The University of
Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Holmes, Nicholas P. and Dakwar, Azar R. (2015) Online control of reaching and pointing to visual, auditory, and multimodal targets: effects of target modality and method of determining correction latency. Vision Research, 117 . pp. 105-116. ISSN 1878-5646

Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/31072/3/Holmes_Dakwar_R3_no_markup2.pdf

1 **Title:** Online control of reaching and pointing to visual, auditory, and multimodal targets:

2 Effects of target modality and method of determining correction latency

3

4 **Authors:** Nicholas P. Holmes[a,b,c], Azar R. Dakwar[a,b]

5

6 **Affiliations:**

7 a.   Department of Psychology, The Hebrew University of Jerusalem, Israel

8 b.   Edmond and Lily Safra Center for Brain Sciences and Interdisciplinary Center for

9 Neural Computation, The Hebrew University of Jerusalem, Israel

10 c.   School of Psychology, University of Nottingham, UK

11

12 **Corresponding author:** Nicholas P Holmes

13 School of Psychology, University of Nottingham

14 University Park, Nottingham, UK, RG7 2RD

15 0044 (0)115 951 2853

16 npholmes@neurobiography.info

17

18 **Running head:** Multimodal online control

Abstract

Movements aimed towards objects occasionally have to be adjusted when the object moves. These online adjustments can be very rapid, occurring in as little as 100ms. More is known about the latency and neural basis of online control of movements to visual than to auditory target objects. We examined the latency of online corrections in reaching-to-point movements to visual and auditory targets that could change side and/or modality at movement onset. Visual or auditory targets were presented on the left or right sides, and participants were instructed to reach and point to them as quickly and as accurately as possible. On half of the trials, the targets changed side at movement onset, and participants had to correct their movements to point to the new target location as quickly as possible. Given different published approaches to measuring the latency for initiating movement corrections, we examined several different methods systematically. What we describe here as the optimal methods involved fitting a straight-line model to the velocity of the correction movement, rather than using a statistical criterion to determine correction onset. In the multimodal experiment, these model-fitting methods produced significantly lower latencies for correcting movements away from the auditory targets than away from the visual targets. Our results confirm that rapid online correction is possible for auditory targets, but further work is required to determine whether the underlying control system for reaching and pointing movements is the same for auditory and visual targets.

Keywords: Multisensory, multimodal, space, online control, methods

1 Introduction

When reaching to point towards or grasp an object, it occasionally moves unexpectedly, or we dislodge it with our hand, or our initial movement was inaccurate. We then have to correct our movement 'online' during its execution. Online control may be the default mode of visuo-motor control, rather than using a model-based or predictive form of control (Zhao & Warren, 2015). Online movement corrections can be very rapid. In cats reaching for a food reward, paw movements can be corrected within as little as 60-70ms following changes in target location (Alstermark, Eide, Górska, Lundberg, & Pettersson, 1984). In humans, significant changes in reaching movement acceleration have been reported as early as 90ms after the target displacement (Paulignan, MacKenzie, Marteniuk, & Jeannerod, 1991). The online control of movements has been thoroughly investigated for changes in the location, size, and other features of visual targets (Paulignan et al., 1991a; Paulignan, Jeannerod, MacKenzie, & Marteniuk, 1991; Veerman, Brenner, & Smeets, 2008; Wijdenes, Gomi, and Brenner, 2015), but the online control of movements towards auditory targets has only just begun to be studied (Boyer, Babayan, Bevilacqua, Noisternig, Warusfel, Roby-Brami, Hanneton, & Vivaud-Delmon, 2013; see Cameron and López-Molinar, 2015, Cluff, Crevecoeur, & Scott, 2015, for similar points regarding proprioception). The present study investigated the ability of healthy human participants to make online movement corrections to visual, auditory, and multimodal targets. In particular, we compared the latencies of these corrections. By 'multimodal' target, we mean a target that begins either as visual or auditory, then switches modality after movement onset, to become auditory or visual, respectively.

The online control of movements to visual targets is thought to be a function of the dorsal visual stream: damage to the superior occipital-parietal cortex impairs the online control of reaching movements (Pisella, Gréa, Tilikete, Vighetto, Desmurget, Rode, Boisson, & Rossetti, 2000), and targets thought to be processed most rapidly by the magnocellular pathway of the dorsal stream are associated with lower latency online control (Veerman et

3

66   al., 2008). There is little evidence concerning the neural basis of online control of movements

67   towards auditory targets. In macaques, neurons in the parietal and premotor cortices may

68   represent the locations of targets across modalities in a common reference frame, for eye

69   and hand movements (Cohen & Anderson, 2000; Graziano, Reiss, & Gross, 1999). Further,

70   the superior colliculus, which receives inputs from vision, audition, and somatosensation, as

71   well as other brain stem regions has been implicated in the online control of reaching

72   movements in the cat (Alstermark et al., 1984; Pettersson, Lundberg, Alstermark, Isa, &

73   Tantisira, 1997), and in primates (Song, Rafal, & McPeek, 2011; Werner, 1993).

74   Given that these brain areas thought to be involved in the online control of movement

75   are responsive to multiple sensory modalities, we speculated that some aspects of the online

76   control of movements may be multimodal or supramodal in nature, and, further, that rapid

77   online control may even be possible for targets that change modality as well as location.

78   Changes in target modality such as this might occur in nature, for example with a cat chasing

79   a mouse (Alstermark et al., 1984): As the mouse runs behind an object, it is visually occluded

80   from the cat, but auditory cues may still be available to guide pursuit.

81   We asked healthy volunteers to make speeded reaching and pointing movements to

82   visual (Experiment 1) and auditory (Experiment 2) targets, which changed location on 50% of

83   the trials (from left-to-right or right-to-left), and, in the third experiment, orthogonally could

84   also change modality (from auditory-to-visual or visual-to-auditory) after movement onset.

85   We determined the time-point at which the movement trajectory changed in the different

86   conditions. Following reviewers' comments, we systematically investigated two different

87   methods of determining latency (statistical, and extrapolation), for three different levels of

88   analysis (whole group, individual participant, and individual trial), and three different types of

89   velocity (lateral, resultant, and statistical components of velocity) – 18 different combinations.

90   For the statistical methods, 61 different statistical thresholds were assessed. This systematic

91   investigation allowed greater certainty in our conclusions, but also highlighted large

92   differences between different methods of estimating correction latency from velocity data.

4

93　　　To summarise, we aimed first to compare different methods of measuring correction

94　latency (see also Wijdenes, Brenner, and Smeets, 2014), second to examine the latency of

95　online corrections for pointing to auditory targets in comparison with visual targets, and third

96　to examine the latencies of movement corrections made to both visual and auditory targets

97　that can change modality and/or position at movement onset.

98

99　2 Materials and Methods

100　2.1 Participants

101　　　Thirteen participants (7 male, 6 female; 11 right-handed; aged between 20 and 33

102　years; including two of the authors) took part in the experiments. All of the participants had

103　normal or corrected vision. All participants gave written, informed consent, the experimental

104　procedures were approved by the local ethical review panel at the Hebrew University of

105　Jerusalem, and were in accordance with the Declaration of Helsinki (as of 2008).

106

107　2.2 Apparatus and materials

108　　　The experiments were performed in a darkened sound-attenuated chamber (Eckel C-

109　26, UK). Participants sat in the middle of the chamber on a straight-backed chair with a

110　horizontal board as a forearm rest supporting a small marker for the starting position in the

111　centre of the chamber (Figure 1). An arced metal hoop of 90cm radius supported an array of

112　three loudspeakers (7.5 degrees left, centrally, and 7.5 degrees right of the midline), and

113　three 5mm diameter LEDs (left, centre, and right, attached centrally in front of each

114　loudspeaker).

115　　　Index fingertip and head position (3 degrees of freedom) and orientation (3 degrees of

116　freedom) were recorded with a Polhemus Patriot (Polhemus, Colchester, VT, USA) magnetic

117　tracking system, sampling at 60Hz. The transmitter was positioned centrally, in front of and

118　below the participant, between their knees. Participants wore plastic goggles which held the

119　head position tracker and a laser pointer, used to assist calibration of head position prior to

5

120 data collection. Horizontal and vertical electrooculographic (EOG) data were acquired with an

121 Active 2 Biosemi system (Biosemi, Amsterdam, The Netherlands), sampling at 1024Hz, with

122 an online low-pass filter of 256Hz. Four electrodes were used for EOG recording: two

123 electrodes for the horizontal EOG, at the outer canthi of the left and right eyes (HEOGL,

124 HEOGR), and two vertical EOG electrodes below (infraorbital, VEOGI) and above

125 (supraorbital, VEOGS) the right eye. Two channels were recorded from the mastoid

126 processes and another from the tip of the nose, but were not used. Bipolar EOG channels

127 were created offline by subtracting HEOGL from HEOGR and VEOGI from VEOGS. The

128 data were referenced online to a common-mode.

