A comparison of methods of assessing cue combination during navigation

Phillip M. Newman¹ · Timothy P. McNamara¹

Published online: 23 July 2020
© The Psychonomic Society, Inc. 2020

Abstract

Mobile organisms make use of spatial cues to navigate effectively in the world, such as visual and self-motion cues. Over the past decade, researchers have investigated how human navigators combine spatial cues, and whether cue combination is optimal according to statistical principles, by varying the number of cues available in homing tasks. The methodological approaches employed by researchers have varied, however. One important methodological difference exists in the number of cues available to the navigator during the outbound path for single-cue trials. In some studies, navigators have access to all spatial cues on the outbound path and all but one cue is eliminated prior to execution of the return path in the single-cue conditions; in other studies, navigators only have access to one spatial cue on the outbound and return paths in the single-cue conditions. If navigators can integrate cues along the outbound path, single-cue estimates may be contaminated by the undesired cue, which will in turn affect the predictions of models of optimal cue integration. In the current experiment, we manipulated the number of cues available during the outbound path for single-cue trials, while keeping dual-cue trials constant. This variable did not affect performance in the homing task; in particular, homing performance was better in dual-cue conditions than in single-cue conditions and was statistically optimal. Both methodological approaches to measuring spatial cue integration during navigation are appropriate.

Keywords Allothetic cue · Cue reliability · Idiothetic cue · Spatial cue integration

Introduction

Mobile animals require mechanisms to remain oriented in the environment as they locomote. One such mechanism is the utilization of spatial cues. Spatial cues can come from the environment (allothetic cues) and the body (idiothetic cues). For example, local and distal landmarks can be used to establish allocentric representations of position and orientation with respect to environmental axes (Biegler & Morris, 1999; Chamizo, 2003; Chamizo, Sterio, & Mackintosh, 1985; Hamilton & Sutherland, 1999; Hardt, Hupbach, & Nadel, 2009). Additionally, vestibular signals, proprioception, and motor efference copy allow navigators to keep track of their velocity and changes in heading direction during locomotion (e.g., Kalia, Schrater, & Legge, 2013; Klatzky, 1998; Philbeck & O'Leary, 2005). However, spatial cues are inherently noisy

and do not always yield reliable representations of location. To circumvent the inherent uncertainty of spatial cues, navigators can combine multiple cues for more precise spatial representations (Bates & Wolbers, 2014; Butler, Smith, Campos, & Bühlhoff, 2010; Chen, McNamara, Kelly, & Wolbers, 2017; Cheng, Shettleworth, Huttenlocher, & Rieser, 2007; Frissen, Campos, Souman, & Ernst, 2011; Kalia, Schrater, & Legge, 2013; Nardini, Jones, Bedford, & Braddick, 2008; Petrini, Caradonna, Foster, Burgess, & Nardini, 2016; Philbeck & O'Leary, 2005; Ratliff & Newcombe, 2008; Sjolund, Kelly, & McNamara, 2018; Tcheang, Bulthoff, & Burgess, 2011; Twyman, Holden, & Newcombe, 2018; Xu, Regier, & Newcombe, 2017; Zhang, Mou, Lei, & Du, 2019; Zhao & Warren, 2015a, 2015b).

Recently, researchers have proposed models of spatial cue integration to explain how human navigators combine allothetic and idiothetic cues when making navigational decisions, for instance, when returning to previously visited locations. For example, the adaptive-combination model (Ratliff & Newcombe, 2008; Twyman, Holden, & Newcombe, 2018; Xu, Regier, & Newcombe, 2017) and Maximum Likelihood Estimation (MLE) model (e.g., Chen et al., 2017; Cheng, et al., 2007; Frissen et al., 2011; Nardini et al., 2008; Petrini

✉ Phillip M. Newman
phillip.m.newman@vanderbilt.edu

¹ Department of Psychology, Vanderbilt University,
Nashville, TN, USA

et al., 2016; Sjolund et al., 2018) propose that navigators combine and weight spatial cues based on cue reliability. Specifically, the MLE model makes mathematical predictions of optimal cue combination. Accordingly, each cue provides a probability distribution over locations, whose reliability is equal to the inverse of its variance (i.e., more precise representations are more reliable). Weights are assigned to cues based on their relative reliabilities and are proportional to a given cue's reliability. The optimal estimate of a target location is represented by the posterior distribution, the mean and variance of which are linear combinations of the weighted means and variances for each cue, respectively (see Eqs. 1–5).¹ Therefore, the MLE model of cue integration predicts that navigators optimally weight and integrate multiple spatial cues according to cue reliability.

