainty. Had the participant group trained with more reliable visual feedback (and therefore clearer evidence that errors were systematic) shown greater improvements in auditory localization accuracy, this could not be explained solely by changes to auditory localization reliability.

A recent study Odegaard et al. (2015) used a different approach to examine whether biased sensory likelihoods and/or priors account for visual and auditory localization biases. They asked participants to localize auditory (noise burst) and visual (Gaussian disk) stimuli presented separately or together at various azimuthal locations (-13, -6.5, 0, 6.5, or 13 degrees), and then determined which of six quantitative models, that varied in terms of sensory likelihood and/or prior parameters, best fitted their data. Results indicated that participants tended to underestimate the eccentricity of visual-only stimuli and, consistent with our findings, overestimate the eccentricity of auditory-stimuli. Auditory and visual biases under bimodal stimulus presentation were dependent on whether the observer inferred common or independent causes for the simultaneously presented auditory and visual stimuli. Unimodal (auditory-only, visual-only) and bimodal (auditory and visual) data were best accounted for by a model that incorporated eight parameters, including a centrally biased visual-only likelihood, a peripherally biased auditory-only likelihood and a general prior for centre. This model was superior to others that assumed non-biased sensory likelihoods and unimodal priors. In the present study, however, we found that auditory-only biases increased as auditory-only localization uncertainty increased, which is consistent with increased reliance on an auditory (or general) peripheral spatial prior. It appears, therefore, that a model incorporating both biased sensory likelihoods and unimodal (visual-central, auditory-peripheral) priors may be necessary to fully account for auditory and visual localization biases. Further research is needed to address why such priors and sensory representation biases exist and, as discussed above, verify whether mechanisms other than Bayesian use of priors can more simply explain perceptual biases that increase with uncertainty.

5.5 Conclusion

Previous research has found that humans show biases in auditory localization of varying magnitude and direction (e.g. Dobreva et al., 2011; Lewald & Ehrenstein, 1998; Odegaard et al., 2015). Here, we find participants showed a tendency to overestimate the eccentricity of a sound source, but importantly that overestimations increased as sensory localization uncertainty increased. This is consistent with the Bayesian Decision Theory principle that as sensory uncertainty increased participants increasingly relied on prior information. Furthermore, we found that auditory localization biases decreased across experimental phases, providing evidence that accuracy can be improved with experience (as well as precision). Further research is needed to test the extent to which sound sources in natural human auditory environments might have non-uniform distributions of azimuths, and the extent to which alternative mechanisms could explain increases of bias under sensory uncertainty.

6. Discussion

6.1 Summary of findings

The aim of this thesis was to address how changes to the visual sense impact the speed, precision and accuracy of human multisensory perception. Much research has shown that normally sighted adults are able to combine visual and non-visual information to reduce the reaction times, uncertainty and/or bias of their response (e.g. Ernst & Banks, 2002; Miller, 1982). The ability to combine information from multiple sensory modalities develops through experience with multisensory stimuli (Gori, 2015; Wallace & Stein, 1997, 2007), during which, vision plays an important role in calibrating the spatial maps of non-visual modalities (Gori, 2015; Wallace, Perrault, et al., 2004). Here, we studied the impact of changes to the visual sense, later in life, on multisensory processing.

First, chapters 2 and 3 assessed whether restoration of vision late in adulthood, following a period of late-blindness, can lead to perceptual improvements in multisensory tasks. Visual deprivation can lead to changes in non-visual processing (e.g. Collignon, Voss, et al., 2009; Merabet & Pascual-Leone, 2010), which may have implications for the success of visual treatments. In particular, visual processing areas can become recruited by non-visual senses, following early-onset (e.g. Gougoux et al., 2005; Lewald & Getzmann, 2013), but also late-onset (e.g. Burton et al., 2002; Cunningham et al., 2015) visual loss. Such compensatory plasticity may be negative for treatments that aim to restore vision by stimulating the visual pathway directly. Typically visually responsive (unisensory or multisensory) neurons that have adapted to responding preferentially to non-visual inputs may need to re-acquire the capacity to respond to visual stimulation. Additionally, to combine multisensory cues optimally, the brain must learn to weight multimodal information according to the relative reliability of the sensory cues. This is a complex process that does not develop until late in childhood (Dekker et al., 2015; Gori et al., 2008; Nardini et al., 2008), and is dependent upon sufficient multisensory experience (Yu et al., 2010). We conducted a series of experiments with blind patients implanted with a retinal prosthesis in which, due to large differences between the reliabilities of restored vision and non-visual cues, patients were not always able to benefit from the new visual information. However, where visual and non-visual reliabilities were more closely matched, benefits to using vision with non-visual information were documented on a reaction time and navigation task. Given limitations in sample size, it was not possible to further analyse the influence of (i) the onset and/or duration of blindness or (ii) the onset and/or duration of visual treatment on the relative visual and non-visual reliabilities measured. This is discussed further below (see 6.2.1).

Next, chapter 4 studied whether adults with degenerative visual loss adjust their reliance on visual and non-visual cues in line with the changing cue relative reliability. Most cases of visual loss occur gradually, with individuals often maintaining some degree of residual vision. Whilst healthy human adults have been shown to re-weight sensory estimates as their reliability changes from one trial to the next (Alais & Burr, 2004; Ernst & Banks, 2002), it was not clear whether longer-term changes in sensory reliability are dealt with similarly. Using an auditory-visual localization task, we tested whether patients accounted for changes to their vision across the visual field. One patient group did account for reliability changes optimally, whereas another patient group – for whom visual loss was restricted to their central field – did not. We propose that the inability to adjust reliance on vision in line with ideal observer predictions may be due to changes in the spatial mapping of auditory and visual cues. Patients with central loss learn to rely increasingly on peripheral vision to attend to central space (developing a PRL, see e.g. Crossland et al., 2011). Consequently, they may have been learning a different correspondence between the auditory, head-centred, spatial map and the visual, eye-centred, representation of space. Thus, at least on some trials, patients may have perceived a disparity between the visual and auditory cues. Further research is needed to assess whether patients with central loss do indeed perceive a spatial misalignment in visual and auditory cues, and whether they are able to adapt to this perceived misalignment.

Finally, in chapter 5, we assessed whether a commonly reported bias in auditory perception could be explained by reliance on prior knowledge. Specifically, psychophysical studies have frequently found that human adults show systematic errors when aligning a visual stimulus with the perceived location of a sound source (e.g. Dobreva et al., 2011; Lewald & Ehrenstein, 1998). Since human adults continuously receive multisensory feedback from the environment, it was unclear why they would not use this feedback to correct any misalignments in sensory representations. We found participants tended to overestimate the eccentricity of the sound source and, consistent with reliance on prior knowledge to reduce sensory uncertainty, overestimations increased as auditory localization uncertainty increased. Next, we assessed whether visual feedback could be used to correct this bias and, importantly, whether the reliability of the visual feedback impacted any improvements in localization accuracy. Results indicated that both localization biases and variability declined following training with visual feedback, irrespective of visual feedback reliability. Therefore, in this task, changes to the visual sense did not impact the ability to improve the accuracy of auditory perception. In terms of Bayesian Decision Theory, improvements following training could reflect either reduced reliance on or modifications to prior knowledge. However, further research is needed to address whether there is a basis in the physical environment for a sound localization prior (e.g. Parise et al., 2014), or whether other mechanisms could account for increased perceptual biases under sensory uncertainty.

6.2 Why do changes to the visual sense impact multisensory perception?

The present studies have found that changes to the visual sense can impact the speed, precision and accuracy of multisensory perception. Human adults were not always able to combine visual and non-visual cues to improve their perception (chapters 2, 3 and 4). However, this could be explained by large differences in the reliability between visual and non-visual cues (chapters 2 and 3), or changes to the spatial-mapping of visual and non-visual cues (chapter 4), rather than an inability to combine sensory information. Moreover, at least in some cases, sensory experience may be sufficient to drive improvements in perceptual accuracy and precision, irrespective of changes to the visual sense (chapter 5).

6.2.1 Improved non-visual processing

The performance of patients implanted with a retinal prosthesis on a size discrimination, speeded reaction, and navigation task was often much worse using vision-only than non-visual cues. Interestingly, results from the navigation task (chapter 3) indicated that patients showed superior non-visual navigation than normally sighted controls. Previous findings have similarly reported improved processing by residual senses following visual loss (e.g. Fortin et al., 2008; Goldreich & Kanics, 2003; Lessard et al., 1998). Improved non-visual performance by blind, compared to normally sighted, adults may result from increased practice and/or cortical reorganisation (Lewald, 2013; Weaver & Stevens, 2007). Importantly, however, the visual information afforded by the prosthesis was also limited, in both spatial and temporal resolution. As future technological developments improve the resolution afforded, (and thereby the discrepancy in visual and non-visual reliabilities decreases), it will be interesting to consider whether patients continue to rely on non-visual information or do use vision to improve their perception on multisensory tasks. In accordance with physiological findings in animals (Yu et al., 2010), it is likely that visual and non-visual experience will be fundamental in driving sensory combination behaviour, particularly for patients who have been visually deprived for a large amount of time.

Patients who have been without vision for a longer amount of time may show more-developed non-visual sensory skills, and therefore rely less on any visual information afforded by the prosthesis. Recently, Cunningham et al. (2015) found initial evidence suggesting that the recruitment of visual cortex by non-visual senses could be reversed by visual experience with a retinal prosthesis (see Section 3.3). As more patients receive retinal implant treatment, it will be interesting to assess the effect of the onset age and/or duration of visual loss on treatment outcomes. Indeed, researchers have found that the hearing of cochlear implant users is influenced greatly by when (pre or post-lingually) the user acquired the hearing impairment, due to the resulting cross-modal plasticity shown by pre-lingually deaf users (Buckley & Tobey, 2011; Doucet et al., 2006).

Similarly, studies have explored (and found) differences in cross-modal plasticity between early and late-onset blind individuals, where ‘early blind’ is frequently used to refer to congenitally blind individuals, or individuals blinded before the age of three years. This distinction arose due to research documenting certain early (pre- 3 years) critical periods for the development of aspects of visual function, including, for example, complex form, object and face recognition. Multisensory capabilities do not appear to develop until much later in childhood (Dekker et al., 2015; Gori et al., 2008; Nardini et al., 2008), and though it is likely that these abilities could develop with experience in adulthood (e.g. Yu et al., 2010), it may be informative to extend the age range considered as ‘early onset’ when evaluating multisensory perception. Moreover, less research has considered the influence of the duration of sensory deprivation on cross-modal plasticity. Whilst cross-modal changes have been documented more frequently in early- than late- blind adults, reflecting increased plasticity in early developmental years, some studies have shown that the brain can retain considerable plasticity in later life (e.g. Ostrovsky et al., 2006). Hence, it follows that the duration of visual deprivation may also impact the outcome of visually restorative treatment. These are interesting questions to consider in future evaluations of restorative treatment outcomes.

6.2.2 Impaired non-visual processing

Results of the auditory-visual localization task (chapter 4) indicate that progressive central vision loss can have a deleterious effect on both the localization reliability of vision and audition. Thus, whilst many studies have reported improved auditory localization capabilities by blind individuals (e.g. Doucet et al., 2005; Voss et al., 2004), here we found that gradual visual loss impacted auditory localization negatively. Similarly, Lessard et al. (1998) found that congenitally blind individuals with residual peripheral vision localized sounds less precisely than sighted or totally blind subjects. The authors suggested that difficulties may arise in developing an auditory map of space that is only partly supported by vision, and that there may be reduced cross-modal plasticity where visual cortex is still stimulated by residual vision. Since the participants in our study were not congenitally blind, but were experiencing gradual visual deterioration, we instead propose that changes in the spatial correspondence between audition and vision, due to the development of eccentric fixation, may have led to a mismatch in the perception of auditory and visual stimuli location. Specifically, as patients learn to use peripheral vision to fixate on central space, there may be a misalignment between eye-centred visual space and head-centred auditory space. Studies have shown that when vision is altered – for example by prism glasses that shift the visual field – this leads to a corresponding physiological shift in the representation of auditory space by collicular neurons (Feldman & Knudsen, 1997; King et al., 1988; Knudsen & Brainard, 1991; Wallace & Stein, 2007). Hence, any misalignment in visual and auditory space may be temporary – whilst the representation of auditory space adjusts to the new visual-auditory correspondence. Future research is needed to confirm whether such spatial mismatches are present, and, if so, whether there is a recalibration of auditory space with sufficient visual-auditory experience.

Additionally, the use of eccentric fixation to perceive central space may introduce a bias in visual perception. For example, Morgan (1978) found that when participants fixated eccentrically, their subjective straight-ahead shifted in the direction of the eye turn. An ideal observer would use information from accurate sensory signals to calibrate a biased sensory cue. For normally sighted adults, typically vision is the most accurate sense for spatial judgments and therefore used to calibrate the other sensory modalities (Gori, 2015). However, for patients with central loss, it may be that vision is no longer the most accurate sense for (at least some) spatial judgments, despite still being the most reliable. If visual localization were shown to be biased in patients with central vision loss, it would be interesting to assess how they account for both the change in visual-auditory spatial mapping and bias in visual perception.

6.2.3 Unchanged non-visual processing

Interestingly, psychophysical studies have found that normally sighted adults can show inconsistencies in auditory and visual spatial representations. Specifically, normally sighted adults asked to align a visual target with a sound source show systematic errors in their judgments (e.g. Dobreva et al., 2011; Lewald & Ehrenstein, 1998). We found that normally sighted adults tended to overestimate the eccentricity of sound sources, and that such overestimations could be explained by reliance on prior knowledge, since biases increased with sensory uncertainty (chapter 5). Whilst biases in auditory localization were reduced following training with visual feedback, the reliability of the visual feedback did not influence learning. Hence, changes to the reliability of the visual sense did not impact auditory localization accuracy. This was unexpected as, in line with previous research, we had expected that observers given more reliable visual feedback would have clearer evidence of a systematic auditory localization error (as opposed to a random error reflecting visual or auditory localization uncertainty), and therefore show faster adaptation reflected in greater improvements in accuracy (e.g. Burge et al., 2008). Instead, irrespective of the reliability of the visual feedback, participants showed improved auditory localization precision following training. This suggests, although not conclusively, that improved auditory localization accuracy was driven by improved reliability of the auditory representation (and therefore reduced reliance on the prior), rather than an adjustment to the prior or sensory representation following feedback. Why did the reliability of auditory localization improve? Reliability could have improved simply due to auditory localization experience that is not reliant on visual feedback (see e.g. review on perceptual learning by Kellman & Garrigan, 2009). To assess whether visual feedback has an impact, it would be necessary to assess whether accuracy improves following training without visual feedback.