129

130 **2.3 Stimuli**

131   Visual and auditory stimuli were presented by passing the same amplitude- modulated

132 white noise stimulus waveform through the sound card of a PC. A parallel port signal

133 triggered a relay switch box that channelled the stimulus to either a loudspeaker or an LED

134 (5mm, red, ~800 mcd). The stimulus was generated on each trial as follows: A 1250ms white

135 noise signal sampled at 44,100Hz, ranging from -1 to +1 was attenuated by 5% to prevent

136 clipping, and shaped with a trapezoidal envelope providing 10ms rise and fall times. To

137 facilitate the perceptual localization of the auditory stimuli, the stimulus was multiplied by a

138 sinusoidal envelope with a frequency of 60Hz, providing an amplitude-modulation depth of

139 80%. Thus, perceptually, the auditory stimulus fluttered, while the visual stimulus was only

140 just visibly flickering if viewed peripherally. In pilot work, participants could easily discriminate

141 between auditory stimuli as little as 7.5 degrees apart (the closest that the speakers could be

142 put together without touching). The 95% confidence intervals for directional pointing error

143 were approximately 5 degrees in azimuth and elevation. Auditory stimuli were presented at a

144 mean approximately 75dB sound pressure level, but to reduce the possibility that auditory

145 stimulus intensity acted as a cue to distance or direction, stimulus amplitude was varied

146 randomly on each trial with a rectangular distribution between 75% and 100% of the original

6

147 amplitude. A central visual fixation stimulus (5mm diameter LED), remained on until the hand

148 and head were correctly positioned at the start of each trial. An auditory preparatory cue

149 (1000ms, 960Hz, central loudspeaker) was used to signal the start of each trial and to cue

150 the participants to fixate. Trials ended with the illumination of the central LED, and

151 presentation of a 1000ms 480Hz tone from the central speaker.

152

153 **2.4 Design**

154 Three experiments were performed in a pseudo-randomised fashion. The first

155 contained only visual targets (96 trials in total), the second only auditory targets (96 trials),

156 and the third (the multimodal experiment) contained visual, auditory, visual-then-auditory,

157 and auditory-then-visual targets (192 trials). The three experiments were performed on the

158 same day in a single session lasting 2-3 hours. The multimodal experiment was always run

159 either first (6 participants) or last (6 participants), while the order of the visual only and

160 auditory only experiments was counterbalanced across participants. The first participant did

161 not participate in the multimodal experiment because of technical problems.

162 Within the visual only and auditory only experiments, four conditions were generated by

163 the exhaustive combination of two binary variables: Initial target position (left or right), and

164 final target position (left or right). Thus, in half of the conditions, the target remained on one

165 side, and in the other half, it switched from one side to the other. The multimodal experiment

166 contained 16 conditions, formed from the exhaustive combination of four binary variables:

167 Initial target position (left or right), final target position (left or right), initial target modality

168 (visual or auditory), and final target modality (visual or auditory). Thus, in half of the trials,

169 there was a change in target position, and orthogonally in half of trials there was a change in

170 target modality at movement onset.

171

172 **2.5 Procedure**

7

173   Participants were briefed and gave informed consent, then the EOG electrodes were

174   attached. The participants were seated in the chamber and the tracking devices were

175   attached. 12-24 practice trials were performed before data collection in each experiment to

176   familiarise the participants with the task. The chamber was kept nearly dark, but due to

177   ambient light and the LEDs, the participants received rudimentary visual feedback of their

178   hand and arm positions during the experiment. To prevent complete dark adaptation, the

179   chamber door was opened every 10-15 minutes, between blocks of trials.

180   The participants were instructed to reach and point, as quickly and as accurately as

181   possible with the index finger of their dominant hand towards a target stimulus appearing on

182   the left or right of the midline, while maintaining their head and eye position (Cameron,

183   Cheng, Chua, van Donkelaar, & Binsted, 2013) towards the central fixation point (which was

184   visible only prior to target onset). Participants were instructed to minimise both their reaction

185   times (RT) and movement times (MT). The participants were told that, in half of the trials, the

186   target would switch, either from left to right, or from right to left, and their task was to correct

187   their movement as rapidly as possible and point to the new target location if it switched. In

188   the multimodal conditions, participants were informed that in half of the trials the target would

189   change from visual to auditory or from auditory to visual, either on the same or a different

190   side. Participants were instructed to keep their index finger as still as possible at the end of

191   their pointing movement, until the end cue sounded.

192   Each trial began with a computer check of the position of the participant's index finger

193   and the orientation of their head, while the visual fixation stimulus was illuminated. If there

194   was any substantial deviation from the starting position for the finger (more than 2cm from

195   the start), or the straight-ahead orientation for the head (more than 7.5 degrees from straight

196   ahead), a warning tone was sounded. If the participant did not correct their finger or head

197   positions, the experimenter gave verbal prompts over an intercom. When both finger and

198   head were correctly positioned, the visual fixation was extinguished and there was a random

199   pre-trial interval with a uniform rectangular distribution between 1.5 and 3s. After the pre-trial

8

interval, the target stimulus was presented, and the participant made their movement. The

index finger position was analysed online, and in trials with a change in target location, the

change occurred as soon as the index finger's 3D velocity exceeded 10cm/s. This velocity

criterion was also used as the 'reaction time', both for trials with and without a target change.

Movement endpoint was the first sample with a 3D velocity below 5cm/s which was

maintained for at least 50ms. Participants were not able to touch the targets with their arm

outstretched. Changes in target location were achieved via a parallel port switch. The

stimulus waveform output was directed either to an LED or a speaker, depending on which

parallel port pin was active. Two seconds of position data were recorded, the end cue was

presented, and the visual fixation was re-illuminated before the start of the next trial.

2.5.1 Calibration of body position, head orientation and EOG data

Before the experiment, the position of the following body parts of each participant was

measured in the recording chamber: left and right index fingers, wrist, elbow, shoulder, and

neck, the top of the head (vertex) and between the two eyes. After the experiment, to

calibrate the head orientation with respect to the speaker locations, each participant was

asked to orient their head (using a laser pointer) to 7 locations, from -15 to +15 degrees right

of centre in 5 degree steps. A series of 7 tones was presented to cue the participants to

move, then fixate each location with the laser pointer. These data were used to calibrate the

raw head orientation data when computing gaze.

A similar procedure was used for EOG calibration. Guided by a series of LED

illuminations, participants made a series of 12 saccadic eye movements while keeping their

head oriented centrally. Each eye movement began at the central location, then fixated an

LED target at either -15, -10, -5, 5, 10, and 15 degrees right of centre for 1s, then returned to

the central location. These 12 saccades were used to calibrate the EOG data by regressing

the saccade-related change in mV EOG signal on the instructed change in eye position. The

slope of this regression was used to estimate eye position changes in degrees from the raw

9

227 EOG data. Supplementary Figure 4 shows mean horizontal and vertical EOG and gaze data

228 for the unimodal auditory experiment. All other experiments and conditions produced similar

229 EOG data.

230

231 **2.6 Data analysis**

232 *2.6.1 Preprocessing*

233   The experiments were run and the data analysis was performed using Matlab

234 (Mathworks, Natick, USA). All the programs are available from the first author's website

235 (http://neurobiography.info), and all raw data will be freely available there. Magnetic

236 interference from the loudspeakers and the metal hoop inside the chamber warped the

237 kinematic data. The warping of the measured space in the chamber was most severe at the

238 periphery, near the loudspeakers and furthest from the transmitter. Warping within the region

239 of most interest - the space traversed in the first few hundred milliseconds of movement

240 duration - was minimal. The kinematic data were unwarped during acquisition using a set of

241 multiple non-linear regression equations derived from a reference set of 500 known positions

242 measured inside the chamber (for a similar method, see Bryson, 1992).

243   Single time-point spikes due to random electromagnetic disturbances were removed

244 from the raw kinematic data in each of the three axes (x: near-far; y: left-right; z: down-up)

245 and replaced by linear interpolation from the adjacent points. The position data were then

246 upsampled to 240Hz then filtered with a 4th order zero-lag Butterworth low-pass filter with a

247 15Hz cut-off. Velocity and acceleration were calculated by simple differentiation, and a

248 number of kinematic parameters were extracted. In order to base our conclusions on the

249 maximum number of useful and valid trials, very broad ranges were used to accept valid

250 trials and to reject only very rare artefactual or clearly erroneous trials based on several

251 kinematic parameters. These parameters were decided upon based on experience and on

252 subjective visual exploration of the auditory only dataset, then programmed in for the

253 analysis of all three datasets. Similar parameters have been used elsewhere (e.g., for

10

254 minimum RT, Day & Brown, 2001; for maximum RT, the upper confidence limits of RTs in

255 Veerman et al., 2008 were about 650ms; See Supplementary Table 1 for further examples).

256 Exclusion criteria based on statistical thresholds can bias datasets (van Selst & Jolicoeur,

257 1994), so they were not used. The ranges were: RT (100-750ms); peak 3D acceleration

258 (100-16,000cm/s/s); time of peak 3D acceleration (0-500ms after RT); peak 3D velocity (25-

259 750cm/s/s); time of peak 3D velocity (16-1000ms after RT); peak 3D deceleration (-125 to -

260 16,000cm/s/s); time of peak 3D deceleration (33-1500ms after RT); MT (125-1500ms); path

261 length (10-100cm). Trials were also excluded if the maximum deviation of the head, eye, or

262 gaze was greater than 5 degrees in the first 200ms post-target presentation. This criterion

263 was relatively liberal for the EOG, but relatively conservative for the head and gaze

264 orientation data which were collected outside of the optimal motion tracking range for the

265 tallest participants.