When conducting cue integration experiments, the number of available cues is manipulated (usually within-subjects; Alais & Burr, 2004; Battaglia, Jacobs, & Aslin, 2003; Ernst & Banks, 2002; Friedmann, Ludvig, & Legge, 2013; Girshick & Banks, 2009; Hillis, Watt, Landy, & Banks, 2004; Jacobs, 1999; Mou & Spetch, 2013; Oruç, Maloney, & Landy, 2003; Rohde, van Dam, & Ernst, 2016). On some trials, all cues are available to the participant (combination condition). On other trials, all cues are available but provide conflicting estimates of the target (conflict condition). Critically, there are n single-cue conditions, one for each of the n cues being investigated. Response distributions from the single-cue conditions are used to compute cue reliabilities and predicted weights, from which predictions of optimal cue combination are computed. The reliability of cue i is computed as

$$r_i = \frac{1}{\sigma_i^2} \quad (1)$$

The optimal weight for cue i is

$$w_i = \frac{r_i}{\sum_{j=1}^n r_j} \quad (2)$$

The optimal combination of the two cues is given by

$$\mu_o = \sum_{i=1}^n w_i \mu_i \quad (3)$$

$$\sigma_o^2 = \sum_{i=1}^n w_i^2 \sigma_i^2 \quad (4)$$

where μ_i and σ_i^2 are the mean and variance for cue i , respectively. The variance of the posterior is necessarily less than or equal to the smaller of the variances of the two single cues. Observed cue weights are obtained from performance in the conflict condition by computing the relative distance of

single-cue distributions to the conflict distribution. The observed weight for cue i is given by

$$\hat{w}_i = \frac{d_i^{-1}}{\sum_{j=1}^n d_j^{-1}} \quad (5)$$

where

$$d_i = |\mu_i - \mu_{\text{conflict}}| \quad (6)$$

Note that predicted and observed weights each sum to one. If navigators optimally combine spatial cues, response variability in the combination and conflict conditions will be equal to the variability of the posterior.

Recent work by Chen et al. (2017) showed across multiple experiments that human navigators can optimally combine landmark and self-motion cues according to the MLE model. Using a within-subjects design, each participant completed all four conditions in immersive virtual reality: landmark cues only, self-motion cues only, combination, and conflict. In each trial, participants performed a homing task, wherein they walked a two-legged outbound path before returning directly to the beginning of the path. In all conditions, participants completed a brief counting task prior to the return path. During landmark trials, participants were disoriented during the counting task, disrupting self-motion cues. Thus, participants were required to rely on the landmarks during the return path. During self-motion trials, the landmarks were removed from the environment following counting, forcing reliance on self-motion cues. During combined trials, neither cue was disrupted. During conflict trials, the landmarks were rotated 15 degrees around the end of the outbound path, placing landmark and self-motion cues in conflict. Across multiple experiments (but not all), response variabilities in the combined and conflict conditions were consistent with optimal integration and mean observed weights were equivalent to optimal weights.

The experiments by Chen et al. are similar to other cue integration experiments in navigation (e.g., Nardini et al., 2008; Sjolund et al., 2018; Zhao & Warren, 2015b) in that both cues were available along the outbound path in all cue conditions. However, other researchers have taken a different approach by leaving only one cue present during the outbound path in the single-cue conditions. For example, Petrini et al. (2016) compared cue integration in adults to 10- to 11-year-olds in immersive virtual reality. Participants learned a two-legged path in three conditions: vision only, self-motion only, and combination. In the vision-only condition, participants watched a pre-recording of someone else's perspective traveling along the two-legged path. In the self-motion condition, participants were guided along the path in darkness by the experimenter. In the combined condition, participants were

¹ We refer to the optimal, combined distribution as the posterior, even though the MLE model assumes uniform priors.

guided along the path by the experimenter while the virtual environment was visible. Participants in the self-motion and combined conditions were led along an irregular path back to the starting point, from which participants in all conditions attempted to recreate the path in darkness. Although 10- to 11-year-olds showed optimal integration of visual and self-motion cues, adults were suboptimal in their cue integration, preferring self-motion cues over visual cues.

To our knowledge, the number of cues available during the outbound path in single-cue conditions has yet to be systematically manipulated in a cue integration context. Such methodological differences might alter performance in single-cue conditions by allowing for cue integration during the outbound path when both cues are available. For example, research conducted by Philbeck and O’Leary (2005) suggests that previously viewed landmark cues can reduce response variability during a blindfolded navigation task. Participants navigated to a target location while blindfolded in one of three conditions: no-landmark, audible landmark, and remembered landmark. In all conditions, participants stood in a room and viewed a target location 6.2 m away for 5 s. In the audible and remembered landmark conditions, a cone was placed 4.2 m away and served as a landmark during viewing. Following viewing the target, participants were blindfolded and attempted to walk directly to the target location. Participants in the audible condition received audible feedback when 2 m away from the 4.2-m mark to signal that they were approaching the landmark, whereas no feedback was given in the remembered landmark condition. Surprisingly, participants in the remembered landmark condition showed the least response variability, nearly halving the response variability of the no-landmark condition. Their results suggest that participants were able to reset some of the accumulated error when passing by the remembered landmark location.