In our study, despite manipulating the reliability of the visual feedback, auditory localization reliability was always worse relative to visual reliability. What would happen if visual feedback reliability were worse than auditory localization reliability? In line with previous research (summarised above), we would expect any inconsistencies between responses and visual feedback to be attributed to visual feedback uncertainty and, consequently, no improvement in bias, unless auditory localization reliability improved. Hence, we have a situation in which even severe changes to the reliability of the visual sense may not impact the ability to improve the spatial accuracy of a non-visual cue. Note, however, that this is because auditory localization inaccuracies reflect the use of a prior, and therefore by improving the accuracy of the auditory representation, there is less reliance on the prior. If instead, auditory localization judgments were based only on a biased auditory sensory representation, auditory localization experience alone would be insufficient to improve accuracy. Interestingly, the finding that humans rely on a prior when making auditory localization judgments may provide a further explanation for studies reporting superior localization accuracy by blind compared to sighted adults (e.g. Lessard et al., 1998). Improved auditory localization accuracy by blind adults could reflect improved auditory localization reliability, through practice, and reduced reliance on a prior. Evidently, however, this would certainly not account for all findings, and could not explain the compensatory plasticity that has been documented following visual deprivation (e.g. Collignon, Davare, et al., 2009; Gougoux et al., 2005).

6.2.4. Summary

The results of the studies conducted with patients undergoing visual treatment or visual loss imply that changes to the visual sense can also impact non-visual senses, and together these have implications for multisensory perception. Firstly, improvements or impairments in visual/non-visual perception can lead to large changes in their relative reliabilities, and large reliability discrepancies (as predicted by Bayesian ideal observer models) reduce the benefit of combining multisensory information. Secondly, some types of visual impairment may lead to changes in the spatial correspondence between visual and non-visual cues, and possibly also biases in visual perception, which may (at least temporarily) disrupt the ability to combine visual and non-visual cues. In an attempt to investigate how changes to the visual sense impact the ability to improve the accuracy of non-visual cues, we assessed whether normally sighted adults could (i) reduce an auditory localization bias using visual feedback and, if so, (ii) whether the reliability of visual feedback influenced the magnitude of bias reduction. The reliability of visual feedback did not affect learning and, instead, it is likely that reductions in auditory localization bias were driven by reduced auditory localization uncertainty causing reduced reliance on a prior. Hence, where non-visual inaccuracies reflect reliance on a prior, changes to non-visual uncertainty can have a greater impact on non-visual accuracy than changes to the visual sense.

6.3 Implications for Treatments and Rehabilitation following Sensory Loss

The results of the studies presented have implications for the treatment and rehabilitation of adults experiencing visual loss. First, the results of the retinal prosthesis studies (chapters 2 & 3) showed that non-visual sensory reliability was often much better than the visual reliability afforded by the prosthesis. Whilst, clearly, the reliability of the prosthetic vision will improve with future technological developments, practice can also lead to improvements in sensory reliability. The patients who participated in these studies commented that they had become accustomed to relying on non-visual senses, and consequently did not use the prosthesis in daily tasks. Recent findings by Cunningham et al. (2015) suggest that compensatory plasticity following visual loss can eventually be reversed by visual experience with a retinal prosthesis but, importantly, that following a period of not using the retinal prosthesis visual cortex is again recruited for non-visual processing. Hence, encouraging patients to use the retinal prosthesis during daily tasks could lead to improvements in visual reliability, and the re-recruitment of visual cortex for processing preferentially visual – as opposed to non-visual – inputs. However, retinal prostheses have been more frequently assessed according to the improvement in performance that they afford on tasks that involve (at least primarily) relying on vision alone, for example reading letters or detecting squares on a screen (Ahuja et al., 2011; da Cruz et al., 2013). It is suggested here that rehabilitation programmes should also assess the influence of retinal prostheses on multisensory tasks and include multisensory training. By using the prosthetic vision in combination with non-visual senses, patients may begin to find correspondences between prosthetic visual and non-visual information. In turn, this may allow them to better interpret the visual information but also to correct any biases in visual perception, and thereby improve performance on visual-only tasks.

Second, the results of the visual-auditory localization task (chapter 4) suggested that certain visual diseases – affecting primarily central vision – could lead to increased non-visual uncertainty. As discussed above, more research is needed to understand when and why visual changes impact non-visual perception. However, low vision rehabilitation services that teach patients to shift their visual field from straight ahead to the ‘best’ peripheral retinal area (thereby developing a ‘trained retinal locus’), may want to consider that the accuracy and reliability of non-visual senses could also be affected. We suggest that as patients learn to rely on peripheral vision to fixate centrally, any misalignments or biases in visual and non-visual spatial information will gradually be corrected with multisensory experience. Therefore, training programmes in eccentric viewing could include a multisensory component, whereby, for example, patients are presented with temporally- and spatially -congruent visual-auditory stimuli and encouraged to use their peripheral vision to locate the stimulus.

Finally, the final study (chapter 5) highlighted that changes to the visual sense need not always impact non-visual senses. Here we found evidence that auditory localization may involve reliance on a prior (however research is needed to justify the existence of this prior, see chapter 5, discussion). In this case, improving the precision of auditory localization, via auditory localization practice, could be sufficient to improve auditory localization accuracy (although we cannot conclude that visual feedback was not necessary from our findings). Hence, whilst we have emphasised the advantages of training with multisensory cues for improving perception (examples above), we do acknowledge that unisensory training could be sufficient to lead to perceptual improvements on certain tasks.

6.4 Other Questions for Future Research

The work presented here has found that changes to the visual sense, later in life, can impact non-visual processing, thereby also affecting multisensory perception. Additionally, however, findings have opened up questions for future research. In chapters 2 and particularly 3 we found that the visual information afforded by the retinal prosthesis is very different to native vision. Specifically, to interpret the ultra low visual information, patients have to learn to use certain strategies including head scanning movements to explore the environment, which can bias their perception (see section 3.3.3). Hence, learning to use prosthetic vision, to some extent, involves learning to use a *new* sensory cue.

It is not clear whether normally sighted adults would immediately combine a new sensory cue with existing non-visual information. Evidence showing that the ability to combine multisensory cues to reduce perceptual uncertainty does not develop until late in childhood (e.g. Dekker et al., 2015; Gori et al., 2008; Nardini et al., 2008; see section 1.2.8), suggests that to combine a new sensory cue with existing sensory information requires extensive multisensory experience with the new and existing cues, to learn which cues belong together and how they relate (see also Ernst, 2007). This has implications, not only for retinal implants, but also for sensory substitution devices, which convert information from a substituted modality (typically vision) into another modality that can then be interpreted (see section 1.6.1). As with retinal implants, sensory substitution devices could introduce perceptual biases, and their effectiveness will depend partly on the extent of compensatory plasticity (see section 6.2.1; though, whereas sensory substitution devices promote cross-modal reorganisation, retinal implants aim to restore the previously deprived (or substituted) modality). Moreover, combining the new sensory information with existing sensory information (including for example residual vision) may require extensive training, and

– given the importance of vision for calibrating non-visual space – where the visual sense is substituted, the combination of non-visual modalities may actually be impaired (see section 1.4). Future research assessing how sensory substitution devices are used in combination with other sensory information is needed to address these questions.

In line with the need for training to combine new sensory signals with existing sensory cues, some recent initial evidence suggests that most normally sighted adults trained to learn a new sense do not immediately combine it with vision. Nardini, Negen, Roome and Thaler (VSS Abstract, 2016), trained ten normally sighted adults to use echolocation to estimate location, and then assessed their localization using echolocation, vision or both. Seven participants performed above chance using the new sensory echolocation cue, but of these only two combined the new sensory signal with vision. Whilst this implies that sensory experience plays a role in acquiring combination abilities, interestingly it also suggests that for at least some individuals, relatively little sensory experience is needed. Hence, even in normally sighted adults, there appears to be variability between individuals in their ability to learn to combine a new sensory cue with existing sensory information. Indeed, similarly, variability has been shown amongst normally sighted adults in how they combine sensory information for typical (non-novel) cues too (e.g. Fetsch et al., 2009; Oruc et al., 2003).

In Chapter 3, we found that two of four patients with a retinal implant were able to use the new visual signal in combination with non-visual information to reduce navigational uncertainty on a triangle completion task. We suggest that this inter-participant variability could partly reflect differences in their experience with the implant and/or compensatory plasticity (see section 6.2.1). Since variability in sensory combination also exists between normally sighted adults, it may be that other factors may partly account for inter-participant variability too. To form an accurate perception of the environment, humans should only combine sensory signals that provide information about the same event or stimulus. Research has shown that humans solve this causal inference problem optimally, as predicted by Bayesian statistics, by combining sensory information weighted by their posterior probability of common or independent sources (Bayesian Causal Inference; Kording et al., 2007). One possibility is that the variability in whether or how participants combine sensory cues, reflects differences in whether they perceive the sensory cues to pertain to a common source, which in turn, will likely depend on the discrepancy in the sensory information provided by each cue. For example, participants who perceive a greater discrepancy in location between echolocation and visual information (possibly reflecting a bias in the echolocation cue) may require more experience to establish the correspondence between these cues in order to infer a common cause between these, and thereby combine them to reduce perceptual uncertainty. To explore this further, future research could consider whether the magnitude of any biases associated with using a new sensory signal impacts the amount of experience needed before the new signal is used in combination with other modalities.

As reviewed (section 1.2.2) much research has shown that humans combine sensory information, perceived to arise from a common source, by weighting cues according to their relative reliability, and neurophysiological research has shown how neurons implement this reliability-weighted integration (see review by Fetsch, DeAngelis, & Angelaki, 2013). More recently Rohe and Noppeney (2015) used psychophysics, Bayesian modelling, functional magnetic resonance imaging and multivariate decoding to assess how the brain performs Bayesian Causal Inference in an audio-visual spatial localization task. They found evidence for a hierarchy of multisensory processes in the human brain, in which location is: firstly (in auditory and visual areas) estimated based predominantly on the signals having independent sources, secondly (in the posterior intraparietal sulcus) based predominantly on the signals having a common source, and finally (in anterior intraparietal sulcus) estimated according to the probability of having an independent or common source. This implies that the brain encodes location estimates under the assumptions of both common and independent sources, before considering the probability of that the sensory information pertains to a common source.

In Chapter 4, we found that patients with central vision loss did not combine visual and auditory location information according to their relative reliabilities. We proposed that this could reflect changes in the spatial correspondence between audition and vision (see section 6.2.2), which may have led to a mismatch in the perceived location of auditory and visual stimuli, at least at some of the azimuths tested. According to the findings by Rohe and Noppeney (2015), this would impact the very last part of the multisensory processing hierarchy only. However, since initial evidence indicates that areas of visual cortex become recruited by non-visual modalities, even following partial (as opposed to complete) visual loss (Cunningham et al., 2015), it seems plausible that changes following sensory loss may also impact earlier parts of this processing hierarchy. Do patients experiencing gradual vision loss make perceptual estimates in the brain based both on signals having a common and independent source? Do top-down prior expectations about the relative accuracies, reliabilities or congruencies between the cues impact the processing hierarchy? Research combining psychophysics with neurophysiological methods to study multisensory processes following sensory loss in the brain, could help answer these questions and further understanding of neural changes and their implications for perception.

6.5 Conclusion

The present studies aimed to understand the impact of changes to the visual sense on multisensory perception. The performance of human adults experiencing changes to their vision (due to visual treatment, disease or experimental-manipulations) on tasks that involve using vision and a non-visual sense was measured. Results indicated that changes to the visual sense can lead to changes in non-visual processing on certain tasks, which can impact the ability to combine sensory information to improve the speed, precision or accuracy of perception. It is proposed that individual experiencing visual treatment or visual loss could benefit from rehabilitation programmes that include multisensory training tasks.

Three main directions have been suggested for future studies: First, as more patients receive retinal implant or sensory substitution treatments, it will be informative for predicting treatment outcomes to measure the impact that the onset and/or duration of blindness and treatment have on patients' restored/substituted, compared to residual, sensory reliability, as well as considering other factors that could explain variability between individuals in combining sensory information. Second, the influence of central vision loss on auditory processing needs further investigation. One area to explore is how any misalignment between central and auditory space impacts the underlying neural activity in the brain associated with auditory and multisensory processes. Another is to what extent training can correct for this misalignment and thereby enable reliability-weighted combination of visual and auditory location cues. Finally, an analysis of the location of different sounds in different environments is needed to understand whether there is a basis for an auditory localization prior, and if not, other mechanisms that could account for sensory biases increasing with uncertainty need to be considered.

Appendix

A.1

The appendix includes further details and results for the experiments described in Chapters 2-5.

A.2 Chapter 2

A.2.1 Different Ball Sets

Data from four of the five participants (001, 002, 003 and 004) with ball set 1 (41-57 mm) were suggestive of better ball size discrimination using vision when comparing larger stimuli (> 49mm) to the standard ball (=49mm), than smaller stimuli (< 49mm). It was possible that larger balls were easier to perceive using the Argus II retinal prosthesis system. Therefore, to check whether absolute ball size affected visual-only discrimination ability, participant 005 completed the task with ball set 2 (49-65 mm). The absolute size of the balls in set 2 was bigger, but the ball sizes to be discriminated were kept constant (ranging from 2 to 8 mm). Participant 005 did not show improved visual-only discrimination of ball sizes (see Table 9), and therefore we decided not to pursue whether absolute ball size was affecting size discrimination performance any further.

Table 9: Mean (SD) Proportion of Correct Visual-Only Discrimination Trials according to Comparison Size.

	001	**002**	**003**	**004**	**Mean**	**005**
					Set 1	**Set 2**
Comparison > Standard	0.40	0.63	0.48	0.45	0.49	0.63
	(0.08)	(0.15)	(0.17)	(0.13)	(0.14)	(0.15)
Comparison > Standard	0.73	0.55	0.70	0.73	0.68	0.40
	(0.05)	(0.10)	(0.22)	(0.13)	(0.14)	(0.14)

A.2.2 Practice Experiment

A.2.2.1 Method

Before starting Experiment 1, participants completed a short practice task (40 trials) and their results were used to inform the procedure of the main experiment. During the practice, participants were presented the standard ball and either the largest or the smallest of the eight comparison balls. They were asked to indicate which of two balls was bigger using either vision-only or touch-only. During ten vision-only trials, the experimenter presented the balls consecutively (one ball presented after the other), whilst during the other ten vision-only trials, balls were presented simultaneously (both balls presented at the same time). During ten haptic-only trials, participants were asked either to tap each ball with the flat palm of their dominant hand, whilst during the other ten haptic-only trials, participants were asked to grasp each ball (without lifting) with their dominant hand. In an attempt to more closely match visual-only and haptic-only discrimination reliabilities, the visual presentation style (consecutive vs. simultaneous ball presentation) with the highest score and the haptic presentation style (tapping vs. grasping) with the lowest score were used for the main experiment.