266   Velocity data were aligned with respect to the RT (and, therefore, to the time that the

267 target changed, or would have changed, location). All timing parameters apart from RT were

268 measured relative to this point. For simplicity, and to improve sensitivity for the analyses of

269 most interest, data from left and right targets were combined by recoding the data as 'velocity

270 relative to the first target location', which involved inverting the y (lateral, left-right) axis data.

271 Similarly, the data from two left-handed participants were mirror-reversed across the midline.

272 We therefore refer to movements towards the ipsilateral and contralateral hemispaces: A

273 right-handed participant reaching towards the right target with their right hand is reaching into

274 ipsilateral hemispace, while reaches with the right hand towards the left target are into

275 contralateral hemispace. EOG data were downsampled to 240Hz, allowing alignment with

276 the kinematic data, then filtered with a 4th order Butterworth bandpass filter with cut-offs at 1

277 and 20Hz.

278   The development of analysis routines and parameters was based on the unimodal

279 auditory dataset. This raises the possibility that the analyses were biased towards the

280 properties or results of the auditory (exploratory) dataset. To counter this bias, the main

11

281 conclusions should therefore be based on the multimodal (confirmatory) dataset, in which the

282 unimodal auditory conditions were repeated. Differences between the unimodal auditory

283 experiment and the equivalent conditions in the visual or multimodal experiments should be

284 interpreted cautiously. To summarise: the auditory-only dataset was used to develop the

285 analytic routines, the visual-only dataset was used as a check to confirm whether and when

286 online corrections were present for the more standard unimodal visual targets, and the

287 multimodal dataset was used to explore the possibility of online control for targets that could

288 change modality as well as location.

289

290 *2.6.3 Measuring the time to initiate a movement correction*

291   The only dependent variable of interest was the latency to initiate a movement

292 correction. There are a number of methods for extracting this variable from movement data,

293 including visual inspection of trajectories (Alstermark *et al*., 1984; Day & Brown, 2001; Hyde

294 & Wilson, 2013), use of an arbitrary distance threshold (Johnson, van Beers, & Haggard,

295 2002), examination of standard kinematic parameters or landmarks (e.g., Paulignan *et al*.,

296 1991a), classifying trials based on kinematic parameters (Pisella *et al*., 2000), using

297 statistical criteria based on comparisons between corrected and uncorrected movements

298 using samples of trials (Baugh, Hoe, & Flanaghan, 2012; Cressman, Cameron, Lam, Franks,

299 & Chua, 2010; Turrell, Bard, Fleury, Teasdale, & Martin, 1998) or samples of participants

300 (Aivar, Brenner, & Smeets, 2008; Glover, Miall, & Rushworth, 2005; Kerr, Fox, & Stein,

301 1994), or extrapolating from linear models fit to velocity curves (Veerman *et al*., 2008).

302 Wijdenes *et al.* (2014) discussed and analysed the variety of methods of analysing correction

303 latencies available. Based on simulations of directional movement data with fixed and known

304 correction latency, they concluded that the model-fitting approaches applied to acceleration

305 data are best. Following reviewers' requests to justify our initial choice of method, we used

306 and developed a similar approach, comparing two methods of measuring the time to initiate a

307 movement correction, at three levels of analysis, and using three types of velocity data.

12

308   Unlike Wijdenes et al., (2014), we also examined a wide range of statistical thresholds for

309   determining correction latency, since choosing any threshold one was arbitrary. We tested

310   and optimised these eighteen methods on the unimodal auditory dataset, compared the

311   results with analysis of the visual dataset, then deployed them in the main, final analysis on

312   the visual and multimodal dataset.

313   The methods all used differences in the velocity between trials with a change in target

314   location versus trials with no change (i.e., correction velocity). In all cases, the modalities of

315   targets in the trials with and without changes were the same (i.e., trials with auditory targets

316   on the left followed by visual targets on the left were compared with trials with auditory

317   targets on the left followed by visual targets on the right. We examined both lateral correction

318   velocity (i.e., in the axis which defined the two target locations) as well as three-dimensional

319   correction velocities (derived from x-, y-, and z-axes), for two reasons – the pointing task was

320   three-dimensional, and the experimental axes were not perfectly aligned with the axes of

321   measurement. Table 1 and the following summarise the six main methods and three data

322   types.

323

324   [INSERT TABLE 1 ABOUT HERE]

325

326   • **Method 1: Group correction threshold.** Each participant's data were averaged to

327   produce a single correction velocity curve per condition. The statistics were then

328   based on the group mean of these trajectories.

329   • **Method 2: Participant correction threshold.** For each participant, the trials with a

330   change in target location are treated as a sample, and compared with the null

331   hypothesis of no correction. The one sample t-statistic reflects trajectory corrections

332   for each participant and condition.

- **Method 3: Trial correction threshold.** Each trial with a change in target location is analysed separately. Each trial is compared with the sample of trials of a similar condition in which the target did not change location. This produces a Z-statistic for each sample, reflecting trajectory corrections for each trial.

- **Method 4: Group zero-crossing**: Same as Method 1, except using the zero-crossing point extrapolated from the line joining the 25% and 75% points (relative to the maximum correction velocity, Veerman et al., 2008) of the group mean correction velocity for each condition.

- **Method 5: Participant zero-crossing**. Same as Method 4, except using the zero-crossing points of the lines joining the 25% and 75% points of each participant's mean correction velocity for each condition.

- **Method 6: Trial zero-crossing**. Same as Method 4, except using the zero-crossing points of the lines joining the 25% and 75% points of each trial's correction velocity for each condition.

- **Velocity a: Lateral correction velocity.** The velocity in the y-axis (left-right) on trials with no change in target location subtracted from the same velocity on trials with a change in target location.

- **Velocity b: 3D correction velocity.** The resultant velocity derived from velocities in the x-, y-, and z-axes (resultant=sqrt(x^2+y^2+z^2) on trials with no change in target location subtracted from trials with a change in target location.

- **Velocity c: Statistical components of correction velocity.** The correction velocity in the x-, y-, and z-axes separately, expressed in confidence interval units (i.e., the mean velocity divided by the SE or SD, then divided by the critical statistical value), for trials with no change in target location subtracted from trials with a change in target location. When the sum of statistical components is greater than 1, the correction velocity is outside the (e.g., 95%) velocity confidence ellipsoid.

14

359

2.6.4 Statistical thresholds

Methods 1-3 all required that an arbitrary statistical threshold was chosen to define the latency at which the correction has occurred (see Wijdenes et al., 2014 for detailed discussion and simulations). Previous reports have used arbitrary thresholds of 2 standard errors (e.g., Kerr *et al.*, 1994), confidence intervals (e.g., Turrell *et al.*, 1998), or 2 standard deviations (Baugh *et al.*, 2012), or used the first significant time-point (e.g., Aiver *et al.*, 2008; Cressman *et al.*, 2010; Glover *et al.*, 2005, see Supplementary Table 1). Methods 1 and 2 compared a sample of participants or trials with zero, so the statistical units were standard errors (i.e., t-tests). Method 3 compared single trials with a sample mean, so the statistical units were standard deviations (Z-tests). For each method, one-tailed tests were used as the predictions were unidirectional: participants always moved from the initial target towards the second target (cf Wijdenes et al., 2014). Choosing a statistical threshold was difficult. Initially, we used a 1% probability cut-off (i.e., 2.68 standard errors for n=13, and 2.72 for n=12), as a means of protecting against the increased false-positive rate in Method 1 (i.e., sequential testing against zero for each possible correction latency over 50ms). However, there was no reason to use this criterion for Methods 2 and 3, which extracted correction latencies from individual participants or trials, then performed the final statistics at the group level.