A more recent study by Kalia, Schrater, and Legge (2013) found similar results. Participants stood in a hallway and studied a target location (5–11 m) marked by a red LED light. Participants were then blindfolded and told to walk in the direction of the target until cued to stop. After stopping, participants indicated their perceived location on a tactile map and stated whether they thought the location they walked to was the same as or different from the target. The weight given to self-motion cues was equal to the slope of the line fitted to the participants’ judged location relative to their actual location. If participants judged their location to be the same as the walked location, weight given to self-motion cues would be equal to one. On the other hand, if participants judged their location to be the same as the visual target, weight given to visual cues would be equal to one. When participants perceived their walked location to be different from the visual target, self-motion was assigned a weight nearing one. When the

walked location was perceived to be the same as the visual target, participants averaged the visual information with self-motion information. Interestingly, response variability was reduced whether or not the walked location was perceived as congruent with the visual target. However, the reduction was greater when they were perceived as congruent. Their results suggest that remembered visual information can enhance navigation, even when such cues are no longer available.

The current study was designed to investigate methodological differences in cue integration experiments in navigation using the MLE model. Participants completed a homing task, wherein the number of cues available during the outbound path in single-cue conditions was systematically manipulated. Participants were randomly assigned to one of two path conditions: In the both-present condition, both cues were available during the outbound path for single-cue trials, with one cue being disrupted prior to the return path; in the one-present condition, only one cue was available during the outbound path for single-cue trials. Combination and conflict trials were identical in the one-present and both-present conditions, i.e., both visual and self-motion cues were present along the outbound and return paths. We tested three major hypotheses:

- 1 If cue integration occurs during decision-making (return path) only and is optimal, performance in dual-cue conditions should be equal to predicted performance from single-cue conditions (see Eqs. 1–6), regardless of the number of cues available along the outbound path in single-cue conditions.
- 2 If optimal cue integration occurs along the outbound path in the both-present condition, and the integrated representation remains intact during the return path, then in the both-present condition, single-cue performance should be equal to dual-cue performance and dual-cue performance would appear to be suboptimal. Predicted performance in the single-cue trials would be improved by the integration of the (eventually) removed cue on the outbound path (e.g., visual cue during self-motion trials). (This pattern of results is also predicted if no cue integration occurs.)
- 3 If optimal cue integration occurs along the outbound path in the both-present condition, but the integrated representation is rendered less reliable when one cue is removed before the return path, then in the both-present condition, dual-cue performance should be better than single-cue performance but suboptimal. The availability of both cues during the return path would produce a more reliable integrated representation in dual-cue trials but predicted performance would still be reduced by integration of cues during the outbound path in single-cue trials, even if the integrated representation is disrupted during the return path.

Methods

Participants

Undergraduate students ($N = 48$; age $M = 20.5$, $SD = 2.21$; 24 females) from Vanderbilt University participated in exchange for credit in a psychology course or monetary compensation. Participants were randomly assigned to one of two path conditions: both-present ($n = 24$; 12 females) or one-present. Previous cue combination studies in navigation (e.g., Bates & Wolbers, 2014; Chen et al., 2017; Sjolund et al., 2018) have used similar sample sizes, finding medium effect sizes ($\eta_G^2s = .11-.18$). Data for six additional participants were excluded due to equipment malfunction ($n = 5$) or failure to correctly follow experimental procedures ($n = 1$). A trial was considered an outlier if the response error fell above three times the interquartile range above the third quartile for a given cue condition. Less than .01% of trials were trimmed.

Materials and procedure

The immersive virtual environment was rendered in Unity, a multiplatform game engine (<https://unity.com/>). The environment was displayed in the HTC Vive head-mounted display (HMD) with a resolution of 1080 x 1200 per eye, refreshed at 90 Hz. The HMD's field-of-view is approximately 110 degrees diagonally. Participants used the HTC Vive wireless controller throughout the experiment. Position and orientation tracking were supported by HTC Vive's Lighthouse tracking system, with a 4 x 4 m tracking space. The size of the room was 7.3 x 8.5 m. The TPCast (<https://www.tpcastvr.com/>) supported wireless tracking of the HMD. Participants were able to physically rotate and move throughout the virtual environment. The experiment was conducted on a computer with an Intel Core i7-6700k processor, 32 GB of RAM, and an NVIDIA GTX 1080 graphics card.