A.2.2.2 Results

Four of five participants (001, 003, 004, & 005) showed better size discrimination performance using vision when the comparison balls were presented simultaneously with the standard ball (mean correct = 67.5%), rather than consecutively (mean correct = 50%). The other participant (002) showed no difference between presentation types (mean correct = 70%). All five participants performed equally well in both the haptic-tapping and haptic-grasping conditions (mean correct = 98%). Based on these results, in the main experiment, all participants were presented comparison and standard stimuli simultaneously, and were asked to tap balls when making haptic judgments.

A.2.3 Visual-Haptic Results with Error Bars

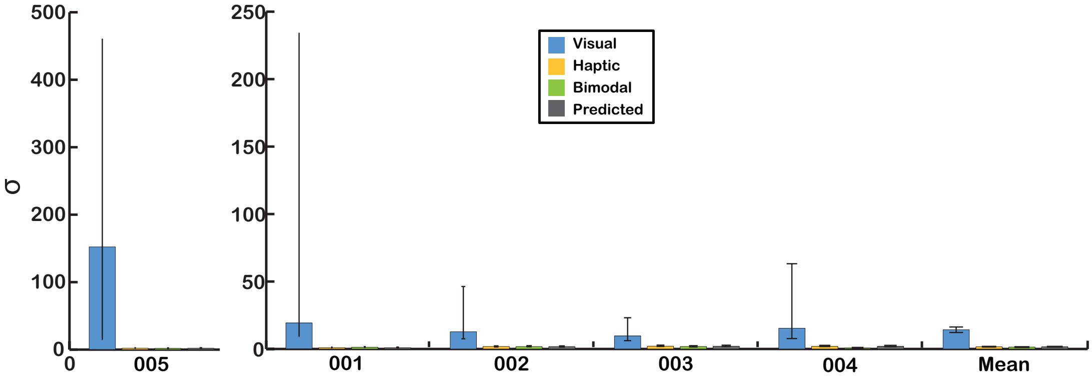


Figure 16: Unisensory (visual-only, haptic-only) and bimodal sigma with error bars for participant data representing 95% confidence limits for the measures. (Error bars for mean measures represent the standard error of the mean).

A.2.4 Proportion of Correct Discrimination Judgments

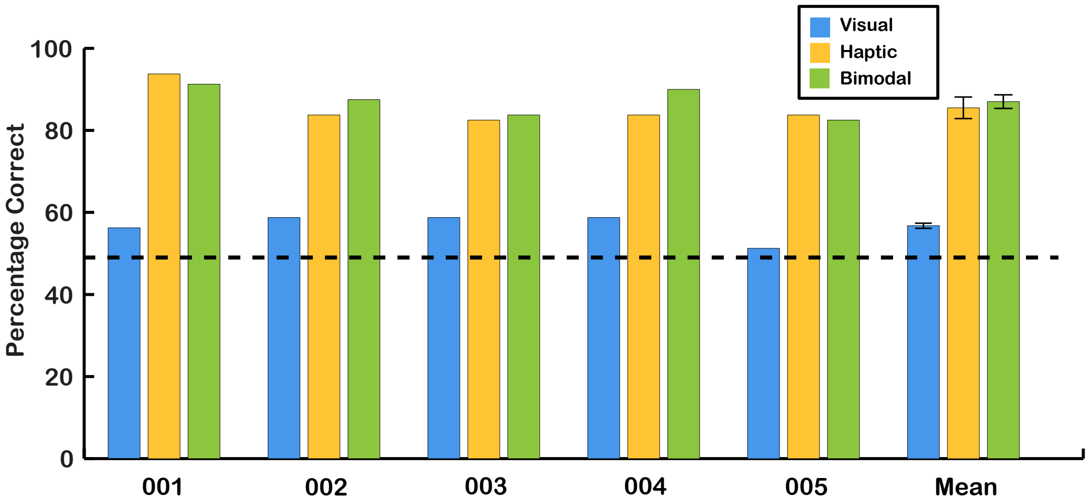


Figure 17: Percentage of trials in which participants correctly discriminated ball size using visual, haptic or bimodal Information. Dashed line indicates performance at chance (50%).

A.2.5 Race Model Predictions for Bimodal Reaction Time Advantages

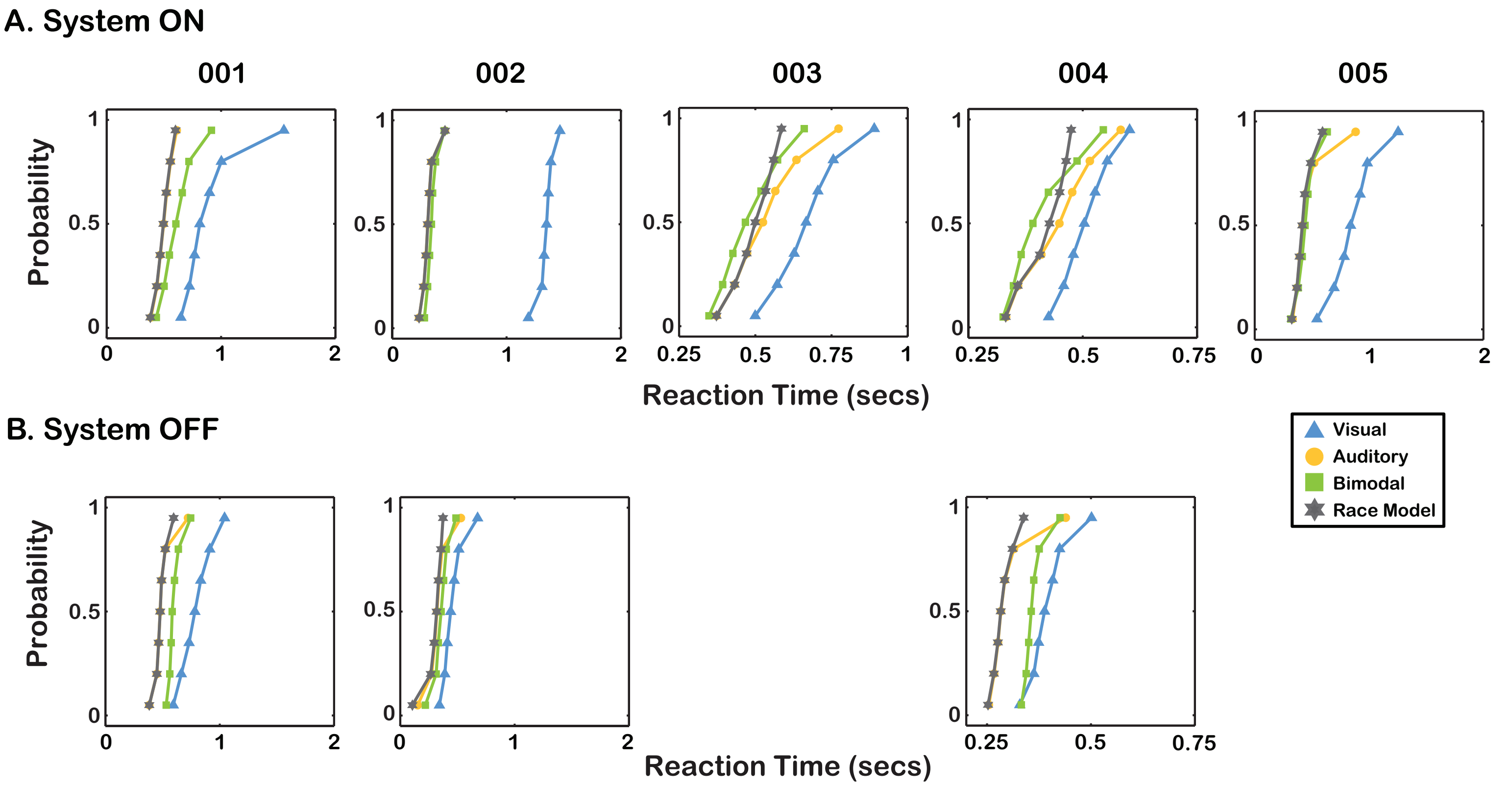


Figure 18: Reaction time cumulative distribution functions (CDFs) to unisensory and bimodal stimuli. The race model prediction, (the sum of the unisensory reaction time CDFs), indicates the largest reaction time advantage participants could achieve by responding to the visual or auditory cue that finishes processing first. Only race model predictions for participants 003 and 004 using the retinal prosthesis (system on) exceeded the reaction times of the best unisensory (auditory) cue. In all other cases, race model predictions and auditory CDFs are matched (overlapping), except at the 95th percentile for participant 005, and system off performance.

A.3 Chapter 3

A.3.1 Patients Navigation with Residual Vision

Three patients who were able to locate the landmark using their residual vision also completed the task with the system off. Therefore they completed the task using prosthetic vision (system-on), residual vision (system-off) and no vision (blindfolded and landmark light off). Given the small number of patients tested, the use of paired-sample *t*-tests to assess differences between conditions is not advised. Instead, we have compared each patient's improvement in performance with residual vision to the 95% confidence intervals calculated from participants with normal vision (see Table 10).

Any improvement by patients with residual vision compared to without vision did not match that based on the controls' performance, except: (i) On the triangle completion task, patients 001 and 002 showed similar improvements in variable error to controls when using residual vision, compared to without vision (but interestingly, were the two patients that had shown worse variable error using prosthetic vision, compared to without vision). (ii) On the path reproduction task, patient 003 showed a similar improvement in constant error to controls when using residual vision compared to without vision (but had shown worse constant error with the retinal prosthesis compared to without vision). These results highlight that there are differences between prosthetic and native vision, and difficulties using prosthetic vision clearly impacted patients' performance on this task.

Table 10: Reduction in Error (m) with Residual Vision compared to Without Vision

Patients' performance compared to the reduction in error when navigating with vision, compared to without vision, that 95% of sighted controls would be expected to show (95% CI).

	Path Reproduction		Triangle Completion	
	Variable Error	**Constant Error**	**Variable Error**	**Constant Error**
95% CI	0.105 – 0.402	0.027 – 0.415	0.038 – 0.346	No Improvement
001	-0.05	-0.6	0.13	-0.07
002	0.05	0.01	0.22	0.17
003	-0.01	0.32	-0.01	-0.23

A.3.2 Controls Navigation with Restricted Field of View Only

In the 'vision' condition, controls completed the task using goggles that both restricted their field of view and visual resolution. An additional condition was included, in which goggles only restricted their field of view, but not their visual resolution. Table 11 shows the improvement in performance compared to without vision for both these conditions. Although, as expected, mean improvements in variable and constant errors in both tasks were greater when resolution was not restricted, the results of paired samples t-tests indicated that differences in improvements (between restricted field of view and resolution versus restricted field of view only) were not significant.

Table 11: Improvements in Variable and Constant Errors (m) by Control Participants when using Vision that was Restricted in Field of View and Resolution, or Restricted in Field of View Only

	Path Reproduction		Triangle Completion	
	Variable Error	Constant Error	Variable Error	Constant Error
Restricted field of view AND resolution	0.25	0.22	0.19	0.29
Restricted field of view only	0.28	0.32	0.32	0.48
Result of paired t-test comparing differences	$t_{[10]} = 1.18$ $p = 0.07$	$t_{[10]} = 2.09$ $p = 0.06$	$t_{[10]} = 1.74$ $p = 0.24$	$t_{[10]} = 1.09$ $p = 0.30$

A.3.3 Constant and Variable Errors With and Without Vision

Table 3: Variable Errors (m) Without Vision and With Vision for Path Reproduction and Triangle Completion Tasks

	Path Reproduction			Triangle Completion		
	Without Vision	With Vision	Improvement	Without Vision	With Vision	Improvement
Controls (M)	0.33	0.08	0.25	0.50	0.31	0.19
001	0.12	0.18	-0.06	0.22	0.09	0.13
002	0.08	0.02	0.06	0.29	0.11	0.18
003	0.03	0.04	-0.01	0.05	0.15	-0.10
004	0.11	0.29	-0.18	0.15	0.43	-0.28

Table 13: Constant Errors (m) Without Vision and With Vision for Path Reproduction and Triangle Completion Tasks

	Path Reproduction			Triangle Completion		
	Without Vision	**With Vision**	**Improvement**	**Without Vision**	**With Vision**	**Improvement**
Controls (M)	0.63	0.41	0.22	1.08	0.78	0.30
001	0.33	1.10	-0.77	1.57	1.48	0.09
002	0.13	0.28	-0.15	1.34	1.27	0.07
003	0.63	0.74	-0.11	0.33	0.31	0.02
004	0.18	0.19	-0.01	0.17	0.76	-0.59

A.3.4 Learning within Conditions

No significant effect of trial number on constant error, or block number on variable error, within conditions was found (see Table 14).

Table 14: Paired t-Test Results of Effect of Block Number on Variable Error and Linear Regression Results of Effect of Trial Number on Constant Error

		Path Reproduction		Triangle Completion	
		Vision	No Vision	Vision	No Vision
Variable Error	Controls	$t_{[10]} = 0.80$, $p = 0.44$	$t_{[10]} = 0.15$, $p = 0.89$	$t_{[10]} = 0.38$, $p = 0.71$	$t_{[10]} = 1.12$, $p = 0.29$
	Patients	$t_{[10]} = 0.96$, $p = 0.41$	$t_{[10]} = 0.52$, $p = 0.64$	$t_{[10]} = 0.05$, $p = 0.96$	$t_{[10]} = 1.47$, $p = 0.24$
Constant Error	Controls	$F_{[2,8]} = 0.90$, $p = 0.37$	$F_{[2,8]} = 0.54$, $p = 0.48$	$F_{[2,8]} = 1.74$, $p = 0.22$	$F_{[2,8]} = 1.91$, $p = 0.21$
	Patients	$F_{[2,8]} = 2.46$, $p = 0.16$	$F_{[2,8]} = 0.76$, $p = 0.41$	$F_{[2,8]} = 1.09$, $p = 0.33$	$F_{[2,8]} = 1.94$, $p = 0.20$

A.4 Chapter 4

A.4.1 Audio-Visual Localization in Sighted Young Adults

A.4.1.2 Method

Participants

12 normally sighted adults aged 18 to 30 years (4 male, M = 23.7 yrs, SD = 4.3) with normal sight and normal hearing were recruited through the UCL psychology online subject pool. Informed written consent, according to the Tenets of the Declaration of Helsinki, was obtained from all participants prior to participation.

Apparatus & Stimuli

As reported here, except that an additional visual stimulus was used, so as to manipulate the reliability of the visual cue within central and peripheral spaces. As before, all 122 LEDs were powered to show white light (2223 cd/m^2) constantly throughout the duration of the experiment, and the visual stimulus used previously (i.e. a 25 msec flash of white light from 50 adjacent LEDs) was used here as the 'more reliable' visual stimulus (V1). The 'less reliable' visual stimulus (V2) was a 25 msec flash from 3 LEDs, selected from a sample of 50 LEDs with probability specified by a normal distribution. Selection was controlled to ensure that the mean of the 3 LEDs selected always reflected the mean of the underlying sample of 50 LEDs.