Following an initial review, a systematic exploration of the statistical threshold for determining correction latency was performed, since previous work has not examined the effect of varying statistical threshold within a dataset (e.g., Wijdenes et al., 2014). Correction velocities for Methods 1-3 were determined using 61 statistical thresholds, from 0.0 to 6.0 SE/SD in steps of 0.1. The minimum possible correction latency was set at 50ms, and in each case, correction latency was defined as the first time-point in sequential testing that exceeded the statistical threshold. For all 3 methods, as the statistical threshold increased from 0 to 6, the mean correction latency increased from 50ms to over 300ms. Previous studies have reported visuomotor correction latencies of 90ms or more, so we expected the

15

386 optimal methods to produce latencies in the 100-200ms range (Wijdenes, Brenner, &

387 Smeets, 2011; Archambault et al., 2015). During this exploratory analysis, one constraint we

388 thought important is that the method should be robust to small changes in the choice of

389 statistical threshold: if the sample size was increased, then the statistical threshold might

390 change (e.g., from 2.72 SE/SD to 2.68 for an increase from 12 to 13 participants).

391 Robustness to small changes in threshold seemed to be evident in how the correction

392 latencies changed as a function of statistical threshold: The more robust methods changed

393 smoothly with changes in threshold, the less robust, less smoothly. Finally, seeing these

394 threshold-latency curves, we reasoned that the peak of the curve may be informative: given a

395 method that is robust to small, arbitrary changes in statistical threshold, the statistical

396 threshold at which the greatest change in correction latency occurred should be informative

397 about the actual correction latency. Theoretically, the problem is one of using noisy data to

398 decide when a signal becomes non-zero. In the case of pointing velocity data, the signal is

399 likely Gaussian in shape. We reasoned that the point of maximum change in latency (x) as a

400 function of threshold (y) should correspond to the maximum slope of the velocity curve.

401 Numerical simulations confirmed this (see Supplementary Materials).

402   For subsequent analysis of the data, we chose the mean peak (across experiments

403 and conditions) of these latency-threshold curves (vertical lines in Figure 5a and 5b) as the

404 'optimal' threshold for each method (Table 1). This choice was arbitrary – given a sigmoid-

405 like increase in correction latency as a function of statistical threshold, we assumed that the

406 steepest part of the sigmoid may represent the best threshold to detect the correction

407 velocity signal above the noise. However, prompted by reviewers, we tested and verified this

408 method of choosing a statistical threshold using numerical simulation. See Supplementary

409 Materials and Supplementary Figures 1-3.

410   Using these three post-hoc criteria (correction latencies approximately between 100

411 and 200ms; correction latencies vary smoothly as a function of statistical threshold; peak

412 change in correction latency indicates optimal statistical threshold), Method 3 seemed the

16

413 most robust of those methods requiring an arbitrary statistical threshold (Table 1, Figures 3-

414 6). The optimum statistical threshold produced by this method was the least variable across

415 different experimental conditions (lowest coefficient of variation, CV, across conditions). The

416 threshold for Method 2 was the least robust and most variable across conditions. While these

417 criteria may have biased the subsequent analysis in general, they were applied equally to all

418 conditions, and all valid correction velocities ≥50ms were analysed.

419   In the analyses reported below, Methods 1-3 used what we determined was the optimal

420 statistical threshold for each method (Table 1, right column, bold values; for Method 1, the

421 maximum statistical threshold across conditions was used since the latency-threshold curve

422 was not smooth for the lowest thresholds). For Methods 4-6, no statistical thresholds were

423 required, but some of the velocity data were unsuited to this analysis as they produced very

424 large variability in correction latencies (including negative latencies, cf Wijdenes et al., 2014).

425 Methods 4 and 5 (group and participant correction velocities, respectively) did not work well

426 for the sum of statistical components (4c and 5c), while Method 6 (trial correction velocities)

427 only worked for the lateral (y-axis) correction velocities (6a). Thus, results for fourteen of the

428 eighteen possible methods are reported. These failures of the model-fitting approaches likely

429 indicate that the correction signal was not strong enough in individual trials, or even

430 participants, to produce a meaningful correction velocity with these three-dimensional

431 velocity measures. Rather than being a weakness of these methods, it may be a strength -

432 the correction signal is primarily, or entirely in the lateral velocity component, and adding

433 other directional velocity components merely adds noise to this signal.

434

435 **3 Results**

436   The only dependent variable of relevance to the aims of the study was the latency to

437 correct movements following a change in target location. Additional analyses, for

438 completeness, to ensure the experimental conditions were comparable, and to answer

439 reviewers' comments, are reported as Supplementary Materials. Participants were able to

17

440   correct their reaching-to-point movements in all conditions: visual, auditory, visual-auditory,

441   and auditory-visual targets. Figure 2 shows the group mean (across participant means)

442   position (upper panel), velocity (middle panel), and acceleration (lower panel) curves for the

443   four unimodal auditory conditions. Red and blue curves show trials where the target

444   remained stationary, and magenta and cyan where it switched sides at movement onset. The

445   data were pooled across ipsilateral and contralateral targets to create displacements and

446   velocities relative to the first target location (i.e., towards the ipsilateral side if the first target

447   was in ipsilateral space). The velocity data were then analysed according to the fourteen

448   different methods described above (Figure 3).

449

450   **3.1 Exclusion of data**

451   A mean(±SD) of 5.54(6.63), 7.62(11.9), and 16.1(18.5) trials were removed from the

452   visual-only, auditory-only, and multimodal experiments (7.83, 6.22, & 9.64%, respectively).

453   Of a total of 4955 trials, 364 (7.92%) were removed. 93 (2.02%) were removed for RTs below

454   100ms, 9 (0.19%) on the peak acceleration criteria, 8 (0.17%) on peak velocity, 5 (0.11%) on

455   peak deceleration, 72 (1.57%) on movement time, 22 (0.48%) on path length, 26 (0.57%) on

456   eye position, 1 (0.02%) on head position, 118 (2.57%) on eye velocity (i.e., saccades or EOG

457   artefacts), and 10 (0.22%) on gaze velocity (i.e., combined eye/head shifts not otherwise

458   detected). Supplementary Figure 4 shows the mean eye position and gaze orientation across

459   participants. While the position of the eye in the head, as measured by EOG, was quite

460   stable, most participants seemed to make quite large head rotations while reaching-to-point,

461   affecting the overall gaze orientation. The head orientation data are, however, less reliable,

462   particularly in tall participants, due to the distance of the head receiver from the Polhemus

463   transmitter. Importantly, however, the mean head and gaze orientation was not significantly

464   different from zero throughout the whole trial, and by the time the mean head orientation had

465   changed more than a few degrees, the reaching trajectory corrections had already begun

466   (i.e., within 300ms).

18

467

3.2 Methods of determining correction latency

469   The first main aim of the study was to evaluate different methods of determining

470   correction latency for three-dimensional velocity data, particularly where those methods

471   require an arbitrary statistical threshold for determining when the movement correction

472   begins. Six methods and three types of velocity data were examined.

473

3.2.1 Method 1: Group correction thresholds

475   Since only one statistical comparison was performed on the group-level data, the

476   statistical t-value plotted against the correction latency gave a single curve for each

477   experiment and condition (e.g., Figure 5A, solid black line). Correction latencies for all six

478   experiments and conditions (Unimodal auditory A, and visual V; multimodal VV, AA, VA, and

479   AV) were between 167 and 242ms, with means for Methods 1a, 1b, and 1c of 202ms,

480   206ms, and 195ms respectively. Data for the unimodal auditory condition are shown in

481   Figure 6.

482

3.2.2 Methods 2 and 3: Participant and trial correction thresholds

484   Methods 2 and 3 depended critically upon an arbitrary statistical threshold chosen to

485   define the correction latency. Method 2 used the mean and variability on a participant-by-

486   participant basis, while Method 3 used a trial-by-trial analysis. The resulting participant

487   means were then analysed at the group level. In order to choose a threshold, we

488   systematically varied the threshold and examined the statistical main effects and interactions

489   across all experiments and conditions. To illustrate the problem of choosing a threshold,

490   Figure 4 shows the effect of the initial threshold (x-axis, from 0 to 6, SE for Method 2, SD for

491   Method 3) on the resulting statistical effects in the multimodal experiment (y-axis, t-values).

492   This figure shows a 'dance of the t-values' (cf Cumming, 2012) – how the primary statistical

493   effects of interest (y-axes) change as a function of the initial (and arbitrary) statistical

19

threshold (x-axis). Figures 4a-4c show that for most thresholds, there were no main effects of the initial target modality (most lines are between the critical t-value criteria), but that for some methods and some threshold ranges, significant main effects or interactions can be found. Indeed, for at least one threshold level per method, four of six methods produced a significant main effect of the initial target modality, four of the final target modality, and all six produced significant interactions between initial and final target modalities.

Given the arbitrary choice of statistical threshold, we required a more constrained method of extracting the correction latency for Methods 2 and 3. Exploring the data, we plotted the mean correction latency (Figure 5a) and the change in mean correction latency (Figure 5b) as a function of the initial statistical threshold. For Methods 2a, 2b, and 2c the resulting curve was quite erratic with no clear peak (Figure 5b shows Method 2a, broken lines). By contrast, for Methods 3a, 3b, and 3c, the curve was smoother, and contained a single clear peak (Figure 5b shows Method 3a, solid light grey line).