Given that video game training has been demonstrated to enhance spatial abilities (see Uttal et al., 2013), we administered a video game history and habits questionnaire (originally developed by Boot et al., 2008) to control for video game experience. The survey asked participants about demographics, hours per week spent playing video games, age when first started playing video games, and owned video game consoles. Only eight participants reported being an active gamer, with 20 participants reporting playing 0 h per week and only three reporting playing more than 5 h per week. Therefore, we do not discuss this metric further.

The virtual environment consisted of an infinite ground plane and three landmarks: A tree, rock, and tower. Participants started each trial at a constant standing location, facing a constant direction (Fig. 1). During each trial, participants performed a homing task by navigating to a home location, and then following a two-legged outbound path. Across trials, the home location and second location of the path were randomly sampled from one of four possible locations each, yielding 16 possible paths. The final location of the path was constant across the trials. Thus, for each trial, the participant walked to three locations in order: Home location, second location, then final location. The home and second locations of the path were 2.15 m and 1.3 m from the final path location, respectively. At the end of the path, participants attempted to return directly to the home location. In both the both-present and one-present conditions, participants completed eleven blocks (including one practice block) of four trial types in a random order: visual-only, self-motion-only, combination, and conflict.

In the both-present condition, the procedure during the outbound path was the same for all trial types. At the start of each trial, participants could see the three landmarks and a red post marking the home location. Upon arriving at the red post, participants completed the outbound path by following yellow posts appearing in succession. Upon arriving at the end of the path, the screen in the HMD went black and displayed text

Fig. 1 Experimental environment from the perspective of the starting location of each trial. The red post marks the home location for a given trial

prompting the participant to wait for the experimenter. During visual-only trials, participants were disoriented by spinning counter-clockwise in a chair while counting down 15 from a randomly chosen number (e.g., from 45 to 30) in steps of three. The virtual environment then reappeared in the HMD and participants attempted to return directly to the home location, pressing a button on the controller to confirm their response. During self-motion-only trials, participants stood in place while counting. Afterwards, the landmarks were removed from the environment and the participants attempted to return to the home location. During combination and conflict trials, participants stood in place when counting and the landmarks remained present during the return path. However, landmarks were rotated 20 degrees around the final yellow post during the return path of conflict trials, placing self-motion and visual cues in conflict.

In the one-present condition, the procedure for combination and conflict trials was identical to the both-present procedure. During visual-only trials, the screen in the HMD went black and the experimenter led participants directly to the final yellow post from the starting position by arm. A video then played through the HMD in which another person's perspective was shown traveling along the outbound path by following the colored posts (with landmarks visible). Upon reaching the final yellow post, the video ended, and participants completed the counting task. The screen in the HMD turned back on and participants attempted to return to the home location as seen in the video. The videos were captured while a naïve research assistant traveled along all possible outbound paths, starting at the same constant location described above. During self-motion-only trials, the screen in the HMD went black and the experimenter led participants from the starting location to the home location, identifying the home location to participants upon arrival, and then along the rest of the outbound path. After reaching the end of the path, participants completed the counting task, and then attempted to return to the home location in darkness. The procedure for the self-motion-only trials in the one-present condition matched that used by Petrini et al. (2016), such that both the outbound and return paths were walked in darkness.²

Analyses

The home location for each trial was treated as the origin in a coordinate system in which the y-axis was parallel to the correct walking direction (i.e., the vector from the end of the

outbound path to the home location) and the x-axis was perpendicular to this axis. Response locations were specified as points on the Euclidean plane defined by this coordinate system. Response accuracy was defined as the Euclidean distance between the mean response location and the target location (i.e., the origin; accuracy is not defined in the conflict condition, as there are two correct locations: one defined by the landmarks and the other defined by self-motion cues). Following previous work (Chen et al., 2017; Nardini et al., 2008; Sjolund et al., 2018), the standard deviation for each condition was calculated using the absolute distance of each response relative to the mean response location. Response errors were computed for targets on each side of the environment separately, and then pooled across sides. Using Equation 4, optimal integration was calculated by combining the response variances from the single-cue conditions. Because access to more cues during decision making is hypothesized to reduce response variability, one-tailed tests were used to compare combination and conflict response variability to single-cue response variability. Using Eqs. 2 and 5, optimal and observed cue weights were calculated as normalized single-cue reliabilities and normalized inverse distances of the mean response locations in the single-cue conditions to the conflict condition, respectively. The Holm–Bonferroni correction was used to correct for multiple comparisons. We did not correct for multiple comparisons when conducting tests comparing performance in the combined and conflict condition to model predictions as higher cost is assigned to falsely accepting the model (cf. Chen et al., 2017).