Procedure

As reported here, except that stimuli sets consisted of either unimodal stimuli (audio-only, V1-only, V2-only) or bimodal stimuli (audio-V1, audio-V2). Again, where visual and auditory stimuli were presented together, stimuli were either presented in congruent locations (no-conflict), or the visual stimulus was displaced leftward (central: by 3°, peripheral: by 4°) compared to the auditory stimulus (conflict). Participants completed 49 test blocks (7 audio-only, 7 V1-only, 7 V2-only, 14 audio-V1, 14 audio-V2) of 64 trials each.

A.4.1.3 Results

Figure 19.A shows the mean reliability for the single cue (audition-only, vision-only) and bimodal conditions, and the ideal (MLE) predictions, in central and peripheral conditions. Bimodal reliability was well predicted by the MLE for both visual stimuli (Central V1: $F_{[2,10]} = 21.6$, $p < 0.001$; V2: $F_{[2,10]} = 16.0$, $p = 0.003$; Peripheral V1: $F_{[2,10]} = 9.8$, $p = 0.01$; V2: $F_{[2,10]} = 24.7$, $p < 0.001$). Visual and auditory reliabilities were best matched in peripheral space, when localizing the more reliable visual stimulus. Here, a significant reduction in the bimodal discrimination threshold relative to the best unisensory cue was observed ($t_{[11]} = 4.2$, p = 0.002). A 2 (central, peripheral) x 2 (high visual reliability, low visual reliability) repeated measured ANOVA indicated both a main effect of visual stimulus reliability (measured: $F_{[1,11]} = 5.4$, $p = 0.04$; predicted: $F_{[1,11]} = 22.5$, $p = 0.001$) and a main effect of position in visual field (measured: $F_{[1,11]} = 5.9$, $p = 0.03$; predicted: $F_{[1,11]} = 42.2$, $p < 0.001$) on the measured and predicted weighting given to vision. Specifically, visual weights were significantly higher for the more reliable than the less reliable visual cue in both central and peripheral locations. Predicted visual weights (V1: $t_{[11]} = 7.8$, $p < 0.001$;

V2: $t_{[11]} = 0.2$, $p = 0.81$) and measured visual weights (V1: $t_{[11]} = 3.1$, $p = 0.01$; V2: $t_{[11]} = 1.2$, $p = 0.26$) were significantly higher in central than peripheral space for the more reliable visual cue (V1), but not the less reliable visual cue (V2).

Paired sample t-tests showed no significant difference between predicted and measured weights for the more reliable visual cue, in both central ($t_{[11]} = 1.6$, $p = 0.15$) and peripheral ($t_{[11]} = 0.2$, $p = 0.86$) conditions. Indeed, as depicted in Figure 19.B, participants adjusted weights accordingly for the more reliable visual cue (V1: $F_{[2,10]} = 4.6$, $p = 0.058$) but not the less reliable visual cue (V2: $F_{[2,10]} = 0.5$, $p = 0.48$). Paired sample t-tests indicated no differences between predicted and measured weights for the less reliable cue in peripheral space ($t_{[11]} = 0.3$, $p = 0.74$), but participants tended to overweight less reliable vision in central space ($t_{[11]} = 2.5$, $p = 0.03$).

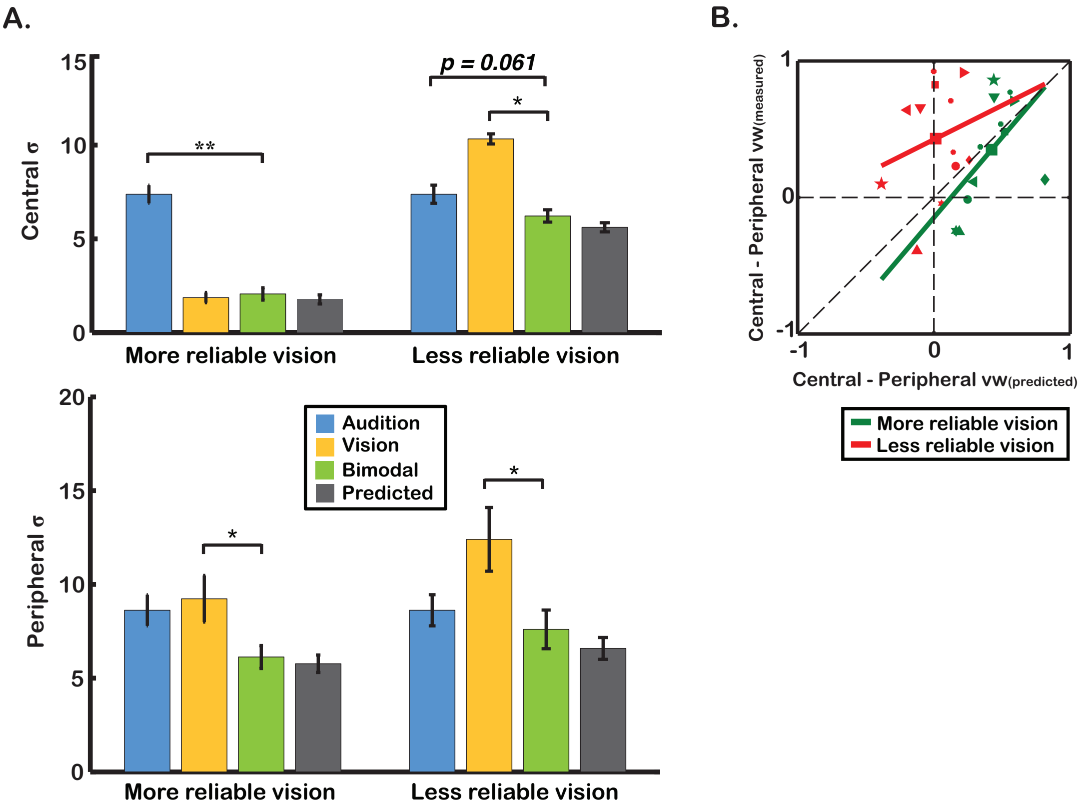


Figure 19: A: Measured and Predicted Reliability in Young Adults. The bars show the standard error of the mean. (* indicates $p < .05$; ** indicates $p < .01$). B: Differences between Measured Visual Weights in Central and Peripheral Space against Differences between Predicted Visual Weights in Central and Peripheral Space.

A.5 Chapter 5

A.5.1 Effect of visual feedback on bias and variability

Phase (before-, during, after- training) had a significant effect on bias ($F_{[1.38,30.26]} = 12.44$, $p < 0.001$), but visual feedback reliability did not ($F_{[1,22]} = 0.02$, $p = 0.896$), and there was no interaction between visual feedback and phase ($F_{[1.38,30.26]} = 0.97$, $p = 0.360$; see Fig. 20A). Phase also had a significant effect on variability ($F_{[2,44]} = 17.06$, $p < 0.001$), but again visual feedback reliability did not ($F_{[1,22]} = 0.26$, $p = 0.614$). However, there was a significant interaction between phase and visual feedback reliability on localisation variability ($F_{[2,44]} = 5.65$, $p = 0.007$): Participants trained with less reliable visual feedback, showed significantly reduced variability during-training than before-training ($t_{[11]} = 6.19$, $p < 0.001$), while for participants trained with more reliable visual feedback, the reduction in variability was not significant ($t_{[11]} = 2.00$, $p = 0.071$).

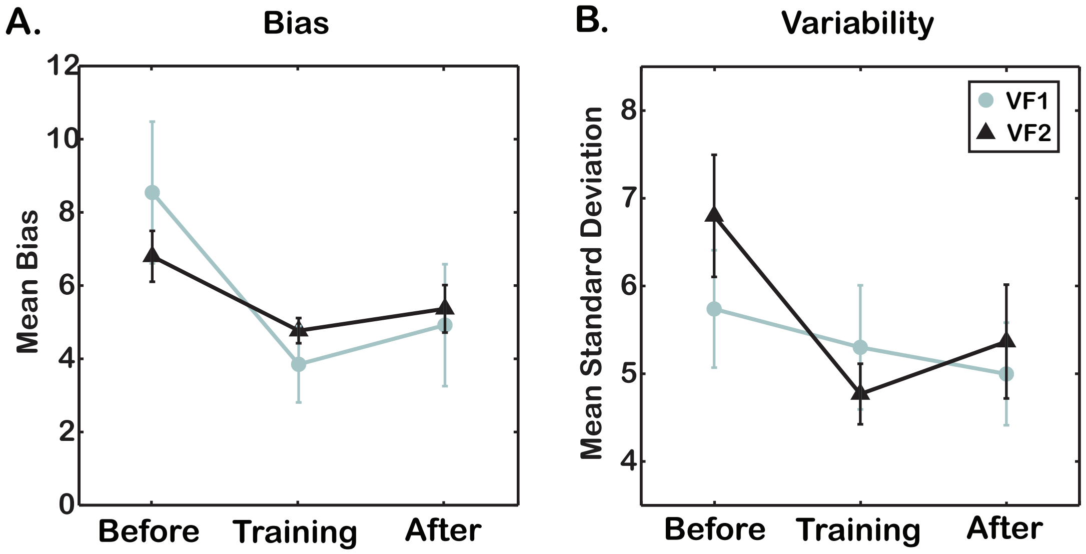


Figure 20: Mean Bias (A) and Mean Variability (B) Before-, During- and After- Training for Participants trained with More Reliable Visual Feedback (VF1) and Less Reliable Visual Feedback (VF2). Bars represent standard error of the mean.

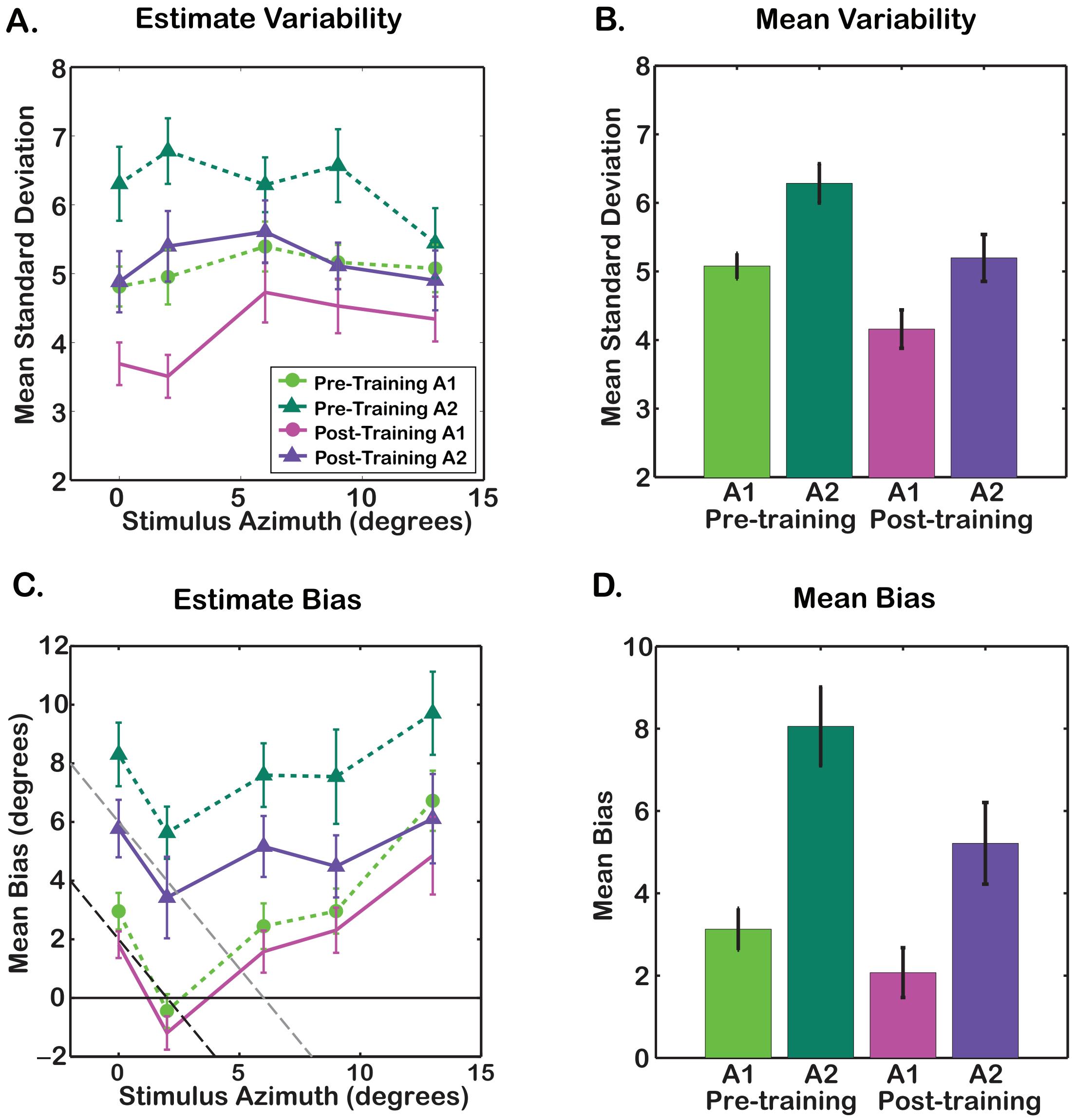


Figure 21: Mean Bias and Variability for Localizing Auditory and Visual Stimuli Before- and After-Training.

A. Mean variability at each location for each stimulus. B. Mean variability at each location for each stimulus. C. Mean bias at each location for each stimulus. D. Mean bias for each stimulus across all locations tested. Grey dotted line in C indicates the line predicted by responding according to the mean of the target stimulus set. Black dotted line in C indicates the line predicted by responding according to the mean of the speakers presenting background noise. Error bars represent the standard error of the mean.

References

Ahuja, A. K., Dorn, J. D., Caspi, A., McMahon, M. J., Dagnelie, G., Dacruz, L., . . . Argus, I. I. S. G. (2011). Blind subjects implanted with the Argus II retinal prosthesis are able to improve performance in a spatial-motor task. *Br J Ophthalmol, 95*(4), 539-543. doi:10.1136/bjo.2010.179622

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol, 14*(3), 257-262. doi:10.1016/j.cub.2004.01.029

Alary, F., Duquette, M., Goldstein, R., Elaine Chapman, C., Voss, P., La Buissonnière-Ariza, V., & Lepore, F. (2009). Tactile acuity in the blind: A closer look reveals superiority over the sighted in some but not all cutaneous tasks. *Neuropsychologia, 47*(10), 2037-2043. doi:10.1016/j.neuropsychologia.2009.03.014

Amedi, A., Raz, N., Azulay, H., Malach, R., & Zohary, E. (2010). Cortical activity during tactile exploration of objects in blind and sighted humans. *Restor Neurol Neurosci, 28*(2), 143-156. doi:10.3233/rnn-2010-0503

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Brain Res Cogn Brain Res, 21*(3), 301-308. doi:10.1016/j.cogbrainres.2004.06.004

Ashmead, D. H., Wall, R. S., Ebinger, K. A., Eaton, S. B., Snook-Hill, M. M., & Yang, X. (1998). Spatial hearing in children with visual disabilities. *Perception, 27*(1), 105-122.