3.2.3 Methods 4-6: Group, participant, and trial zero-crossings

One clear advantage of a model-fitting approach is that there is no need for an arbitrary statistical threshold to determine the correction latency. Instead, in the three model-fitting methods used here, a straight line is fit to two points on the correction velocity curve, the first at 25% of the maximum correction velocity, the second at the 75% point (Figure 3d, Veerman *et al.*, 2008). This analysis was performed on the three correction velocity types at group, participant, and trial levels. Only five of the nine possible methods succeeded in measuring correction latency, with the lateral correction velocity (a) providing seemingly the most robust inputs for Methods 4, 5, and 6.

3.2.4 Comparison of methods of determining correction latency

Our first aim was to compare different methods of determining correction latency based on real velocity data for three-dimensional pointing movements (cf Wijdenes et al., 2014).

20

521 Figure 6 shows the performance of fourteen different methods for determining correction

522 latency in the unimodal auditory dataset, which was used to optimise the analysis routines.

523 From these data, while Method 2 (participant-level data) produced the lowest estimates of

524 correction latency (Figure 6a, across-condition range=110-230ms), they were also the most

525 variable (SD and CV, Figures 6b, 6c; cf Veerman et al., 2008). To measure the correlation

526 between different methods, r-values determined across participants within each condition

527 were Z-transformed, then averaged across experimental conditions. The mean correlations

528 between correction latency measurements determined by Method 2 and by the other

529 methods (Figures 6d and 6e) were the lowest. The opposite pattern applied to Methods 4-6

530 (zero-crossing, model-fitting): the highest correction latencies (range=182-271ms), but the

531 least variable and most highly-correlated with other methods. Method 3 produced mean

532 (range=130-218ms) and variable correction latencies between those of Methods 2 and

533 Methods 4-6. In summary, Method 2 produced short but variable correction latencies which

534 may provide an estimate of the earliest movement corrections; Methods 4-6 provided longer

535 but less variable latencies; Method 3 may provide a compromise between these two options.

536

537 **3.3 Effects of target modality**

538   Our second aim was to determine correction latency for our auditory, in comparison to

539 our visual target objects. We decided that Method 6a – fitting a straight-line model to lateral

540 correction velocity data on a trial-by-trial basis was the best method. Similar conclusions

541 were reached by others (Wijdenes et al., 2014). We applied this method to analyse the

542 effects of target modality in the three experiments. The unimodal auditory experiment

543 resulted in similar correction latencies (mean±SD=244±49.2ms) to the unimodal visual

544 experiment (241±50.1ms, $t(12)$=-0.274, p=.789). For the same comparison in the multimodal

545 experiment, however, the unimodal auditory corrections (233±62.0ms) were initiated

546 significantly earlier than the unimodal visual (255±53.1ms, $t(11)$=2.29, p=.043).

21

547   A factorial analysis with the variables initial target modality (auditory, visual), and final

548   target modality (auditory, visual) revealed a significant effect of initial target modality,

549   t(11)=6.68, p<.001, with corrections away from auditory targets initiated 35.6±18.5ms faster

550   than away from visual targets. There was no significant effect of final target modality (p=.08)

551   or interaction between initial and final target modality (p=.166).

552

553   **3.4 Relationship between target localisability, target salience, and correction latency**

554   Our auditory and visual targets were matched for location and timing, were driven by

555   the same signal, and were both clearly suprathreshold, but they differed in physical size. We

556   did not explicitly equate auditory and visual targets for localisability, detectability, or salience

557   prior to running the experiments. Rather, in pilot testing we ensured that the auditory stimulus

558   was maximally localisable, and used a relatively dim LED for the visual stimulus. Equating

559   different stimulus attributes is very difficult within vision alone (Veerman et al., 2008, p220),

560   and is perhaps even more so between modalities. Can differences in correction latency

561   between auditory and visual stimuli be explained by differences in localisability of the

562   targets? Analysis of the means across participants' mean endpoint constant errors, both in

563   distance and direction, suggested that, indeed, visual targets were localised between 2.0-

564   2.6mm and 0.1-0.2 degrees better (closer to the target) across experimental conditions and

565   endpoint measurements than auditory. These small differences (approximately half of the

566   LED's diameter; 0.5% of the total movement length; 1% of the angle between targets) were

567   significant in 3 of 5 variables examined (uncorrected p-values .012 to .033). However,

568   arguing against the possibility that localisability *explains* correction latency, only one of 20

569   correlations performed between endpoint error and correction latency (on different measures,

570   both trial-by-trial, and using participants' means) was significant (uncorrected p=.035,

571   correlation between angular error relative to the head, and correction latencies measured

572   with Method 3a). Note, too that minimising endpoint error was not emphasised to the

573   participants, that no instructions were given about exactly how to point (e.g., "position the

22

574 fingertip on a line between the eye/head and the target"), that the greatest magnetic

575 distortions in the positional data were at the movement endpoint, and that many stimulus and

576 task manipulations affect endpoint accuracy in purely visual experiments (see

577 Supplementary Materials).

578   Since it is not clear how auditory and visual conspicuity (or salience) are experimentally

579 to be equated (cf Veerman et al., 2008), we used the proxies for conspicuity that Veerman et

580 al. (2008) used – correction velocity slope – and that Cameron et al., (2013) used – peak

581 correction velocity magnitude. Correction magnitude systematically affects the calculation of

582 correction latency (Wijdenes et al., 2014). Comparing the unimodal (visual only and auditory

583 only) experiments, there was no significant difference in either correction velocity slope

584 $(t(12)=1.01, p=.334)$ or magnitude $(t(12)=1.59, p=.138)$. For correction velocity slopes in the

585 multimodal experiment, there was no significant main effect of initial $(F(1,11)=1.42, p=.258)$,

586 or of final target modality $(F(1,11)=0.002, p=.968,$ and no significant interaction between

587 these factors $(F(1,11)=0.20, p=.663)$. For peak correction velocity magnitudes, there was

588 also no significant main effect of initial $(F(1,11)=0.87, p=.371)$ or final target modality

589 $(F(1,11)=1.21, p=.294)$, while the interaction between these variables only showed a trend

590 $(F(1,11)=4.48, p=.058)$ in which corrections within a modality resulted in non-significantly

591 greater peak correction velocities (visual mean±SD=103±28.9cm/s; auditory=103±24.5cm/s)

592 than corrections between modalities (visual-then-auditory=99.6±29.2cm/s, auditory-then-

593 visual=96.0±26.3cm.s), this trend is in the same direction as the non-significant interaction in

594 correction latencies reported above. These results argue against there being differences

595 between the localisability, detectability, salience, or conspicuity of the auditory and visual,

596 and to a lesser extent the multimodal, target objects in our experiment (Veerman et al., 2008;

597 Cameron et al., 2013).

598

599 **4 Discussion**

23

600   We examined the latency for making online corrections in reaching-to-point movements

601   towards visual, auditory, visual-then-auditory, and auditory-then-visual targets. Across

602   experiments, conditions, and methods, significant movement corrections were evident at

603   110-271ms after the change in target location. These correction latencies were mostly much

604   lower than the mean reaction times for initiating the movements (238-296ms).

605   Our first finding is that fitting a straight-line model (Methods 4-6) to velocity data

606   provides better (lower variability) correction latencies than using a statistical threshold or

607   sequential statistical testing (Methods 1-3). Second, with the model-fitting approaches,

608   examining the component of velocity only in the direction of the target jump is more likely to

609   be successful than using all three directional components (e.g., Methods 4a, 5a, and 6a

610   here). Third, if using a statistical threshold, then the choice of threshold to determine

611   correction latency (whether in SE, SD, or confidence interval units, at the single trial or single

612   participant level) can have dramatic and unpredictable effects on the outcome of subsequent

613   statistical testing (Figure 4). Without examining a wide range of possible statistical thresholds

614   systematically, any choice of threshold is arbitrary and potentially misleading. Similar

615   conclusions were reached by Wijdenes et al., (2014), who decided, based on simulations of

616   lateral movement corrections with known onset, magnitude, and intensity, that fitting straight-

617   line models to acceleration data was optimal. Our work extends their findings by looking

618   systematically at 61 different statistical thresholds, three different types of velocity in three

619   dimensions, and testing real data from three-dimensional movements.

620   Our second novel finding is that the online control of reaching-to-point movements can

621   be just as effective, in latency and magnitude, for auditory as for visual targets (e.g., Boyer *et*

622   *al*., 2013; Veerman *et al*., 2008), at least in a unimodal context, where target modality was

623   fixed. Our third finding is that the latency of online corrections of movements to targets that

624   can change modality as well as location (i.e., movement corrections in a multimodal context)

625   depends upon the initial target modality – corrections away from auditory targets are initiated

626   earlier than corrections away from visual targets. These conclusions must be tempered by

24

627 the fact that we chose only one specific kind of auditory and visual target. Different stimulus

628 attributes may result in different correction latencies (Veerman et al., 2008). Indeed, it may

629 be that whether the participant (or just the relevant parts of their nervous system) treats the

630 two targets as the same is critical for initiating rapid corrections. This is discussed further

631 below.