We also computed the Bayes factor for all comparisons (Jarosz & Wiley, 2014). We considered a Bayes factor greater than three as sufficient evidence for the null hypothesis, and a Bayes factor less than one-third as sufficient evidence for the alternative hypothesis. When comparing performance in the combination and conflict condition to model predictions, if the p value did not reach significance and the Bayes factor was between one and three, we considered cues to be combined near-optimally (cf. Chen et al., 2017). The prior for the Bayes factor was a central Cauchy distribution with scale r on effect size set to 0.707. This is a common prior and scale used for Bayes factor analysis; interpretations of the Bayes factor is seldom affected by changes in scale r (Rouder, Speckman, Sun, Morey, & Iverson, 2009).

Results

Response accuracy

Response accuracy was examined using a mixed factorial ANOVA, with path condition as a between-subjects factor and cue condition as a within-subjects factor (see Table 1 for

² Although the return path during self-motion trials in the one-present and both-present conditions were not identical (i.e., the screen in the HMD was black in the one-present condition while the virtual environment, absent of landmarks, was visible in the both-present condition, providing possible optic flow cues), this difference did not impact the results; response variabilities in self-motion-only trials were identical across the one- and both-present conditions.

Table 1 Means (standard deviations) for accuracy and variability (standard deviation) measures for each cue condition by path condition

	Both-present		One-present	
	Acc.	Var. (SD)	Acc.	Var. (SD)
Visual-only	0.85 (0.40)	0.82 (0.45)	0.63 (0.28)	0.74 (0.31)
Self-motion-only	0.65 (0.31)	0.86 (0.28)	0.74 (0.51)	0.86 (0.36)
Combination	0.25 (0.25)	0.49 (0.14)	0.53 (0.16)	0.49 (0.13)
Conflict	--	0.52 (0.16)	--	0.58 (0.15)
Optimal	--	0.54 (0.18)	--	0.49 (0.12)

means). Mauchly's test for departure from sphericity was significant ($p = .018$), so the Greenhouse–Geisser correction was used. The main-effect of path condition was not significant, $F < 1$. The main-effect of cue condition was significant, $F(2,96) = 6.55$, GG epsilon = .84, $p = .004$, $\eta_p^2 = .12$. Post hoc comparisons showed improved accuracy in the combination condition ($M = 0.51$, $SD = 0.21$) relative to the visual-only condition ($M = 0.74$, $SD = 0.38$), $t(47) = 4.03$, $p < .001$, $\alpha_{crit} = .017$ $d = .58$ (BF < 0.01), and self-motion-only condition ($M = 0.70$, $SD = 0.42$), $t(47) = 2.80$, $p = .007$, $\alpha_{crit} = .025$ $d = .40$ (BF = 0.20). The visual-only and self-motion-only conditions showed similar accuracy, $t(47) = 0.52$, $p = .607$, $\alpha_{crit} = .050$ $d = .07$ (BF = 5.61). The path condition by cue condition interaction was not significant, $F(2,96) = 2.99$, GG epsilon = .84, $p = .064$, $\eta_p^2 = .06$.

Response variability

Response variability was examined using a mixed factorial ANOVA, with path condition as a between-subjects factor and cue condition as a within-subjects factor. Mauchly's test for departure from sphericity was significant ($p < .001$), so the Greenhouse–Geisser correction was used. The main-effect of path condition was not significant, $F < 1$. The main-effect of cue condition was significant, $F(3,138) = 22.69$, GG epsilon = .64, $p < .001$, $\eta_p^2 = .33$ (Fig. 2). Planned comparisons showed

Fig. 2 Response variability (standard deviation) collapsed across path conditions as a function of cue condition

that response variability was smaller in the combination condition ($M = 0.49$, $SD = 0.14$) than the visual-only condition ($M = 0.78$, $SD = 0.39$), $t(47) = 5.44$, $p < .001$, $\alpha_{crit} = .050$, $d = .79$ (BF < 0.01), and self-motion-only condition ($M = 0.86$, $SD = 0.32$), $t(47) = 7.90$, $p < .001$, $\alpha_{crit} = .025$, $d = 1.14$ (BF < 0.01). Response variability was smaller in the conflict condition ($M = 0.55$, $SD = 0.16$) than the visual-only, $t(47) = 4.22$, $p < .001$, $\alpha_{crit} = .100$ $d = .61$ (BF < 0.01), and self-motion-only conditions, $t(47) = 6.13$, $p < .001$, $\alpha_{crit} = .033$ $d = .89$ (BF < 0.01). Optimal variability predicted from single-cue performance ($M = 0.52$, $SD = 0.16$) was not significantly different from observed variability in the combination condition, $t(48) = -0.98$, $p = .334$, $\alpha_{crit} = .050$, $d = .14$ (BF = 4.06), or observed variability in the conflict condition, $t(47) = 1.17$, $p = .247$, $\alpha_{crit} = .050$, $d = .17$ (BF = 3.36). These comparisons indicate that performance in the dual-cue conditions was optimal. The path condition by cue condition interaction was not significant, $F < 1$.