Barth, D. S., Goldberg, N., Brett, B., & Di, S. (1995). The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain Res, 678*(1-2), 177-190.

Bates, S. L., & Wolbers, T. (2014). How cognitive aging affects multisensory integration of navigational cues. *Neurobiology of Aging*(0). doi:10.1016/j.neurobiolaging.2014.04.003

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J Opt Soc Am A Opt Image Sci Vis, 20*(7), 1391-1397.

Bausenhart, K. M., de la Rosa, M. D., & Ulrich, R. (2014). Multimodal integration of time. *Exp Psychol, 61*(4), 310-322. doi:10.1027/1618-3169/a000249

Bavelier, D., & Neville, H. J. (2002). Cross-modal plasticity: where and how? *Nat Rev Neurosci, 3*(6), 443-452. doi:10.1038/nrn848

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept Psychophys, 29*(6), 578-584.

Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb Cortex, 17*(9), 2172-2189. doi:10.1093/cercor/bhl128

Bock, O. (1986). Contribution of retinal versus extraretinal signals towards visual localization in goal-directed movements. *Exp Brain Res, 64*(3), 476-482.

Bolognini, N., Cecchetto, C., Geraci, C., Maravita, A., Pascual-Leone, A., & Papagno, C. (2012). Hearing shapes our perception of time: temporal discrimination of tactile stimuli in deaf people. *J Cogn Neurosci, 24*(2), 276-286. doi:10.1162/jocn_a_00135

Bottari, D., Caclin, A., Giard, M. H., & Pavani, F. (2011). Changes in early cortical visual processing predict enhanced reactivity in deaf individuals. *PLoS One, 6*(9), e25607. doi:10.1371/journal.pone.0025607

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433-436.

Brenton, R. S., & Phelps, C. D. (1986). The normal visual field on the Humphrey field analyzer. *Ophthalmologica, 193*(1-2), 56-74.

Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: auditory signals can modulate tactile tap perception. *Exp Brain Res, 162*(2), 172-180. doi:10.1007/s00221-004-2128-2

Buchel, C., Price, C., Frackowiak, R. S., & Friston, K. (1998). Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain, 121 (Pt 3)*, 409-419.

Buckley, K. A., & Tobey, E. A. (2011). Cross-modal plasticity and speech perception in pre- and postlingually deaf cochlear implant users. *Ear Hear, 32*(1), 2-15. doi:10.1097/AUD.0b013e3181e8534c

Burge, J., Ernst, M. O., & Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching. *J Vis, 8*(4), 20 21-19. doi:10.1167/8.4.20

Burgess, N. (2008). Spatial Cognition and the Brain. *Ann N Y Acad Sci, 1124*(1), 77-97. doi:10.1196/annals.1440.002

Burton, H., Snyder, A. Z., Conturo, T. E., Akbudak, E., Ollinger, J. M., & Raichle, M. E. (2002). Adaptive changes in early and late blind: a fMRI study of Braille reading. *J Neurophysiol, 87*(1), 589-607.

Butler, J. S., Smith, S. T., Campos, J. L., & Bulthoff, H. H. (2010). Bayesian integration of visual and vestibular signals for heading. *J Vis, 10*(11), 23. doi:10.1167/10.11.23

Butler, R. A. (1986). The bandwidth effect on monaural and binaural localization. *Hear Res, 21*(1), 67-73.

Carlile, S., Leong, P., & Hyams, S. (1997). The nature and distribution of errors in sound localization by human listeners. *Hear Res, 114*(1-2), 179-196.

Carlile, S., Martin, R., & McAnally, K. (2005). Spectral information in sound localization. *Int Rev Neurobiol, 70*, 399-434. doi:10.1016/s0074-7742(05)70012-x

Carrasco, M. M., & Pallas, S. L. (2006). Early visual experience prevents but cannot reverse deprivation-induced loss of refinement in adult superior colliculus. *Vis Neurosci, 23*(6), 845-852. doi:10.1017/s0952523806230177

Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., & Wallace, M. T. (2007). Visual deprivation alters the development of cortical multisensory integration. *J Neurophysiol, 98*(5), 2858-2867. doi:10.1152/jn.00587.2007

Chader, G. J., Weiland, J., & Humayun, M. S. (2009). Artificial vision: needs, functioning, and testing of a retinal electronic prosthesis. *Prog Brain Res, 175*, 317-332. doi:10.1016/s0079-6123(09)17522-2

Champoux, F., Collignon, O., Bacon, B. A., Lepore, F., Zatorre, R. J., & Theoret, H. (2011). Early- and late-onset blindness both curb audiotactile integration on the parchment-skin illusion. *Psychol Sci, 22*(1), 19-25. doi:10.1177/0956797610391099

Champoux, F., Lepore, F., Gagne, J. P., & Theoret, H. (2009). Visual stimuli can impair auditory processing in cochlear implant users. *Neuropsychologia, 47*(1), 17-22. doi:10.1016/j.neuropsychologia.2008.08.028

Charbonneau, G., Veronneau, M., Boudrias-Fournier, C., Lepore, F., & Collignon, O. (2013). The ventriloquist in periphery: impact of eccentricity-related reliability on audio-visual localization. *J Vis, 13*(12), 20. doi:10.1167/13.12.20

Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian integration of spatial information. *Psychol Bull, 133*(4), 625-637. doi:10.1037/0033-2909.133.4.625

Codina, C., Pascalis, O., Mody, C., Toomey, P., Rose, J., Gummer, L., & Buckley, D. (2011). Visual advantage in deaf adults linked to retinal changes. *PLoS One, 6*(6), e20417. doi:10.1371/journal.pone.0020417

Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Falz, L., Dambrosia, J., . . . Hallett, M. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature, 389*(6647), 180-183. doi:10.1038/38278

Collignon, O., Charbonneau, G., Lassonde, M., & Lepore, F. (2009). Early visual deprivation alters multisensory processing in peripersonal space. *Neuropsychologia, 47*(14), 3236-3243. doi:10.1016/j.neuropsychologia.2009.07.025

Collignon, O., Davare, M., Olivier, E., & De Volder, A. G. (2009). Reorganisation of the right occipito-parietal stream for auditory spatial processing in early blind humans. A transcranial magnetic stimulation study. *Brain Topogr, 21*(3-4), 232-240. doi:10.1007/s10548-009-0075-8

Collignon, O., Dormal, G., Albouy, G., Vandewalle, G., Voss, P., Phillips, C., & Lepore, F. (2013). Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. *Brain, 136*(Pt 9), 2769-2783. doi:10.1093/brain/awt176

Collignon, O., Lassonde, M., Lepore, F., Bastien, D., & Veraart, C. (2007). Functional cerebral reorganization for auditory spatial processing and auditory substitution of vision in early blind subjects. *Cereb Cortex, 17*(2), 457-465. doi:10.1093/cercor/bhj162

Collignon, O., Voss, P., Lassonde, M., & Lepore, F. (2009). Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. *Exp Brain Res, 192*(3), 343-358. doi:10.1007/s00221-008-1553-z

Crossland, M. D., Engel, S. A., & Legge, G. E. (2011). The preferred retinal locus in macular disease: toward a consensus definition. *Retina, 31*(10), 2109-2114. doi:10.1097/IAE.0b013e31820d3fba

Cui, Q. N., O'Neill, W. E., & Paige, G. D. (2010). Advancing age alters the influence of eye position on sound localization. *Exp Brain Res, 206*(4), 371-379. doi:10.1007/s00221-010-2413-1

Cunningham, S., Tjan, B., Bao, P., Falabella, P., & Weiland, J. (2015). Tactile-Evoked V1 responses in Argus II Retinal Prosthesis Patients assessed with fMRI: A Case Study. *J Vis, 15*(12), 359. doi:10.1167/15.12.359

Cunningham, S., Weiland, J., Bao, P., & Tjan, B. (2011). Visual cortex activation induced by tactile stimulation in late-blind individuals with retinitis pigmentosa. *Conf Proc IEEE Eng Med Biol Soc, 2011*, 2841-2844. doi:10.1109/IEMBS.2011.6090785

da Cruz, L., Coley, B. F., Dorn, J., Merlini, F., Filley, E., Christopher, P., . . . Argus, I. I. S. G. (2013). The Argus II epiretinal prosthesis system allows letter and word reading and long-term function in patients with profound vision loss. *Br J Ophthalmol, 97*(5), 632-636. doi:10.1136/bjophthalmol-2012-301525

Dacey, D. M., & Petersen, M. R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proc Natl Acad Sci U S A, 89*(20), 9666-9670.

De Meo, R., Murray, M. M., Clarke, S., & Matusz, P. J. (2015). Top-down control and early multisensory processes: chicken vs. egg. *Front Integr Neurosci, 9*, 17. doi:10.3389/fnint.2015.00017

De Volder, A. G., Catalan-Ahumada, M., Robert, A., Bol, A., Labar, D., Coppens, A., . . . Veraart, C. (1999). Changes in occipital cortex activity in early blind humans using a sensory substitution device. *Brain Res, 826*(1), 128-134.

de Winter, J. C. F. (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation, 18*(10).

Dekker, Tessa M., Ban, H., van der Velde, B., Sereno, Martin I., Welchman, Andrew E., & Nardini, M. (2015). Late Development of Cue Integration Is Linked to Sensory Fusion in Cortex. *Current Biology, 25*(21), 2856-2861. doi:10.1016/j.cub.2015.09.043

Desai, S., Stickney, G., & Zeng, F. G. (2008). Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *J Acoust Soc Am, 123*(1), 428-440. doi:10.1121/1.2816573

Dobreva, M. S., O'Neill, W. E., & Paige, G. D. (2011). Influence of aging on human sound localization. *J Neurophysiol, 105*(5), 2471-2486. doi:10.1152/jn.00951.2010

Dobreva, M. S., O'Neill, W. E., & Paige, G. D. (2012). Influence of age, spatial memory, and ocular fixation on localization of auditory, visual, and bimodal targets by human subjects. *Exp Brain Res, 223*(4), 441-455. doi:10.1007/s00221-012-3270-x

Dodd, B. (1979). Lip reading in infants: attention to speech presented in- and out-of-synchrony. *Cogn Psychol, 11*(4), 478-484.

Dormal, G., & Collignon, O. (2011). Functional selectivity in sensory-deprived cortices. *J Neurophysiol, 105*(6), 2627-2630. doi:10.1152/jn.00109.2011

Dormal, G., Lepore, F., Harissi-Dagher, M., Albouy, G., Bertone, A., Rossion, B., & Collignon, O. (2015). Tracking the evolution of crossmodal plasticity and visual functions before and after sight restoration. *J Neurophysiol, 113*(6), 1727-1742. doi:10.1152/jn.00420.2014

Dorn, J. D., Ahuja, A. K., Caspi, A., da Cruz, L., Dagnelie, G., Sahel, J. A., . . . gus, I. I. S. G. (2013). The Detection of Motion by Blind Subjects With the Epiretinal 60-Electrode (Argus II) Retinal Prosthesis. *JAMA Ophthalmol, 131*(2), 183-189. doi:10.1001/2013.jamaophthalmol.221

Doucet, M. E., Bergeron, F., Lassonde, M., Ferron, P., & Lepore, F. (2006). Cross-modal reorganization and speech perception in cochlear implant users. *Brain, 129*(Pt 12), 3376-3383. doi:10.1093/brain/awl264

Doucet, M. E., Guillemot, J. P., Lassonde, M., Gagne, J. P., Leclerc, C., & Lepore, F. (2005). Blind subjects process auditory spectral cues more efficiently than sighted individuals. *Exp Brain Res, 160*(2), 194-202. doi:10.1007/s00221-004-2000-4

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron, 57*(1), 11-23. doi:10.1016/j.neuron.2007.12.013

Dufour, A., Despres, O., & Candas, V. (2005). Enhanced sensitivity to echo cues in blind subjects. *Exp Brain Res, 165*(4), 515-519. doi:10.1007/s00221-005-2329-3

Elbert, T., Sterr, A., Rockstroh, B., Pantev, C., Muller, M. M., & Taub, E. (2002). Expansion of the tonotopic area in the auditory cortex of the blind. *J Neurosci, 22*(22), 9941-9944.

Emmorey, K., Mehta, S., & Grabowski, T. J. (2007). The neural correlates of sign versus word production. *Neuroimage, 36*(1), 202-208. doi:10.1016/j.neuroimage.2007.02.040

Ernst, M. O. (2006a). A Bayesian view on multimodal cue integration *Human Body Perception from the Inside Out.* (pp. 105-131). New York: Oxford University Press.

Ernst, M. O. (2006b). A Bayesian view on multimodal cue integration. In G. Knoblich, M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human Body Perception From The Inside Out* (pp. 105-131). New York: Oxford University Press.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *J Vis, 7*(5), 7 1-14. doi:10.1167/7.5.7

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429-433. doi:10.1038/415429a

Ernst, M. O., Banks, M. S., & Bulthoff, H. H. (2000). Touch can change visual slant perception. *Nat Neurosci, 3*(1), 69-73. doi:10.1038/71140

Falchier, A., Schroeder, C. E., Hackett, T. A., Lakatos, P., Nascimento-Silva, S., Ulbert, I., . . . Smiley, J. F. (2010). Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cereb Cortex, 20*(7), 1529-1538. doi:10.1093/cercor/bhp213

Feldman, D. E., & Knudsen, E. I. (1997). An anatomical basis for visual calibration of the auditory space map in the barn owl's midbrain. *J Neurosci, 17*(17), 6820-6837.

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nat Rev Neurosci, 14*(6), 429-442. doi:10.1038/nrn3503

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., & Angelaki, D. E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. *J Neurosci, 29*(49), 15601-15612. doi:10.1523/JNEUROSCI.2574-09.2009

Fieger, A., Roder, B., Teder-Salejarvi, W., Hillyard, S. A., & Neville, H. J. (2006). Auditory spatial tuning in late-onset blindness in humans. *J Cogn Neurosci, 18*(2), 149-157. doi:10.1162/089892906775783697

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., . . . MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nat Neurosci, 6*(9), 915-916. doi:10.1038/nn1102

Finney, E. M., Fine, I., & Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nat Neurosci, 4*(12), 1171-1173. doi:10.1038/nn763

Finocchietti, S., Cappagli, G., & Gori, M. (2015). Encoding audio motion: spatial impairment in early blind individuals. *Front Psychol, 6*, 1357. doi:10.3389/fpsyg.2015.01357

Fishman, M. C., & Michael, P. (1973). Integration of auditory information in the cat's visual cortex. *Vision Res, 13*(8), 1415-1419.