632

633 **4.1 Latency of corrections to targets presented unimodally**

634   Across the two unimodal experiments, correction latencies did not differ significantly,

635 however within the context of the multimodal experiment, latencies to correct movements to

636 purely auditory targets were significantly shorter (22ms, using the model-fitting Method 6a)

637 than to purely visual targets. This difference is slightly lower than might be expected based

638 on the difference in initial processing time (i.e., RT) for visual over auditory targets: Mean

639 (±SD) RTs to visual targets were 39.1±37.8ms longer than for auditory targets (see

640 Supplementary Materials). Very similar differences in RT and the latency of auditory and

641 visual signals in superior parietal lobe were reported by Molholm, Sehatpour, Mehta,

642 Shpaner, Gomez-Ramirez, Ortigue, Dyke, Schwartz, & Foxe (2006). Why this RT advantage

643 did not translate into a similar advantage in correction latency in the purely unimodal

644 experiments is unclear. One caveat is that the auditory dataset was used for exploration and

645 to optimise the analysis methods, so comparisons between this and other datasets need to

646 be made cautiously (i.e., are potentially biased by 'double-dipping').

647   In general, the correction latencies reported here are slightly longer than those reported

648 in other similar research (e.g., a minimum of 90ms in Paulignan et al., 1991a; ~130-200ms

649 across participants in Veerman et al., 2008; see Supplementary Table 9 for further details).

650 Two relatively trivial factors, and one likely more important factor may contribute to this: First,

651 our kinematic data were smoothed with a 15Hz low-pass cut-off filter, which is higher than

652 some other similar studies (5Hz, Boyer et al., 2013; 8Hz, Paulignan et al., 1991a) - lower-

653 frequency cut-offs smooth the data more. In exploratory work, repeating the analyses with a

25

654 5Hz cut-off resulted in correction latencies an average of 26ms shorter. Second, our

655 kinematic data were recorded, and target location changed, only at 60Hz, meaning that our

656 correction latencies are over-estimated by an average of 8.33ms (half a sample). Finally, and

657 likely most important, our targets were presented in three-dimensional space rather than, for

658 example, on a graphics tablet, computer screen, or other flat surface (e.g., Veerman et al.,

659 2008). This complicates the movements performed, requiring different muscle groupings. It

660 also means that, since the axes of measurement and axes of movement were not perfectly

661 coregistered, components of the movement corrections may have occurred in the x-, y-, and

662 z-axes rather than purely in the lateral y-axis. We attempted to overcome this limitation by

663 examining both lateral correction velocity (Methods 1a-6a, Figure 3a, similar to much

664 previous research), as well as resultant 3D velocity (Methods 1b-6b, Figure 3b), and the

665 individual statistical components of correction velocities (Methods 1c-6c, Figure 3c). Overall,

666 different sampling and analysis parameters may have produced perhaps 20-30ms greater

667 correction latencies, but by far the largest contribution to estimation of correction latency is

668 the method and statistical threshold chosen (Figures 5-6). For example, Leonard, Gritsenko,

669 Ouckama, & Stapley (2011) studied reaching and pointing movements in 3D space,

670 measuring correction latencies in 2D, using Method 3 (individual trials) with a 1SD threshold.

671 They reported correction latencies of around 180-190ms. Using the same statistical threshold

672 and method, our latencies for the unimodal auditory experiment are 128ms for the 1D data

673 (Method 3a), and 121ms for the 3D data (Method 3b).

674

675 **4.2 Latency of corrections to targets presented multimodally**

676   The primary aim of this research was to assess whether the online control of reaching-

677 to-point movements might be multimodal or supramodal, or whether it is, at least in

678 significant part, unimodal. This aim was operationalised by hypothesising that, if online

679 control is substantially unimodal, then having to switch between visual and auditory targets,

26

680   or auditory and visual targets, should incur some cost relative to changing movement

681   trajectories between targets of the same modality. This hypothesis was not supported.

682   The first target modality had a greater effect on correction latencies (36ms, d=1.9) than

683   the second target modality (14ms, d=0.56). For the first target, it was easier for participants

684   to correct their movements *away* from the auditory targets than away from the visual targets.

685   This finding might be explained by hypothesising that reaching movements to visual targets

686   are more 'locked on' to their target; that movements towards initially visual targets are more

687   ballistic and less amenable to online control than movements to auditory targets; or that

688   reaching and pointing movements are generated and controlled using predominantly visual

689   representations of target location, that it is easier to select, maintain, and acquire visual

690   targets than auditory targets. This could be due to a relative imprecision in auditory

691   localization relative to visual localization. This possibility can be tested in future research by

692   systematically manipulating the relative localisability of targets across modalities, and

693   therefore the precision of movements towards them (e.g., Izawa & Shadmehr, 2008).

694   Examining the latency to correct movements towards auditory targets in congenitally blind,

695   recently blind, and blindfolded participants may also shed light on this question. Finally,

696   under an attentional account, auditory localization may be more dependent on focusing

697   attention than visual localization which may be more automatic. Therefore, any attentional

698   switch cost may be larger when one has to switch from visual to auditory than vice versa.

699

700   **4.3 Target detectability, localisability, and number**

701   Following Veerman et al. (2008), and Cameron et al. (2013), we measured the

702   correction velocity slope and peak correction velocity as proxies for the conspicuity (i.e.,

703   detectability, or salience) of our different modality targets, and found no significant effects of

704   target modality, apart from a trend towards corrections within a modality having higher peak

705   correction velocity than between modalities. Without another available measure of

706   conspicuity between modalities, these data suggest that our targets were well-matched, and

27

that the significant differences in correction latencies that we reported are not due to differences in conspicuity. Nevertheless, we can of course not claim that all kinds of auditory and visual targets will produce the same patterns of results. Within vision alone, the particular attributes of the stimuli determine to a great extent the correction latency (Veerman et al., 2008), and this is almost certainly true for auditory, proprioceptive, and tactile targets as well (Cameron & López-Moliner, 2015).

Previous visual studies of online control have used, for example, different initial and final target size and colour (Day & Brown, 2000), or compared correction latencies between objects differing on numerous stimulus attributes and task relevance (Aivar et al., 2008, 2015), or examined corrections away from visual targets that were presented shortly after imperative auditory movement cues (Wijdenes et al., 2011). Similarly, many previous studies have used two or more discrete target objects, LEDs, or illuminated locations on a flat screen, with target illumination switching instantaneously between the two stimuli, rather than moving a single, physical object (Supplementary Table 9; see Day & Lyon, 2001, for a counter-example; and a brief review in Sarlegna & Muthi, 2015). Thus, the present study, as well as most previous studies, of online control implicitly assumes that participants perceive continuity (i.e., apparent motion) between the (illuminated) target locations; that the same, single object is apparently moved. This is the assumption of unity (for examples in multimodal perception, see Vatakis & Spence, 2007). If the online control of movements depends on the assumption of unity, this concern about whether online corrections are the same for a single, moved object versus two sequentially presented objects or locations applies equally to unimodal and multimodal studies alike, and represents an important question for future research to address.

5 Acknowledgements

28

738

6 References

740 Aivar, M. P., Brenner, E., Smeets, J. B. J. (2008). Avoiding moving obstacles. *Experimental*

741   *Brain Research*, **190**(3), 251–264.

742 Alstermark, B., Eide, E., Górska, T., Lundberg, A., & Pettersson, L. G. (1984). Visually

743   guided switching of forelimb target reaching in cats. *Acta Physiologica Scandinavica*,

744   *120*(1), 151–153.

745 Archambault, P. S., Ferrari-Toniolo, S. Caminiti, R. Battaglia-Mayer, A. (2015). Visually-

746   guided correction of hand reaching movements: The neurophysiological bases in the

747   cerebral cortex. *Vision Research*, **110**, 244–256.

748 Baugh, L. A., Hoe, E., & Flanagan, J. R. (2012). Hand-held tools with complex kinematics are

749   efficiently incorporated into movement planning and online control. *Journal of*

750   *Neurophysiology*, **108**(7), 1954–1964.

751 Boyer, E. O., Babayan, B. M., Bevilacqua, F., Noisternig, M., Warusfel, O., Roby-Brami, A.,

752   Hanneton, S., & Viaud-Delmon, I. (2013). From ear to hand: the role of the auditory-

753   motor loop in pointing to an auditory source. *Frontiers in Computational Neuroscience*,

754   *7*, 26.

755 Cameron, B. D., Cheng, D. T., Chua, R., van Donkelaar, P., Binsted, G. (2013). Explicit

756   knowledge and real-time action control: anticipating a change does not make us

757   respond more quickly. *Experimental Brain Research*, **229**(3):359-372.