Cue weights

Observed and optimal visual cue weights were examined using a mixed factorial ANOVA, with path condition as the between-subjects factor and weight type (observed vs. optimal) as the within-subjects factor. Both the main-effect of path condition and weight type were not significant, $F(1,46) = 1.14$, $p = .292$, $\eta_p^2 = .02$ and $F < 1$, respectively. The interaction between path condition and weight type was not significant, $F(1,46) = 2.38$, $p = .130$, $\eta_p^2 = .04$ (see Table 2 for means).

To examine the linear relationship between weight types, observed visual weights were regressed onto optimal visual weights (Fig. 3). Because no differences were observed between path conditions with respect to cue weights in the conflict condition, the data were collapsed across path conditions. Observed cue weights were predicted by optimal weights, $F(1,22) = 13.34$, $p < .001$, Adj. $R^2 = .21$. To test for optimal cue weighting, the slope and intercept of the regression equation were compared to one and zero, respectively. A slope value less than one and an intercept value greater than zero would reflect a suboptimal relationship between observed and optimal weights predicted by single-cue variabilities. The slope and intercept values were significantly less than one and greater than zero, respectively, suggesting suboptimal cue weights ($\beta = 0.37$, $p < .001$ and $\alpha = 0.37$, $p < .001$).

Table 2 Means (standard deviations) for observed and optimal visual cue weights by path condition

	Both-present	One-present
Observed vis. weight	0.52 (0.19)	0.63 (0.17)
Optimal vis. weight	0.54 (0.22)	0.55 (0.26)

Fig. 3 Observed visual weights from the conflict condition regressed onto optimal visual weights, collapsed across path conditions. Each point corresponds to one participant ($N = 48$)

Although mean observed cue weights were consistent with mean optimal cue weights, observed cue weights were suboptimal at the individual level (see also, Chen et al., 2017).

Discussion

The current experiment compared two methods of assessing spatial cue integration. A notable difference among experiments in previous work lies in the number of cues available to the navigator for single-cue trials. For example, Chen et al. (2017) had both landmark and self-motion cues available during the outbound path, whereas Petrini et al. (2016) only had the targeted cue available. Related work has shown that humans can integrate previously studied visual cues with self-motion cues during blindfolded navigation (Kalia, Schrater, & Legge, 2013; Philbeck & O’Leary, 2005), thus it is possible that cue integration can occur during the outbound path when both cues are available. Such an effect would contaminate single-cue variances by reducing them, and subsequently underestimate optimal variances. To address this issue, we systematically manipulated the number of cues available during the outbound path, producing two path conditions: both-present and one-present. We analyzed response accuracy, response variability, and observed vs. predicted visual cue weights.

Across path conditions, mean response locations were found to be most accurate in the combination condition than in either of the single-cue conditions. Furthermore, single-cue

conditions showed similar response accuracy. Although the MLE model of cue integration does not make specific predictions about response accuracy (Cheng et al., 2007), it is not surprising that accessibility to more cues provides more accurate representations of spatial locations. For instance, Mou and Spetch (2013) showed that human observers were better able to detect changes in spatial locations of objects when they had access to visual and body-based cues than either one alone. However, their experiments employed a two-alternative forced-choice task, whereas participants in the current experiment made continuous bivariate responses.

Collapsing across path conditions, response variability showed significant reduction when participants had access to both visual and self-motion cues as opposed to one or the other. Participants were more precise when estimating the target location in the combination and conflict conditions than in either of the single-cue conditions. Furthermore, response variabilities in the combination and the conflict conditions were consistent with optimal cue integration ($ps > .050$ and $BFs > 3.00$) as predicted by performance in the single-cue conditions. We did not find an interaction between path conditions and cue conditions, suggesting that participants were able to optimally integrate visual and self-motion cues regardless of the number of cues available along the outbound path for single-cue trials. These results corroborate the hypothesis that cue integration does not occur along the outbound path, but instead occurs during decision making, validating both experimental methodologies.