Foo, P., Warren, W. H., Duchon, A., & Tarr, M. J. (2005). Do humans integrate routes into a cognitive map? Map- versus landmark-based navigation of novel shortcuts. *J Exp Psychol Learn Mem Cogn, 31*(2), 195-215. doi:10.1037/0278-7393.31.2.195

Forster, B., Cavina-Pratesi, C., Aglioti, S. M., & Berlucchi, G. (2002). Redundant target effect and intersensory facilitation from visual-tactile interactions in simple reaction time. *Exp Brain Res, 143*(4), 480-487. doi:10.1007/s00221-002-1017-9

Fortenbaugh, F. C., & Robertson, L. C. (2011). When here becomes there: attentional distribution modulates foveal bias in peripheral localization. *Atten Percept Psychophys, 73*(3), 809-828. doi:10.3758/s13414-010-0075-5

Fortin, M., Voss, P., Lord, C., Lassonde, M., Pruessner, J., Saint-Amour, D., . . . Lepore, F. (2008). Wayfinding in the blind: larger hippocampal volume and supranormal spatial navigation. *Brain, 131*(Pt 11), 2995-3005. doi:10.1093/brain/awn250

Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., & Ladavas, E. (2005). Audiovisual integration in patients with visual deficit. *J Cogn Neurosci, 17*(9), 1442-1452. doi:10.1162/0898929054985446

Frissen, I., Vroomen, J., de Gelder, B., & Bertelson, P. (2005). The aftereffects of ventriloquism: generalization across sound-frequencies. *Acta Psychol (Amst), 118*(1-2), 93-100. doi:10.1016/j.actpsy.2004.10.004

Gagnon, L., Kupers, R., Schneider, F. C., & Ptito, M. (2010). Tactile maze solving in congenitally blind individuals. *Neuroreport, 21*(15), 989-992. doi:10.1097/WNR.0b013e32833eaaf9

Gallistel, C. R. (1990). Representations in animal cognition: an introduction. *Cognition, 37*(1-2), 1-22.

Garcia, S., Jones, P., Rubin, G., & Nardini, M. (2015). Visual-auditory localization in central and peripheral space. *J Vis, 15*(12), 719. doi:10.1167/15.12.719

Garcia, S., Petrini, K., da Cruz, L., Rubin, G., & Nardini, M. (2014). Cue combination with a new sensory signal: multisensory processing in blind patients with a retinal prosthesis. *J Vis, 14*(10), 1132-1132. doi:10.1167/14.10.1132

Garcia, S., Petrini, K., Rubin, G. S., Da Cruz, L., & Nardini, M. (2015). Visual and Non-Visual Navigation in Blind Patients with a Retinal Prosthesis. *PLoS One, 10*(7), e0134369. doi:10.1371/journal.pone.0134369

Gepshtein, S., & Banks, M. S. (2003). Viewing geometry determines how vision and haptics combine in size perception. *Curr Biol, 13*(6), 483-488.

Gepshtein, S., Burge, J., Ernst, M. O., & Banks, M. S. (2005). The combination of vision and touch depends on spatial proximity. *J Vis, 5*(11), 1013-1023. doi:10.1167/5.11.7

Geruschat, D. R., Flax, M., Tanna, N., Bianchi, M., Fisher, A., Goldschmidt, M., . . . Dorn, J. (2015). FLORA: Phase I development of a functional vision assessment for prosthetic vision users. *Clin Exp Optom, 98*(4), 342-347. doi:10.1111/cxo.12242

Getzmann, S. (2007). The effect of brief auditory stimuli on visual apparent motion. *Perception, 36*(7), 1089-1103.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn Sci, 10*(6), 278-285. doi:10.1016/j.tics.2006.04.008

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning; differentiation or enrichment? *Psychol Rev, 62*(1), 32-41.

Girard, S., Collignon, O., & Lepore, F. (2011). Multisensory gain within and across hemispaces in simple and choice reaction time paradigms. *Exp Brain Res, 214*(1), 1-8. doi:10.1007/s00221-010-2515-9

Gold, J. I., Law, C. T., Connolly, P., & Bennur, S. (2008). The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *J Neurophysiol, 100*(5), 2653-2668. doi:10.1152/jn.90629.2008

Goldreich, D., & Kanics, I. M. (2003). Tactile acuity is enhanced in blindness. *J Neurosci, 23*(8), 3439-3445.

Gondan, M., Lange, K., Rosler, F., & Roder, B. (2004). The redundant target effect is affected by modality switch costs. *Psychon Bull Rev, 11*(2), 307-313.

Gori, M. (2015). Multisensory Integration and Calibration in Children and Adults with and without Sensory and Motor Disabilities. *Multisens Res, 28*(1-2), 71-99.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Curr Biol, 18*(9), 694-698. doi:10.1016/j.cub.2008.04.036

Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Front Integr Neurosci, 6*, 77. doi:10.3389/fnint.2012.00077

Gori, M., Sandini, G., Martinoli, C., & Burr, D. (2010). Poor haptic orientation discrimination in nonsighted children may reflect disruption of cross-sensory calibration. *Curr Biol, 20*(3), 223-225. doi:10.1016/j.cub.2009.11.069

Gori, M., Sandini, G., Martinoli, C., & Burr, D. C. (2014). Impairment of auditory spatial localization in congenitally blind human subjects. *Brain, 137*(Pt 1), 288-293. doi:10.1093/brain/awt311

Gori, M., Tinelli, F., Sandini, G., Cioni, G., & Burr, D. (2012). Impaired visual size-discrimination in children with movement disorders. *Neuropsychologia, 50*(8), 1838-1843. doi:10.1016/j.neuropsychologia.2012.04.009

Gougoux, F., Zatorre, R. J., Lassonde, M., Voss, P., & Lepore, F. (2005). A functional neuroimaging study of sound localization: visual cortex activity predicts performance in early-blind individuals. *PLoS Biol, 3*(2), e27. doi:10.1371/journal.pbio.0030027

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*: New York: Wiley.

Gregory, R. L., & Wallace, J. G. (1963). Recovery from early blindness: a case study. In R. L. Gregory (Ed.), *Concepts and Mechanisms of Perception* (pp. 65-129). London: Duckworth.

Gu, Y., Angelaki, D. E., & Deangelis, G. C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nat Neurosci, 11*(10), 1201-1210. doi:10.1038/nn.2191

Harrington, L. K., & Peck, C. K. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Exp Brain Res, 122*(2), 247-252.

Heimler, B., Weisz, N., & Collignon, O. (2014). Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience, 283*, 44-63. doi:10.1016/j.neuroscience.2014.08.003

Helbig, H. B., & Ernst, M. O. (2007). Optimal integration of shape information from vision and touch. *Exp Brain Res, 179*(4), 595-606. doi:10.1007/s00221-006-0814-y

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. *Science, 298*(5598), 1627-1630. doi:10.1126/science.1075396

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: optimal cue combination. *J Vis, 4*(12), 967-992. doi:10:1167/4.12.1

Hong Lore, W., & Song, S. (1991). Central and peripheral visual processing in hearing and nonhearing individuals. *Bulletin of the Psychonomic Society, 29*(5), 437-440. doi:10.3758/BF03333964

Hoover, A. E., Harris, L. R., & Steeves, J. K. (2012). Sensory compensation in sound localization in people with one eye. *Exp Brain Res, 216*(4), 565-574. doi:10.1007/s00221-011-2960-0

Hotting, K., & Roder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychol Sci, 15*(1), 60-64.

Hotting, K., & Roder, B. (2009). Auditory and auditory-tactile processing in congenitally blind humans. *Hear Res, 258*(1-2), 165-174. doi:10.1016/j.heares.2009.07.012

Hotting, K., Rosler, F., & Roder, B. (2004). Altered auditory-tactile interactions in congenitally blind humans: an event-related potential study. *Exp Brain Res, 159*(3), 370-381. doi:10.1007/s00221-004-1965-3

Humayun, M. S., Dorn, J. D., da Cruz, L., Dagnelie, G., Sahel, J.-A., Stanga, P. E., . . . Greenberg, R. J. (2012). Interim Results from the International Trial of Second Sight's Visual Prosthesis. *Ophthalmology, 119*(4), 779-788. doi:10.1016/j.ophtha.2011.09.028

Hyvarinen, J., Carlson, S., & Hyvarinen, L. (1981). Early visual deprivation alters modality of neuronal responses in area 19 of monkey cortex. *Neurosci Lett, 26*(3), 239-243.

Isaiah, A., Vongpaisal, T., King, A. J., & Hartley, D. E. (2014). Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *J Neurosci, 34*(33), 11119-11130. doi:10.1523/jneurosci.4767-13.2014

Izraeli, R., Koay, G., Lamish, M., Heicklen-Klein, A. J., Heffner, H. E., Heffner, R. S., & Wollberg, Z. (2002). Cross-modal neuroplasticity in neonatally enucleated hamsters: structure, electrophysiology and behaviour. *Eur J Neurosci, 15*(4), 693-712.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat Neurosci, 13*(8), 1020-1026. doi:10.1038/nn.2590

Jones, P. R., Garcia, S. E., & Nardini, M. (2015). Digital LED Pixels: Instructions for use and a characterization of their properties. *Behav Res Methods*. doi:10.3758/s13428-015-0653-5

Jousmaki, V., & Hari, R. (1998). Parchment-skin illusion: sound-biased touch. *Curr Biol, 8*(6), R190.

Kadunce, D. C., Vaughan, J. W., Wallace, M. T., Benedek, G., & Stein, B. E. (1997). Mechanisms of within- and cross-modality suppression in the superior colliculus. *J Neurophysiol, 78*(6), 2834-2847.

Kaiser, A. R., Kirk, K. I., Lachs, L., & Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *J Speech Lang Hear Res, 46*(2), 390-404.

Kalia, A., Lesmes, L. A., Dorr, M., Gandhi, T., Chatterjee, G., Ganesh, S., . . . Sinha, P. (2014). Development of pattern vision following early and extended blindness. *Proc Natl Acad Sci U S A, 111*(5), 2035-2039. doi:10.1073/pnas.1311041111

Kalia, A., Schrater, P., & Legge, G. (2013). Combining Path Integration and Remembered Landmarks When Navigating without Vision. *PLoS One, 8*(9), e72170. doi:10.1371/journal.pone.0072170

Karns, C. M., Dow, M. W., & Neville, H. J. (2012). Altered cross-modal processing in the primary auditory cortex of congenitally deaf adults: a visual-somatosensory fMRI study with a double-flash illusion. *J Neurosci, 32*(28), 9626-9638. doi:10.1523/jneurosci.6488-11.2012

Keetels, M., & Vroomen, J. (2012). Frontiers in Neuroscience: Perception of Synchrony between the Senses. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press/Taylor & Francis Llc.

Keller, I., & Lefin-Rank, G. (2010). Improvement of visual search after audiovisual exploration training in hemianopic patients. *Neurorehabil Neural Repair, 24*(7), 666-673. doi:10.1177/1545968310372774

Kellman, P. J., & Garrigan, P. (2009). Perceptual learning and human expertise. *Phys Life Rev, 6*(2), 53-84. doi:10.1016/j.plrev.2008.12.001

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annu Rev Psychol, 55*, 271-304. doi:10.1146/annurev.psych.55.090902.142005

Kerzel, D. (2002). Memory for the position of stationary objects: disentangling foveal bias and memory averaging. *Vision Res, 42*(2), 159-167. doi:10.1016/S0042-6989(01)00274-7

King, A. J. (2015). Crossmodal plasticity and hearing capabilities following blindness. *Cell Tissue Res, 361*(1), 295-300. doi:10.1007/s00441-015-2175-y

King, A. J., & Carlile, S. (1993). Changes induced in the representation of auditory space in the superior colliculus by rearing ferrets with binocular eyelid suture. *Exp Brain Res, 94*(3), 444-455.

King, A. J., Hutchings, M. E., Moore, D. R., & Blakemore, C. (1988). Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature, 332*(6159), 73-76. doi:10.1038/332073a0

King, A. J., & Parsons, C. H. (1999). Improved auditory spatial acuity in visually deprived ferrets. *Eur J Neurosci, 11*(11), 3945-3956.

Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Doherty, S., & Pellegrino, J. W. (1990). Acquisition of route and survey knowledge in the absence of vision. *J Mot Behav, 22*(1), 19-43.

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). "What's new in Psychtoolbox-3?". *Perception, 36*(ECVP Abstract Supplement.), 14. doi:10.1177/03010066070360s101

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res, 43*(24), 2539-2558.

Knudsen, E. I., & Brainard, M. S. (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. *Science, 253*(5015), 85-87.

Kording, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS One, 2*(9), e943. doi:10.1371/journal.pone.0000943

Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature, 427*(6971), 244-247. doi:10.1038/nature02169

Korte, M., & Rauschecker, J. P. (1993). Auditory spatial tuning of cortical neurons is sharpened in cats with early blindness. *J Neurophysiol, 70*(4), 1717-1721.

Kowalska, J., & Szelag, E. (2006). The effect of congenital deafness on duration judgment. *J Child Psychol Psychiatry, 47*(9), 946-953. doi:10.1111/j.1469-7610.2006.01591.x

Kujala, T., Alho, K., Paavilainen, P., Summala, H., & Naatanen, R. (1992). Neural plasticity in processing of sound location by the early blind: an event-related potential study. *Electroencephalogr Clin Neurophysiol, 84*(5), 469-472.

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279-292. doi:10.1016/j.neuron.2006.12.011

Landry, S. P., Bacon, B. A., Leybaert, J., Gagne, J. P., & Champoux, F. (2012). Audiovisual segregation in cochlear implant users. *PLoS One, 7*(3), e33113. doi:10.1371/journal.pone.0033113

Landry, S. P., Guillemot, J. P., & Champoux, F. (2013). Temporary deafness can impair multisensory integration: a study of cochlear-implant users. *Psychol Sci, 24*(7), 1260-1268. doi:10.1177/0956797612471142

Landry, S. P., Guillemot, J. P., & Champoux, F. (2014). Audiotactile interaction can change over time in cochlear implant users. *Front Hum Neurosci, 8*, 316. doi:10.3389/fnhum.2014.00316

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *J Opt Soc Am A Opt Image Sci Vis, 18*(9), 2307-2320.

Langer, M. S., & Bulthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception, 30*(4), 403-410.

Lazzouni, L., & Lepore, F. (2014). Compensatory plasticity: time matters. *Front Hum Neurosci, 8*, 340. doi:10.3389/fnhum.2014.00340

Leclerc, C., Saint-Amour, D., Lavoie, M. E., Lassonde, M., & Lepore, F. (2000). Brain functional reorganization in early blind humans revealed by auditory event-related potentials. *Neuroreport, 11*(3), 545-550.