758 Cameron, B. D., López-Moliner, J. (2015). Target modality affects visually guided online

759   control of reaching. *Vision Research*, **110**, 233-243.

29

760 Cluff, T., Crevecoeur, F., Scott, S. H. (2015). A perspective on multisensory integration and

761   rapid perturbation responses. *Vision Research*, **110**, 215-222.

762 Cohen, Y. E., & Andersen, R. A. (2000). Reaches to sounds encoded in an eye-centered

763   reference frame. *Neuron*, *27*(3), 647–652.

764 Cressman, E. K., Cameron, B. D., Lam, M. Y., Franks, I. M., & Chua, R. (2010). Movement

765   duration does not affect automatic online control. *Human Movement Science*, **29**(6),

766   871–881.

767 Cumming, G. (2012). Understanding the new statistics: effect sizes, confidence intervals, and

768   meta-analysis. Routledge, New York.

769 Day, B. L., Brown, P. (2001). Evidence for subcortical involvement in the visual control of

770   human reaching. *Brain*, **124**(9),1832–1840.

771 Day, B. L., Lyon, I. N. (2000). Voluntary modification of automatic arm movements evoked by

772   motion of a visual target. *Experimental Brain Research*, **130**(2):159-168.

773 Glover, S. R., Miall, R. C., Rushworth, M. F. S. (2005). Parietal rTMS disrupts the initiation

774   but not the execution of on-line adjustments to a perturbation of object size. *Journal of*

775   *Cognitive Neuroscience*, **17**(1), 124–136.

776 Graziano, M. S. A., Reiss, L. A., & Gross, C. G. (1999). A neuronal representation of the

777   location of nearby sounds. *Nature*, *397*(6718), 428–430.

778 Hyde, C. E. A., Wilson, P. H. (2013). Impaired online control in children with developmental

779   coordination disorder reflects developmental immaturity. *Developmental*

780   *Neuropsychology*, **38**(2), 81–97.

781 Izawa, J., & Shadmehr, R. (2008). On-line processing of uncertain information in visuomotor

782   control. *Journal of Neuroscience*, *28*(44), 11360–11368.

783 Johnson, H., van Beers, R. J., Haggard, P. (2002). Action and awareness in pointing tasks.

784   *Experimental Brain Research*, **146**(4), 451–459.

785 Kerr, G. K., Fox, P., & Stein, J. F. (1994). Corrections to unexpected visual changes in the

786   perceived position of the hand during rapid movements. *Human Movement Science*,

787   **15**(5), 763–786.

788 Leonard, J. A., Gritsenko, V., Ouckama, R., & Stapley, P. J. (2011). Postural adjustments for

789   online corrections of arm movements in standing humans. *Journal of Neurophysiology*,

790   **105**(5):2375-2388.

791 Molholm, S., Sehatpour, P., Mehta, A. D., Shpaner, M., Gomez-Ramirez, M., Ortigue, S.,

792   Dyke, J. P., Schwartz, T. H., & Foxe, J. J. (2006). Audio-visual multisensory integration

793   in superior parietal lobule revealed by human intracranial recordings. *Journal of*

794   *Neurophysiology*, *96*(2), 721–729.

795 Paulignan, Y., MacKenzie, C., Marteniuk, R. G., & Jeannerod, M. (1991a). Selective

796   perturbation of visual input during prehension movements. 1. The effects of changing

797   object position. *Experimental Brain Research*, *83*(3), 502–512.

798 Paulignan, Y., Jeannerod, M., MacKenzie, C., & Marteniuk, R. G. (1991b). Selective

799   perturbation of visual input during prehension movements. 2. The effects of changing

800   object size. *Experimental Brain Research*, *87*(2), 407–420.

801 Pettersson, L. G., Lundberg, A., Alstermark, B., Isa, T., & Tantisira, B. (1997). Effect of spinal

802   cord lesions on forelimb target-reaching and on visually guided switching of target-

803   reaching in the cat. *Neuroscience Research*, *29*(3), 241–256.

804 Pisella, L., Gréa, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., Boisson, D., &

805   Rossetti, Y. R. C. (2000). An 'automatic pilot' for the hand in human posterior parietal

806   cortex: Toward reinterpreting optic ataxia. *Nature Neuroscience*, *3*(7), 729–736.

807 Sarlegna, F. R., Mutha, P. K. (2015). The influence of visual target information on the online

808   control of movements. *Vision Research*, **110**, 144-154.

809 Song, J., Rafal, R. D., & McPeek, R. M. (2011). Deficits in reach target selection during

810   inactivation of the midbrain superior colliculus. *Proceedings of the National Academy of*

811   *Sciences USA*, *108*(51), 1433–1440.

31

812   Turrell, Y., Bard, C., Fleury, M., Teasdale, N., & Martin, O. (1998). Corrective loops involved

813   in fast aiming movements: Effect of task and environment. *Experimental Brain*

814   *Research*, **120**(1), 41–51.

815   Vatakis, A., Spence, C. (2007). Crossmodal binding: evaluating the "unity assumption" using

816   audiovisual speech stimuli. *Perception & Psychophysics*, **69**(5):744-56.

817   Veerman, M. M., Brenner, E., & Smeets, J. B. J. (2008). The latency for correcting a

818   movement depends on the visual attribute that defines the target. *Experimental Brain*

819   *Research*, *187*(2), 219–228.

820   Werner, W. (1993). Neurons in the primate superior colliculus are active before and during

821   arm movements to visual targets. *European Journal of Neuroscience*, *5*(4), 335–340.

822   Wijdenes, L. O, Brenner, E. Smeets J. B. V. (2011). Fast and fine-tuned corrections when the

823   target of a hand movement is displaced. *Experimental Brain Research*, **214,** 453–462.

824   Wijdenes, L. O, Brenner, E. Smeets J. B. V. (2014). Analysis of methods to determine the

825   latency of online movement adjustments. *Behaviour Research*, **46,** 131–139.

826   Wijdenes, L. O, Gomi, H. Brenner, E. (2015). Vision Research special issue on the ''On-line

827   Visual Control of Action''. *Vision Research*, **110,** 143.

828   Zhao, H., Warren, W. H. (2015). On-line and model-based approaches to the visual control of

829   action. *Vision Research*, **110**, 190-202.

7 Figure Legends

Figure 1. Experimental apparatus. Participants sat at the centre of a 90cm radius metal hoop (grey arc) supporting three loudspeakers (filled trapeziums) and three LEDs (filled circles) at the centre, and 7.5 degrees to the left and right of the participant's midline. Participants rested their hand on a starting board (grey rectangle), keeping their index finger in a 'start' location (filled circle). Participants wore a tracker on their index finger and vertex (solid squares), and a pair of goggles supporting a laser pointer (filled rectangles).

Figure 2. Mean (±SE) position (a), velocity (b), and acceleration curves (c). Data show reaching-to-point movements towards an auditory target 7.5 degrees on the ipsilateral (red, "Ipsi-ipsi") or contralateral side of the midline (blue, "Contra-contra"). The magenta curves show movements initially directed to an ipsilateral target, and corrected after movement onset to the contralateral target (magenta, "Ipsi-contra"), and cyan curves show movements initially directed to the contralateral, then corrected to the ipsilateral target (cyan, "Contra-ipsi").

Figure 3. Methods of measuring latency to correct a reaching movement. Each panel shows a correction velocity curve with correction latency derived in different ways. (a) The mean (thick black line) lateral (ipsi-to-contra) velocity data are used to find the first point where the lower confidence limit (lower thin black line) of the correction curve is above zero (broken red line). The vertical red line indicates the correction latency. The confidence interval around the mean correction velocity is set according to an arbitrary choice of statistical threshold (e.g., 2.18 standard errors from the mean would be a 95% CI for a single-sample t-test with 13 participants). This procedure is used in Methods 1a, 1b, 2a (example data shown), and 2b. (b) A similar procedure can be used on single trials with a change in target location, by comparing a single correction velocity curve to the mean±SD velocity on trials without a change in target location. Correction latency is determined when

33

857 *the correction velocity exceeds the confidence limit (broken red line). This procedure is used*

858 *in Methods 3a and 3b (example data shown). (c) A 'confidence ellipsoid' is calculated using*

859 *the x (blue), y (green), and z (magenta) statistical components of the correction velocities.*

860 *The statistical components of the x, y, and z correction velocities (i.e., the mean of each*

861 *component of velocity, divided by the SD or SE of the components, then divided by the*

862 *critical statistical threshold value (e.g., t(12)=2.18 for a 95% confidence ellipsoid). Correction*

863 *latency is determined by the first point where the sum of the three components is greater*

864 *than one (i.e., outside the confidence ellipsoid, broken red line). This procedure is used in*

865 *Methods 1c, 2c, and 3c. (d) A model-fitting approach is used to fit a straight line to the first*