Although navigators can make use of a remembered cue to reduce uncertainty about spatial location, even when the cue is no longer available (Kalia, Schrater, & Legge, 2013; Philbeck & O’Leary, 2005), we did not find evidence of cue integration along the outbound path in single-cue conditions. If participants were integrating cues along the outbound path, such that response variability was reduced in single-cue trials, response variability in dual-cue conditions should have been suboptimal, as optimal predictions would have been contaminated by the utilization of visual cues during self-motion trials and vice versa. It remains possible that cue integration did occur along the outbound path, but that this integrated representation was not utilized during the return path. Although Kalia, Schrater, and Legge observed reduced response precision both when the stopping location was perceived as consistent and inconsistent with the remembered landmark, it is possible that integration of remembered landmarks only enhances self-motion guided behavior when the traveled path places the navigator near the remembered landmark. Monitoring of remembered landmark locations might only be relevant to the navigator when the path runs tangentially to a landmark as the navigator might need to avoid collision. Although the authors employed trials in which participants were stopped before reaching the landmark, response variability was not compared across stopping positions. It is also possible that remembered landmarks

are only integrated with self-motion information when participants are explicitly told to utilize such cues. In Philbeck and O’Leary’s study, participants were told to “pay attention to the nearest [landmark] as they walked by it.” Reduction in response variability may not have been observed otherwise.

Mean observed cue weights were consistent with optimal cue weights in both path conditions. However, the linear relationship between observed and optimal cue weights revealed suboptimal weighting, as indicated by slope and intercept values less than one and greater than zero, respectively. Suboptimal linear relationships between observed and optimal cue weights have been demonstrated before (Chen et al., 2017). Similar to Chen et al.’s results, participants appeared to be biased toward the midpoint of the two cues during dual-cue trials (both cues should have indicated the same target location in combination trials but variability in sensory-perceptual systems could have produced disparate estimates even in the combination conditions). Their interpretation of such results was that participants were implementing a cost function that weighted errors of estimation more near the midpoint of the two cues (see Berger, 1985; Robert, 2007). This behavior could reflect sensitivity to the fact that large errors of navigation are possible when a cue with high relative reliability is used; a cue can have high relative reliability and still be poor in absolute terms because the other cues have even worse reliability (see Eq. 2). In such cases, navigators may be unwilling to trust either cue fully. We agree that a potential way to account for such deviations from optimality may be to employ a complete Bayesian model that incorporates cost functions.

Conclusions

The current experiment confirmed the validity of two methods of measuring cue integration in spatial navigation. Specifically, optimal cue integration was observed when participants had access to both visual and self-motion cues along the outbound path for single-cue trials, suggesting that the cue eliminated prior to the return path was not contaminating predicted optimal performance. Generally, patterns of cue performance across path conditions were similar given the lack of interaction between the path conditions. We also observed similar patterns of cue weighting across the path conditions, with mean observed weights consistent with optimal weights. However, participants showed deviations from optimality when regressing observed weights onto optimal weights. Although observed cue weights shared a linear relationship with optimal weights, the relationship was suboptimal, a pattern that has been observed before. There are other methodological approaches (e.g., path retracing vs. homing), however, that have not been systematically compared, and researchers

should take these different approaches into consideration when designing cue integration experiments in navigation.

Acknowledgements This research was supported in part by National Science Foundation Grant 1526448.

Open Practices Statement All data for the current experiment are available at <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/WI65ZU>.

Please e-mail the corresponding author to request experimental materials.