Lee, D. S., Lee, J. S., Oh, S. H., Kim, S. K., Kim, J. W., Chung, J. K., . . . Kim, C. S. (2001). Cross-modal plasticity and cochlear implants. *Nature, 409*(6817), 149-150. doi:10.1038/35051653

Lessard, N., Pare, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature, 395*(6699), 278-280. doi:10.1038/26228

Lewald, J. (1998). The effect of gaze eccentricity on perceived sound direction and its relation to visual localization. *Hear Res, 115*(1-2), 206-216.

Lewald, J. (2002a). Rapid adaptation to auditory-visual spatial disparity. *Learn Mem, 9*(5), 268-278. doi:10.1101/lm.51402

Lewald, J. (2002b). Vertical sound localization in blind humans. *Neuropsychologia, 40*(12), 1868-1872.

Lewald, J. (2007). More accurate sound localization induced by short-term light deprivation. *Neuropsychologia, 45*(6), 1215-1222. doi:10.1016/j.neuropsychologia.2006.10.006

Lewald, J. (2013). Exceptional ability of blind humans to hear sound motion: implications for the emergence of auditory space. *Neuropsychologia, 51*(1), 181-186. doi:10.1016/j.neuropsychologia.2012.11.017

Lewald, J., Dorrscheidt, G. J., & Ehrenstein, W. H. (2000). Sound localization with eccentric head position. *Behav Brain Res, 108*(2), 105-125.

Lewald, J., & Ehrenstein, W. H. (1998). Auditory-visual spatial integration: a new psychophysical approach using laser pointing to acoustic targets. *J Acoust Soc Am, 104*(3 Pt 1), 1586-1597.

Lewald, J., & Ehrenstein, W. H. (2001). Spatial coordinates of human auditory working memory. *Brain Res Cogn Brain Res, 12*(1), 153-159.

Lewald, J., & Getzmann, S. (2006). Horizontal and vertical effects of eye-position on sound localization. *Hear Res, 213*(1-2), 99-106. doi:10.1016/j.heares.2006.01.001

Lewald, J., & Getzmann, S. (2013). Ventral and dorsal visual pathways support auditory motion processing in the blind: evidence from electrical neuroimaging. *Eur J Neurosci, 38*(8), 3201-3209. doi:10.1111/ejn.12306

Lewkowicz, D. J. (2002). Heterogeneity and heterochrony in the development of intersensory perception. *Brain Res Cogn Brain Res, 14*(1), 41-63.

Lickliter, R., & Bahrick, L. E. (2004). Perceptual Development and the Origins of Multisensory Responsiveness. In G. A. Calvert, C. Spence, & B. Stein (Eds.), *The Handbook of Multisensory Processes*. Cambridge, MA: MIT Press.

Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat Neurosci, 13*(11), 1421-1427. doi:10.1038/nn.2653

Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (2001). Navigating without vision: basic and applied research. *Optom Vis Sci, 78*(5), 282-289.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *J Exp Psychol Gen, 122*(1), 73-91.

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PLoS One, 8*(1), e54798. doi:10.1371/journal.pone.0054798

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat Neurosci, 9*(11), 1432-1438. doi:10.1038/nn1790

MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S., Williams, S. C., . . . Brammer, M. J. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain, 125*(Pt 7), 1583-1593.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus, 16*(12), 1091-1101. doi:10.1002/hipo.20233

Mahoney, J. R., Li, P. C., Oh-Park, M., Verghese, J., & Holtzer, R. (2011). Multisensory integration across the senses in young and old adults. *Brain Res, 1426*, 43-53. doi:10.1016/j.brainres.2011.09.017

Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *J Acoust Soc Am, 87*(5), 2188-2200.

Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J. P., Maeder, P. P., Clarke, S., & Meuli, R. A. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb Cortex, 17*(7), 1672-1679. doi:10.1093/cercor/bhl077

Maurer, D., Mondloch, C. J., & Lewis, T. L. (2007). Effects of early visual deprivation on perceptual and cognitive development. *Prog Brain Res, 164*, 87-104. doi:10.1016/s0079-6123(07)64005-9

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

Merabet, L., Hamilton, R., Schlaug, G., Swisher, J. D., Kiriakopoulos, E. T., Pitskel, N. B., . . . Pascual-Leone, A. (2008). Rapid and reversible recruitment of early visual cortex for touch. *PLoS One, 3*(8), e3046. doi:10.1371/journal.pone.0003046

Merabet, L., & Pascual-Leone, A. (2010). Neural reorganization following sensory loss: the opportunity of change. *Nat Rev Neurosci, 11*(1), 44-52.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci, 7*(10), 3215-3229.

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science, 221*(4608), 389-391.

Meredith, M. A., & Stein, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Res, 365*(2), 350-354.

Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *J Neurophysiol, 75*(5), 1843-1857.

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn Psychol, 14*(2), 247-279.

Mills, A. W. (1958). On the Minimum Audible Angle. *J Acoust Soc Am, 30*(4), 237-246. doi:10.1121/1.1909553

Mittelstaedt, M. L., & Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften, 67*(11), 566-567. doi:10.1007/BF00450672

Miyauchi, R., Kang, D. G., Iwaya, Y., & Suzuki, Y. (2014). Relative localization of auditory and visual events presented in peripheral visual field. *Multisens Res, 27*(1), 1-16.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res, 14*(1), 115-128.

Moody-Antonio, S., Takayanagi, S., Masuda, A., Auer, E. T., Jr., Fisher, L., & Bernstein, L. E. (2005). Improved speech perception in adult congenitally deafened cochlear implant recipients. *Otol Neurotol, 26*(4), 649-654.

Morgan, C. L. (1978). Constancy of egocentric visual direction. *Percept Psychophys, 23*(1), 61-68.

Moro, S. S., Harris, L. R., & Steeves, J. K. (2014). Optimal audiovisual integration in people with one eye. *Multisens Res, 27*(3-4), 173-188.

Moro, S. S., & Steeves, J. K. (2015). Audiovisual integration in people with one eye: Normal temporal binding window and sound induced flash illusion but reduced McGurk effect. *J Vis, 15*(12), 721. doi:10.1167/15.12.721

Morrell, F. (1972). Visual system's view of acoustic space. *Nature, 238*(5358), 44-46.

Morris, R. G. M. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation, 12*(2), 239-260. doi:10.1016/0023-9690(81)90020-5

Muchnik, C., Efrati, M., Nemeth, E., Malin, M., & Hildesheimer, M. (1991). Central auditory skills in blind and sighted subjects. *Scand Audiol, 20*(1), 19-23.

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2015). The multisensory function of the human primary visual cortex. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2015.08.011

Musseler, J., van der Heijden, A. H., Mahmud, S. H., Deubel, H., & Ertsey, S. (1999). Relative mislocalization of briefly presented stimuli in the retinal periphery. *Percept Psychophys, 61*(8), 1646-1661.

Nardini, M., Bedford, R., & Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. *Proc Natl Acad Sci U S A, 107*(39), 17041-17046. doi:10.1073/pnas.1001699107

Nardini, M., Begus, K., & Mareschal, D. (2013). Multisensory uncertainty reduction for hand localization in children and adults. *J Exp Psychol Hum Percept Perform, 39*(3), 773-787. doi:10.1037/a0030719

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Curr Biol, 18*(9), 689-693. doi:10.1016/j.cub.2008.04.021

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Dev Sci, 9*(5), 454-464. doi:10.1111/j.1467-7687.2006.00512.x

Neville, H. J., & Lawson, D. (1987). Attention to central and peripheral visual space in a movement detection task: an event-related potential and behavioral study. II. Congenitally deaf adults. *Brain Res, 405*(2), 268-283.

Nilsson, M. E., & Schenkman, B. N. (2015). Blind people are more sensitive than sighted people to binaural sound-location cues, particularly inter-aural level differences. *Hear Res*. doi:10.1016/j.heares.2015.09.012

Noppeney, U. (2007). The effects of visual deprivation on functional and structural organization of the human brain. *Neurosci Biobehav Rev, 31*(8), 1169-1180. doi:10.1016/j.neubiorev.2007.04.012

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map.* Oxford: Clarendon Press.

Occelli, V., Bruns, P., Zampini, M., & Roder, B. (2012). Audiotactile integration is reduced in congenital blindness in a spatial ventriloquism task. *Neuropsychologia, 50*(1), 36-43. doi:10.1016/j.neuropsychologia.2011.10.019

Occelli, V., Spence, C., & Zampini, M. (2013). Auditory, tactile, and audiotactile information processing following visual deprivation. *Psychol Bull, 139*(1), 189-212. doi:10.1037/a0028416

Odegaard, B., Wozny, D. R., & Shams, L. (2015). Biases in Visual, Auditory, and Audiovisual Perception of Space. *PLoS Comput Biol, 11*(12), e1004649. doi:10.1371/journal.pcbi.1004649

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration. *Nat Neurosci, 14*(6), 775-782. doi:10.1038/nn.2815

Ortiz, T., Poch, J., Santos, J. M., Requena, C., Martinez, A. M., Ortiz-Teran, L., . . . Pascual-Leone, A. (2011). Recruitment of occipital cortex during sensory substitution training linked to subjective experience of seeing in people with blindness. *PLoS One, 6*(8), e23264. doi:10.1371/journal.pone.0023264

Oruc, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Res, 43*(23), 2451-2468.

Ostrovsky, Y., Andalman, A., & Sinha, P. (2006). Vision following extended congenital blindness. *Psychol Sci, 17*(12), 1009-1014. doi:10.1111/j.1467-9280.2006.01827.x

Otto, T. U., & Mamassian, P. (2012). Noise and correlations in parallel perceptual decision making. *Curr Biol, 22*(15), 1391-1396. doi:10.1016/j.cub.2012.05.031

Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proc Natl Acad Sci U S A, 111*(16), 6104-6108. doi:10.1073/pnas.1322705111

Pavani, F., & Bottari, D. (2012). Frontiers in Neuroscience: Visual Abilities in Individuals with Profound Deafness A Critical Review. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press Llc.

Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I. P., Mottonen, R., Tarkiainen, A., & Sams, M. (2005). Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroreport, 16*(2), 125-128.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis, 10*(4), 437-442.

Perez Fornos, A., Sommerhalder, J., da Cruz, L., Sahel, J. A., Mohand-Said, S., Hafezi, F., & Pelizzone, M. (2012). Temporal properties of visual perception on electrical stimulation of the retina. *Invest Ophthalmol Vis Sci, 53*(6), 2720-2731. doi:10.1167/iovs.11-9344

Perrott, D. R. (1984). Concurrent minimum audible angle: A re‐examination of the concept of auditory spatial acuity. *J Acoust Soc Am, 75*(4), 1201-1206. doi:10.1121/1.390771

Perrott, D. R., Costantino, B., & Cisneros, J. (1993). Auditory and visual localization performance in a sequential discrimination task. *J Acoust Soc Am, 93*(4), 2134-2138. doi:10.1121/1.406675

Petrini, K., Caradonna, A., Foster, C., Burgess, N., & Nardini, M. (2014). The visual influence on path reproduction in darkness is stronger during childhood. *J Vis, 14*(10), 1345. doi:10.1167/14.10.1345

Petrini, K., Jones, P. R., Smith, L., & Nardini, M. (2015). Hearing Where the Eyes See: Children Use an Irrelevant Visual Cue When Localizing Sounds. *Child Dev, 86*(5), 1449-1457. doi:10.1111/cdev.12397

Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option: children's integration of auditory and haptic information is suboptimal. *Dev Sci, 17*(3), 376-387. doi:10.1111/desc.12127

Petrus, E., Isaiah, A., Jones, A. P., Li, D., Wang, H., Lee, H. K., & Kanold, P. O. (2014). Crossmodal induction of thalamocortical potentiation leads to enhanced information processing in the auditory cortex. *Neuron, 81*(3), 664-673. doi:10.1016/j.neuron.2013.11.023

Piaget, J., & Cook, M. T. (1952). *The origins of intelligence in children.*: NY: International University Press.

Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., . . . De Volder, A. G. (2006). Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage, 31*(1), 279-285. doi:10.1016/j.neuroimage.2005.11.036

Poirier, C., De Volder, A. G., & Scheiber, C. (2007). What neuroimaging tells us about sensory substitution. *Neurosci Biobehav Rev, 31*(7), 1064-1070. doi:10.1016/j.neubiorev.2007.05.010

Proksch, J., & Bavelier, D. (2002). Changes in the spatial distribution of visual attention after early deafness. *J Cogn Neurosci, 14*(5), 687-701. doi:10.1162/08989290260138591

Putzar, L., Goerendt, I., Lange, K., Rosler, F., & Roder, B. (2007). Early visual deprivation impairs multisensory interactions in humans. *Nat Neurosci, 10*(10), 1243-1245. doi:10.1038/nn1978

Putzar, L., Gondan, M., & Roder, B. (2012). Basic multisensory functions can be acquired after congenital visual pattern deprivation in humans. *Dev Neuropsychol, 37*(8), 697-711. doi:10.1080/87565641.2012.696756

Putzar, L., Hotting, K., & Roder, B. (2010). Early visual deprivation affects the development of face recognition and of audio-visual speech perception. *Restor Neurol Neurosci, 28*(2), 251-257. doi:10.3233/rnn-2010-0526

Putzar, L., Hotting, K., Rosler, F., & Roder, B. (2007). The development of visual feature binding processes after visual deprivation in early infancy. *Vision Res, 47*(20), 2616-2626. doi:10.1016/j.visres.2007.07.002

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Trans N Y Acad Sci, 24*, 574-590.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychol Rev, 111*(2), 333-367. doi:10.1037/0033-295x.111.2.333

Rauschecker, J. P., & Harris, L. R. (1983). Auditory compensation of the effects of visual deprivation in the cat's superior colliculus. *Exp Brain Res, 50*(1), 69-83.

Rauschecker, J. P., & Kniepert, U. (1994). Auditory localization behaviour in visually deprived cats. *Eur J Neurosci, 6*(1), 149-160.

Razavi, B., O'Neill, W. E., & Paige, G. D. (2007). Auditory spatial perception dynamically realigns with changing eye position. *J Neurosci, 27*(38), 10249-10258. doi:10.1523/jneurosci.0938-07.2007

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc Natl Acad Sci U S A, 95*(3), 869-875.