866 *points along the correction velocity curve from panel **a** which are greater than or equal to*

867 *25% (25% Vmax, lower solid horizontal red line) and 75% (75% Vmax, upper solid horizontal*

868 *red line) of the maximal correction velocity. The straight line model is then extrapolated back*

869 *to find the zero crossing with the x-axis (broken red line). The zero-crossing is the correction*

870 *latency (vertical red line).*

871

872 ***Figure 4. Effect of statistical thresholds for determining correction latency on main***

873 ***effects of and interactions between initial and final target modality in the multimodal***

874 ***experiment**. Each panel shows the statistical threshold used to determine the correction*

875 *latency on the x-axis (from 0 to 6, corresponding to the number of standard errors from the*

876 *mean for participant mean data (black lines, Methods 2a, 2b, 2c), and standard deviations for*

877 *trial-by-trial data (grey lines, Methods 3a, 3b, 3c). The y-axes show (**a**) the statistical main*

878 *effects (t-values) of the initial target modality, (**b**) the final target modality, (**c**) and the*

879 *interaction between initial and final target modality (**c**). Horizontal red lines show the critical t-*

880 *values for significant effects (5%, two-tailed). Depending on both the method and the*

881 *statistical threshold chosen, both main effects and their interaction can be found 'significant'*

882 *or 'not significant'.*

883

34

884 ***Figure 5. Effect of statistical thresholds on correction latency in the unimodal auditory***

885 ***experiment***. *Each plot shows the statistical threshold on the x-axis and (**a**) the resulting*

886 *mean lateral correction latency or (**b**) the change in mean lateral correction latency on the y-*

887 *axis. Method 1a (black line) produced only a single group correction latency (the first valid*

888 *(i.e.≥50ms) correction latency occurred with the threshold indicated by the vertical line,*

889 *hence the curve is truncated; the curve is smoothed for display purposes), but Methods 2a*

890 *(broken grey line) and 3a (sold grey line) produced a different correction latency for each*

891 *participant and condition (Method 2a) or participant, condition, and trial (Method 3a). The*

892 *data in panel b show that the group mean correction latency changes more smoothly as a*

893 *function of the statistical threshold for Method 3a than for Method 2a. Method 3a produces a*

894 *smoother curve (panel a), and a single large peak (panel b), while Method 2a produces a*

895 *less smooth curve and multiple peaks. The point of maximal change in correction latency as*

896 *a function of statistical threshold (i.e., the steepest part of Figure 5A – horizontal lines on*

897 *each curve) was taken as the optimal statistical threshold (vertical lines) to use for*

898 *determining correction latency. See Supplementary Materials for a numerical simulation and*

899 *validation of this approach.*

900

901 ***Figure 6. Descriptive and diagnostic statistics for fourteen methods of determining***

902 ***correction latency***. *Each plot shows the group mean (±95% CI where available) of*

903 *correction latency statistics from each of fourteen methods, extracted from the correction*

904 *velocity data. (**a**) Mean correction latency for all 14 methods. (**b**) Standard deviation (SD) of*

905 *correction latency across participants for Methods 2, 3, 5, and 6. (**c**) Coefficient of variation*

906 *(CV) correction latency (SD/mean). (**d**) Correlations between different procedures of*

907 *estimating correction latency within each Method (e.g., the correlations for Method 2a are*

908 *with Method 2b, and 2c). (**e**) Correlations between different procedures and methods (e.g.,*

909 *the correlations for Method 2a are with Methods 3a, 3b, 3c, 5a, 5b, and 6a). Correlations are*

35

910 *expressed as Z-values after Fisher's r-to-Z transformation, to allow valid use of parametric*

911 *statistics.*

36

Table 1: Performance of nine methods of determining correction latency for visual,

auditory, visual-then-auditory, and auditory-then-visual targets using statistical

thresholds

Method	Level	Velocity	Units	Optimal threshold per condition						Mean (SD) (CV)
				VV	AA	VV	VA	AV	AA	
1a	Group	lateral	SE	1.70	2.56	2.42	2.82	2.06	1.87	**2.24** (0.432) (0.193)
1b	Group	3D-r	SE	2.58	4.47	2.68	1.13	1.91	1.28	**2.34** (1.22) (0.522)
1c	Group	3D-c	SE	2.73	1.88	1.85	2.92	3.07	1.68	**2.36** (0.618) (0.262)
2a	Subject	lateral	SE	1.2	0.9	2.2	1.1	2.3	1.6	**1.55** (0.589) (0.380)
2b	Subject	3D-r	SE	1.1	2.0	2.5	1.4	2.1	2.2	**1.88** (0.527) (0.280)
2c	Subject	3D-c	SE	0.6	0.5	1.0	0.9	1.5	0.8	**0.883** (0.354) (0.401)
3a	Trial	lateral	SD	1.2	1.0	1.4	1.2	1.3	1.2	**1.22** (0.133) (0.109)
3b	Trial	3D-r	SD	1.2	1.5	1.9	1.6	1.3	1.0	**1.42** (0.319) (0.225)
3c	Trial	3D-c	SD	1.9	2.1	2.5	1.7	2.0	2.4	**2.10** (0.303) (0.144)

915 Data shown are the optimal thresholds for determining correction latencies in SE (standard
916 error) or SD (standard deviation) statistical units. CV: Coefficient of variation=SD/mean. VV:
917 Visual-visual; AA: Auditory-auditory; VA: Visual-auditory; AV: Auditory-visual; Mod. 1: Main
918 effect of initial target modality; Mod. 2: Main effect of final target modality. 3D-r: Resultant 3D
919 velocity=sqrt(x^2+y^2+z^2); 3D-c: Statistical components of velocity in x-, y-, and z-axes.

Table 2. Fourteen measures of correction latency for visual, auditory, visual-then-auditory, and auditory-then-visual targets

Method	Unimodal			Multimodal						
	VV	AA	diff (p)	VV	AA	VA	AV	Mod 1 diff (p)	Mod 2 diff (p)	Interaction diff (p)
1a	179	192	-13.0	217	217	221	183	19.0	-19.0	15.0
1b	200	204	-4.00	217	179	242	196	42.0	-4.00	-21.0
1c	183	192	-9.00	217	204	217	171	31.5	-18.5	-14.5
2a	143 (46.5)	121 (37.6)	21.6 (.10)	110 (37.3)	118 (57.9)	181 (180)	126 (50.1)	23.8 (.47)	-31.5 (.31)	-40.0 (.17)
2b	167 (42.6)	153 (71.5)	14.6 (.47)	121 (57.1)	138 (65.2)	176 (55.5)	167 (45.6)	-4.08 (.84)	-12.8 (.44)	**-41.6 (.028)**
2c	230 (59.3)	214 (53.4)	16.5 (.51)	112 (59.3)	155 (80.4)	155 (53.0)	139 (39.3)	-7.73 (.71)	**-34.6 (.032)**	-7.73 (.62)
3a	152 (21.4)	175 (54.4)	-22.4 (.10)	137 (28.4)	130 (34.2)	164 (53.9)	145 (31.1)	13.3 (.16)	-5.74 (.44)	-21.0 (.13)
3b	186 (32.9)	211 (34.6)	25.7 (.06)	168 (35.8)	183 (43.1)	218 (59.8)	202 (54.4)	0.76 (.95)	-15.5 (.24)	-34.3 (.10)
3c	175 (35.0)	191 (43.0)	-16.5 (.28)	149 (26.8)	145 (38.8)	208 (65.8)	177 (50.6)	17.8 (.24)	-13.8 (.25)	**-45.3 (.013)**
4a	206	210	-3.47	223	199	222	186	30.1	-6.31	6.54
4b	219	223	-3.71	227	189	235	213	30.3	7.91	-16.5
5a	229 (46.6)	223 (44.8)	5.60 (.56)	248 (46.8)	215 (52.3)	244 (47.2)	211 (46.8)	**32.8 (.005)**	0.02 (.99)	3.98 (.50)
5b	247 (49.7)	246 (46.9)	1.03 (.92)	254 (52.8)	227 (53.7)	260 (48.2)	237 (48.0)	**25.0 (.002)**	1.73 (.78)	-7.63 (.15)
6a	246 (50.1)	248 (49.2)	-2.67 (.79)	255 (53.1)	233 (62.0)	275 (55.2)	225 (57.2)	**35.6 (<.001)**	-14.0 (.08)	-5.99 (.17)

Data shown are the mean (SD) correction latencies in ms using six methods (Table 1). p and ANOVA columns (Mod 1, Mod 2, Interaction) show p-values for t-test and ANOVA terms respectively. diff=effect size in ms (visual-auditory or unimodal-multimodal). Bold font indicates significant effects (p<.05, uncorrected). VV: Visual-visual; AA: Auditory-auditory; VA: Visual-auditory; AV: Auditory-visual; Mod. 1: Main effect of initial target modality; Mod. 2: Main effect of final target modality.