References

- Alais, D. & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257–262.
- Bates, S. L., & Wolbers, T. (2014). How cognitive aging affects multi-sensory integration of navigational cues. *Neurobiology of Aging*, 35(12), 2761–2769.
- Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Optical Society of America*, 20(7), 1391–1397.
- Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis* (2nd ed.). New York, NY: Springer-Verlag.
- Biegler, R. & Morris, R. G. M. (1999). Blocking in the spatial domain with arrays of discrete landmarks. *Journal of Experimental Psychology*, 25(3) 334–351.
- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effect of video game playing on attention, memory, and executive control. *Acta Psychologica*, 129 387–398.
- Butler, J. S., Smith, S. T., Campos, J. L., & Bühlhoff, H. H. (2010). Bayesian integration of visual and vestibular signals for heading. *Journal of Vision*, 10(11), 23. 1–13.
- Chamizo, V. D. (2003). Acquisition of knowledge about spatial location: Assessing the generality of the mechanism of learning. *The Quarterly Journal of Experimental Psychology*, 56B(1), 102–113.
- Chamizo, V. D., Sterio, D., & Mackintosh, N. J. (1985). Blocking and overshadowing between intra-maze and extra-maze cues: A test of the independence of locale and guidance learning. *The Quarterly Journal of Experimental Psychology*, 37B, 235–253.
- Chen, X., McNamara, T. P., Kelly, J. W., & Wolbers, T. (2017). Cue combination in human spatial navigation. *Cognitive Psychology*, 95, 105–144.
- Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian integration of spatial information. *Psychological Bulletin*, 133(4), 625–637.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433.
- Friedmann, A., Ludvig, E. A., & Legge, E. L. G. (2013). Bayesian combination of two-dimensional location estimates. *Behavior Research*, 45, 98–107. <https://doi.org/10.3758/s13428-012-0241-x>
- Frissen, I., Campos, J. L., Souman, J. L., & Ernst, M. O. (2011). Integration of vestibular and proprioceptive signals for spatial updating. *Experimental Brain Research*, 212(2), 163–176.
- Girshick, A. R. & Banks, M. S. (2009). Probabilistic combination of slant information: Weighted averaging and robustness as optimal percepts. *Journal of Vision*, 9(9):8, 1–20.
- Hamilton, D. A. & Sutherland, R. J. (1999). Blocking in human place learning: Evidence from virtual navigation. *Psychobiology*, 27(4), 453–461.

- Hardt, O., Hupbach, A., & Nadel, L. (2009). Factors moderating blocking in human place learning: The role of task instructions. *Learning and Behavior*, 37(1), 42–59. <https://doi.org/10.3758/LB.37.1.42>
- Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, 4(12), 967–992. <https://doi.org/10.1167/4.12.1>
- Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Research*, 39, 3621–3629.
- Jarosz, A. F. & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting bayes factors. *Journal of Problem Solving*, 7, 2–9.
- Kalia, A. A., Schrater, P. R., & Legge, G. E. (2013). Combining path integration and remembered landmarks when navigating without vision. *PLoS ONE* 8(9): e72170. <https://doi.org/10.1371/journal.pone.0072170>
- Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, Distinctions, and Interconnections. C. Freksa, C. Habel, & K. F. Wender (Eds.) *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge*, 1–17.
- Mou, W., & Spetch, M. L. (2013). Object location memory: Integration and competition between multiple context objects but not between observers' body and context objects. *Cognition*, 126(2), 181–197.
- Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology*, 18, 689–693.
- Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research* 43, 2451–2468.
- Petrini, K., Caradonna, A., Foster, C., Burgess, N., & Nardini, M. (2016). How vision and self-motion combine or compete during path reproduction changes with age. *Scientific Reports*, 6, 29163.
- Philbeck, J. W. & O'Leary, S. (2005). Remembered landmarks enhance the precision of path integration. *Psicológica*, 26, 7–24.
- Ratliff, K. R., & Newcombe, N. S. (2008). Reorienting when cues conflict: Evidence for an adaptive combination view. *Psychological Science*, 19, 1301–1307.
- Robert, C. P. (2007). *The Bayesian choice* (2nd ed.). New York, NY: Springer.
- Rohde, M., van Dam, L. C. J., & Ernst, M. O. (2016). Statistically optimal multisensory cue integration: A practical tutorial. *Multisensory Research*, 29, 279–317.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237.
- Sjölund, L. A., Kelly, J. W., & McNamara, T. P. (2018). Optimal combination of environmental cues and path integration during navigation. *Memory & Cognition*, 46, 89–99. <https://doi.org/10.3758/s13421-017-0747-7>.
- Tcheang, L., Bulthoff, H. H., & Burgess, N. (2011). Visual influence on path integration in darkness indicates a multimodal representation of large-scale space. *Proceedings of the National Academy of Sciences of the United States of America*, 108(3), 1152–1157.
- Twyman, A. D., Holden, M. P., & Newcombe, N. S. (2018). First direct evidence of cue integration in reorientation: A new paradigm. *Cognitive Science*, 42, 923–936.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352–402.
- Xu, Y., Regier, T., & Newcombe, N. S. (2017). An adaptive cue combination model of human spatial orientation. *Cognition*, 163, 56–66.
- Zhang, L., Mou, W., Lei, X., & Du, Y. (2019, December 2). Cue Combination Used to Update the Navigator's Self-Localization, Not the Home Location. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <https://doi.org/10.1037/xlm0000794>
- Zhao, M., & Warren, W. H. (2015a). Environmental stability modulates the role of path integration in human navigation. *Cognition*, 142, 96–109.
- Zhao, M., & Warren, W. H. (2015b). How you get there from here: Interaction of visual landmarks and path integration in human navigation. *Psychological Science*, 26(6), 915–924.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.