Renier, L. A., Anurova, I., De Volder, A. G., Carlson, S., VanMeter, J., & Rauschecker, J. P. (2010). Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind. *Neuron, 68*(1), 138-148. doi:10.1016/j.neuron.2010.09.021

Renzi, C., Bruns, P., Heise, K. F., Zimerman, M., Feldheim, J. F., Hummel, F. C., & Roder, B. (2013). Spatial remapping in the audio-tactile ventriloquism effect: a TMS investigation on the role of the ventral intraparietal area. *J Cogn Neurosci, 25*(5), 790-801. doi:10.1162/jocn_a_00362

Ricciardi, E., Bonino, D., Pellegrini, S., & Pietrini, P. (2013). Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture? *Neurosci Biobehav Rev*. doi:10.1016/j.neubiorev.2013.10.006

Roder, B., Teder-Salejarvi, W., Sterr, A., Rosler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature, 400*(6740), 162-166. doi:10.1038/22106

Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biol, 13*(2), e1002073. doi:10.1371/journal.pbio.1002073

Romei, V., Murray, M. M., Merabet, L., & Thut, G. (2007). Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *J Neurosci, 27*(43), 11465-11472. doi:10.1523/jneurosci.2827-07.2007

Rosas, P., Wagemans, J., Ernst, M. O., & Wichmann, F. A. (2005). Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *J Opt Soc Am A Opt Image Sci Vis, 22*(5), 801-809.

Rouger, J., Fraysse, B., Deguine, O., & Barone, P. (2008). McGurk effects in cochlear-implanted deaf subjects. *Brain Res, 1188*, 87-99. doi:10.1016/j.brainres.2007.10.049

Rousseeuw, P. J., & Driessen, K. v. (1999). A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics, 41*(3), 212-223. doi:10.2307/1270566

Rowland, B., Jiang, W., & Stein, B. E. (2014). Brief cortical deactivation early in life has long-lasting effects on multisensory behavior. *J Neurosci, 34*(21), 7198-7202. doi:10.1523/jneurosci.3782-13.2014

Rowland, B., Perrault, T., Vaughan, J., & Stein, B. (2015). Multisensory Integration is based on Information, not Efficacy. *J Vis, 15*(12), 363. doi:10.1167/15.12.363

Royal, D. W., Krueger, J., Fister, M. C., & Wallace, M. T. (2010). Adult plasticity of spatiotemporal receptive fields of multisensory superior colliculus neurons following early visual deprivation. *Restor Neurol Neurosci, 28*(2), 259-270. doi:10.3233/rnn-2010-0488

Sadato, N., Okada, T., Honda, M., & Yonekura, Y. (2002). Critical period for cross-modal plasticity in blind humans: a functional MRI study. *Neuroimage, 16*(2), 389-400. doi:10.1006/nimg.2002.1111

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M. P., Dold, G., & Hallett, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature, 380*(6574), 526-528. doi:10.1038/380526a0

Sathian, K. (2005). Visual cortical activity during tactile perception in the sighted and the visually deprived. *Dev Psychobiol, 46*(3), 279-286. doi:10.1002/dev.20056

Sathian, K., & Zangaladze, A. (2002). Feeling with the mind's eye: contribution of visual cortex to tactile perception. *Behav Brain Res, 135*(1-2), 127-132.

Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2003). Sound induces perceptual reorganization of an ambiguous motion display in human infants. *Dev Sci, 6*(3), 233-241. doi:10.1111/1467-7687.00276

Schorr, E. A., Fox, N. A., van Wassenhove, V., & Knudsen, E. I. (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proc Natl Acad Sci U S A, 102*(51), 18748-18750. doi:10.1073/pnas.0508862102

Schroeder, C. E., & Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res, 14*(1), 187-198.

Schroger, E., & Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology, 35*(6), 755-759.

Schurmann, M., Caetano, G., Hlushchuk, Y., Jousmaki, V., & Hari, R. (2006). Touch activates human auditory cortex. *Neuroimage, 30*(4), 1325-1331. doi:10.1016/j.neuroimage.2005.11.020

Sciutti, A., Burr, D., Saracco, A., Sandini, G., & Gori, M. (2014). Development of context dependency in human space perception. *Exp Brain Res, 232*(12), 3965-3976. doi:10.1007/s00221-014-4021-y

Scott, G. D., Karns, C. M., Dow, M. W., Stevens, C., & Neville, H. J. (2014). Enhanced peripheral visual processing in congenitally deaf humans is supported by multiple brain regions, including primary auditory cortex. *Front Hum Neurosci, 8*, 177. doi:10.3389/fnhum.2014.00177

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature, 385*(6614), 308. doi:10.1038/385308a0

Senna, I., Parise, C. V., & Ernst, M. O. (2015). Hearing in slow-motion: Humans underestimate the speed of moving sounds. *Sci Rep, 5*, 14054. doi:10.1038/srep14054

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature, 408*(6814), 788. doi:10.1038/35048669

Sheth, B. R., & Shimojo, S. (2001). Compression of space in visual memory. *Vision Res, 41*(3), 329-341.

Shinn-Cunningham, B. (2000). Adapting to remapped auditory localization cues: a decision-theory model. *Percept Psychophys, 62*(1), 33-47.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci, 27*(3), 161-168. doi:10.1016/j.tins.2004.01.006

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Ann N Y Acad Sci, 1296*, 31-49. doi:10.1111/nyas.12121

Spinelli, D. N., Starr, A., & Barrett, T. W. (1968). Auditory specificity in unit recordings from cat's visual cortex. *Exp Neurol, 22*(1), 75-84.

Stein, B. E., Labos, E., & Kruger, L. (1973). Sequence of changes in properties of neurons of superior colliculus of the kitten during maturation. *J Neurophysiol, 36*(4), 667-679.

Stein, B. E., Meredith, M. A., Huneycutt, W. S., & McDade, L. (1989). Behavioral Indices of Multisensory Integration: Orientation to Visual Cues is Affected by Auditory Stimuli. *J Cogn Neurosci, 1*(1), 12-24. doi:10.1162/jocn.1989.1.1.12

Stevens, A. A., & Weaver, K. E. (2005). Auditory perceptual consolidation in early-onset blindness. *Neuropsychologia, 43*(13), 1901-1910. doi:10.1016/j.neuropsychologia.2005.03.007

Stevens, J. C., Foulke, E., & Patterson, M. Q. (1996). Tactile acuity, aging, and braille reading in long-term blindness. *Journal of experimental psychology: applied, 2*(2), 91.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nat Neurosci, 9*(4), 578-585. doi:10.1038/nn1669

Striem-Amit, E., Dakwar, O., Reich, L., & Amedi, A. (2012). The large-scale organization of "visual" streams emerges without visual experience. *Cereb Cortex, 22*(7), 1698-1709. doi:10.1093/cercor/bhr253

Stronks, H. C., & Dagnelie, G. (2014). The functional performance of the Argus II retinal prosthesis. *Expert review of medical devices, 11*(1), 23-30. doi:10.1586/17434440.2014.862494

Tao, Q., Chan, C. C., Luo, Y. J., Li, J. J., Ting, K. H., Wang, J., & Lee, T. M. (2013). How Does Experience Modulate Auditory Spatial Processing in Individuals with Blindness? *Brain Topogr*. doi:10.1007/s10548-013-0339-1

Targher, S., Occelli, V., & Zampini, M. (2012). Audiovisual integration in low vision individuals. *Neuropsychologia, 50*(5), 576-582. doi:10.1016/j.neuropsychologia.2011.10.021

Tcheang, L., Bulthoff, H. H., & Burgess, N. (2011). Visual influence on path integration in darkness indicates a multimodal representation of large-scale space. *Proc Natl Acad Sci U S A, 108*(3), 1152-1157. doi:10.1073/pnas.1011843108

Thaler, L., Arnott, S. R., & Goodale, M. A. (2011). Neural correlates of natural human echolocation in early and late blind echolocation experts. *PLoS One, 6*(5), e20162. doi:10.1371/journal.pone.0020162

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Res, 22*(3), 377-380.

Tremblay, C., Champoux, F., Lepore, F., & Theoret, H. (2010). Audiovisual fusion and cochlear implant proficiency. *Restor Neurol Neurosci, 28*(2), 283-291. doi:10.3233/rnn-2010-0498

Tyler, R., Parkinson, A. J., Fryauf-Bertchy, H., Lowder, M. W., Parkinson, W. S., Gantz, B. J., & Kelsay, D. M. (1997). Speech perception by prelingually deaf children and postlingually deaf adults with cochlear implant. *Scand Audiol Suppl, 46*, 65-71.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In M. A. G. D. J. Ingle, & R. J. W. Mansfield (Ed.), *Analysis of visual behaviour.* Cambridge, MA: MIT Press.

Vachon, P., Voss, P., Lassonde, M., Leroux, J. M., Mensour, B., Beaudoin, G., . . . Lepore, F. (2013). Reorganization of the auditory, visual and multimodal areas in early deaf individuals. *Neuroscience, 245*, 50-60. doi:10.1016/j.neuroscience.2013.04.004

Van Boven, R. W., Hamilton, R. H., Kauffman, T., Keenan, J. P., & Pascual-Leone, A. (2000). Tactile spatial resolution in blind braille readers. *Neurology, 54*(12), 2230-2236.

van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. *Percept Psychophys, 70*(6), 955-968.

Verboven, S., & Hubert, M. (2005). LIBRA: a MATLAB library for robust analysis. *Chemometrics and Intelligent Laboratory Systems, 75*(2), 127-136. doi:10.1016/j.chemolab.2004.06.003

Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Curr Biol, 24*(11), 1256-1262. doi:10.1016/j.cub.2014.04.020

von Helmholtz, H. (1993). *Treatise on Physiological Optics: Translated from the 3rd German edition by Southall JPC* (Vol. 3). WI: George Banta Publishing.

Voss, P., Gougoux, F., Zatorre, R. J., Lassonde, M., & Lepore, F. (2008). Differential occipital responses in early- and late-blind individuals during a sound-source discrimination task. *Neuroimage, 40*(2), 746-758. doi:10.1016/j.neuroimage.2007.12.020

Voss, P., Lassonde, M., Gougoux, F., Fortin, M., Guillemot, J. P., & Lepore, F. (2004). Early- and late-onset blind individuals show supra-normal auditory abilities in far-space. *Curr Biol, 14*(19), 1734-1738. doi:10.1016/j.cub.2004.09.051

Voss, P., Pike, B. G., & Zatorre, R. J. (2014). Evidence for both compensatory plastic and disuse atrophy-related neuroanatomical changes in the blind. *Brain, 137*(Pt 4), 1224-1240. doi:10.1093/brain/awu030

Voss, P., & Zatorre, R. J. (2012). Organization and reorganization of sensory-deprived cortex. *Curr Biol, 22*(5), R168-173. doi:10.1016/j.cub.2012.01.030

Vroomen, J., de Gelder, B., & Vroomen, J. (2004). Temporal ventriloquism: sound modulates the flash-lag effect. *J Exp Psychol Hum Percept Perform, 30*(3), 513-518. doi:10.1037/0096-1523.30.3.513

Wallace, M. T., Meredith, M. A., & Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *J Neurophysiol, 80*(2), 1006-1010.

Wallace, M. T., Perrault, T. J., Jr., Hairston, W. D., & Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *J Neurosci, 24*(43), 9580-9584. doi:10.1523/jneurosci.2535-04.2004

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Exp Brain Res, 158*(2), 252-258. doi:10.1007/s00221-004-1899-9

Wallace, M. T., & Stein, B. E. (1997). Development of multisensory neurons and multisensory integration in cat superior colliculus. *J Neurosci, 17*(7), 2429-2444.

Wallace, M. T., & Stein, B. E. (2001). Sensory and multisensory responses in the newborn monkey superior colliculus. *J Neurosci, 21*(22), 8886-8894.

Wallace, M. T., & Stein, B. E. (2007). Early experience determines how the senses will interact. *J Neurophysiol, 97*(1), 921-926. doi:10.1152/jn.00497.2006

Wallace, M. T., Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *J Neurophysiol, 76*(2), 1246-1266.

Wan, C. Y., Wood, A. G., Reutens, D. C., & Wilson, S. J. (2010). Congenital blindness leads to enhanced vibrotactile perception. *Neuropsychologia, 48*(2), 631-635. doi:10.1016/j.neuropsychologia.2009.10.001

Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage, 31*(3), 1247-1256. doi:10.1016/j.neuroimage.2006.01.016

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *J Vis, 5*(10), 834-862. doi:10.1167/5.10.7

Weaver, K. E., & Stevens, A. A. (2007). Attention and sensory interactions within the occipital cortex in the early blind: an fMRI study. *J Cogn Neurosci, 19*(2), 315-330. doi:10.1162/jocn.2007.19.2.315

Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, C. M., Cohen, L. G., . . . Rauschecker, J. P. (2000). A positron emission tomographic study of auditory localization in the congenitally blind. *J Neurosci, 20*(7), 2664-2672.

Welchman, A. E., Lam, J. M., & Bulthoff, H. H. (2008). Bayesian motion estimation accounts for a surprising bias in 3D vision. *Proc Natl Acad Sci U S A, 105*(33), 12087-12092. doi:10.1073/pnas.0804378105

Withington-Wray, D. J., Binns, K. E., & Keating, M. J. (1990). The Maturation of the Superior Collicular Map of Auditory Space in the Guinea Pig is Disrupted by Developmental Visual Deprivation. *Eur J Neurosci, 2*(8), 682-692.

Wong, M., Gnanakumaran, V., & Goldreich, D. (2011). Tactile spatial acuity enhancement in blindness: evidence for experience-dependent mechanisms. *J Neurosci, 31*(19), 7028-7037. doi:10.1523/jneurosci.6461-10.2011

Xu, J., Yu, L., Stanford, T. R., Rowland, B. A., & Stein, B. E. (2015). What does a neuron learn from multisensory experience? *J Neurophysiol, 113*(3), 883-889. doi:10.1152/jn.00284.2014

Yu, L., Rowland, B. A., & Stein, B. E. (2010). Initiating the development of multisensory integration by manipulating sensory experience. *J Neurosci, 30*(14), 4904-4913. doi:10.1523/JNEUROSCI.5575-09.2010

Yu, L., Stein, B. E., & Rowland, B. A. (2009). Adult plasticity in multisensory neurons: short-term experience-dependent changes in the superior colliculus. *J Neurosci, 29*(50), 15910-15922. doi:10.1523/jneurosci.4041-09.2009

Zaidel, A., Ma, W. J., & Angelaki, D. E. (2013). Supervised calibration relies on the multisensory percept. *Neuron, 80*(6), 1544-1557. doi:10.1016/j.neuron.2013.09.026

Zaidel, A., Turner, A. H., & Angelaki, D. E. (2011). Multisensory calibration is independent of cue reliability. *J Neurosci, 31*(39), 13949-13962. doi:10.1523/jneurosci.2732-11.2011

Zangaladze, A., Epstein, C. M., Grafton, S. T., & Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature, 401*(6753), 587-590. doi:10.1038/44139

Zwiers, M. P., Van Opstal, A. J., & Cruysberg, J. R. (2001). A spatial hearing deficit in early-blind humans. *J Neurosci, 21*(9), RC142: 141-145.