Using virtual reality to assess dynamic self-motion and landmark cues for spatial updating in children and adults

Erica M. Barhorst-Cates[1] · Jessica Stoker[2] · Jeanine K. Stefanucci[2] · Sarah H. Creem-Regehr[2]

Abstract

The relative contribution of different sources of information for *spatial updating* – keeping track of one's position in an environment – has been highly debated. Further, children and adults may differ in their reliance on visual versus body-based information for spatial updating. In two experiments, we tested children (age 10–12 years) and young adult participants on a virtual point-to-origin task that varied the types of self-motion information available for translation: full-dynamic (walking), visual-dynamic (controller induced), and no-dynamic (teleporting). In Experiment 1, participants completed the three conditions in an indoor virtual environment with visual landmark cues. Adults were more accurate in the full- and visual-dynamic conditions (which did not differ from each other) compared to the no-dynamic condition. In contrast, children were most accurate in the visual-dynamic condition and also least accurate in the no-dynamic condition. Adults outperformed children in all conditions. In Experiment 2, we removed the potential for relying on visual landmarks by running the same paradigm in an outdoor virtual environment with no geometrical room cues. As expected, adults' errors increased in all conditions, but performance was still relatively worse in teleporting. Surprisingly, children showed overall similar accuracy and patterns across locomotion conditions to adults. Together, the results support the importance of dynamic translation information (either visual or body-based) for spatial updating across both age groups, but suggest children may be more reliant on visual information than adults.

Keywords Spatial updating · Visual and body-based · Development · Virtual reality

Introduction

Virtual reality (VR) is increasingly used in research, medical treatment, education, and entertainment. Despite the widespread popularity of VR and its far-reaching applications, little is known about individual differences in how people interact with space in virtual worlds. Even fewer studies have focused on developmental differences, despite children being one of the primary end-users of VR. Interacting with a virtual world requires spatial updating, or the ability of navigators to keep track of their self-location relative to the environment as they move. Both *piloting* – using visual landmark cues, and *path integration* – using self-motion cues, are involved in spatial

updating, making navigation a complicated task that involves integrating multiple sources of information. Recent data demonstrate that adult females show comparable performance with different types of self-motion information required for spatial updating in a landmark-filled virtual environment (Barhorst-Cates et al., 2020), but are impaired without self-motion information for translation (when "teleporting"). Whether this phenomenon is generalizable to landmark-free environments or is also present in children is unknown. The present study assessed spatial updating in young adults and in 10- to 12-year-old children using VR with a goal to determine (1) which sources of dynamic self-motion information are necessary and sufficient for spatial updating, (2) whether children rely on different information than adults, and (3) whether effects generalize across environments with differences in landmark cues.

Path integration refers to spatial updating of one's position by integrating both translations and rotations for a traversed path (Chrastil & Warren, 2012). While humans generally perform above chance in path integration tasks, performance is not optimal, even with many available cues (Chrastil & Warren, 2013). Error accumulates with distance traveled or

✉ Erica M. Barhorst-Cates
  BarhorsE@einstein.edu

[1] Moss Rehabilitation Research Institute, 60 Township Line Road, Elkins Park, PA, USA

[2] Psychology Department, University of Utah, Salt Lake City, UT, USA

number of turns as a path is traversed (Fujita et al., 1990; Lappe et al., 2007). The method by which translation information is encoded (i.e., the self-motion information) may thus affect the extent of error. This self-motion information arises from both visual (optic flow, binocular disparity, and surface texture of the environment) and body-based sources (efferent motor commands, proprioceptive and vestibular feedback; Chrastil & Warren, 2013).

In adults, it has been shown that spatial updating through physical movement is largely an automatic process (Rieser, 1989; Farrell & Thomson, 1998). Rotation information is particularly important for spatial updating (Chance et al., 1998; Klatzky et al., 1998; Wraga et al., 2004; but see Riecke et al., 2007), whereas the importance of physical translation is less understood. Some find that actual body translation is necessary, especially in complex large environments (Ruddle & Lessels, 2006; Ruddle et al., 2011), whereas others show that real bodily translation is not necessary (Chance et al., 1998) as long as there are some translational motion cues (e.g., Nguyen-Vo et al., 2019). Further, the difficulty in spatial updating after imagined movement has been attributed more to rotation than to translation (Presson & Montello, 1994). Research on spatial updating without vision shows good accuracy in spatial updating with low numbers of turns, supporting reliance on body-based information when vision is not available (e.g., Loomis, Klatzky, Golledge, & Philbeck, 1999; Petrini et al., 2016; Philbeck & Loomis, 1997; Rieser et al., 1986; Thomson, 1983). However, humans can perceive self-motion, turn angles, and distances from optic flow (Bremmer & Lappe, 1999; Hettinger, 2002; Warren et al., 2001) even when no body-based information is available. Some studies suggest that realistic and contextually rich three-dimensional models induce a stronger sense of self-motion perception (Riecke, Heyde, & Bülthoff, 2005; Trutoiu et al., 2009) and promote spatial updating (Riecke et al., 2007; Riecke, Schulte-Pelkum, Caniard, & Bulthoff, 2005). We have recently demonstrated that young adult females can perform equally well on a point-to-origin task in a visually rich environment with either visual-only or visual and body-based translation information, as long as some self-motion information is present (Barhorst-Cates et al., 2020). When young adults translate in a virtual world with a "teleporting" method that eliminates self-motion information, point-to-origin and triangle completion performance is impaired (Barhorst-Cates et al., 2020; Cherep et al., 2020). Still others have argued that visual and body-based sources of information are equally useful (Chrastil et al., 2019), perhaps because they share similar neurological codes (Huffman & Ekstrom, 2019). As such, it is unclear whether body-based *translation* information is necessary for spatial updating in adults, particularly in environments that do not have sufficient visual information (e.g., landmarks) for using piloting strategies.

Children may depend on body-based self-motion information for spatial updating more than adults because children require overt movement to understand spatial concepts. Locomotor status (Foreman et al., 1989; Yan et al., 1998) and balance ability (Jansen & Heil, 2010) are predictive of spatial cognitive tasks in children above and beyond other factors such as general intelligence and executive functioning (Frick & Möhring, 2016; Gabbard et al., 2012). These findings highlight the role of movement experience for children's acquisition of spatial knowledge. Dependence on body-based information may decrease by the age of approximately 10–11 years (Lehnung et al., 1998; Lehnung et al., 2003), but recent research suggests that children in this age group still perform best when they can use body-based information only (Petrini et al., 2016). Petrini et al. (2016) argue that 10- to 11-year-old children have difficulty ignoring visual information during spatial updating tasks, even when it is irrelevant, whereas adults can fluctuate between and ignore certain cues as needed. This age difference may be explained by differences in sensory calibration (Gori et al., 2008; Newell & Wade, 2018) or multisensory integration (Downing et al., 2015) of visual and body-based cues.

There are open questions about the use of body-based versus visual cues in successful spatial updating at different ages. Research in adults suggests that either visual or body-based information for translation is sufficient for comparable performance in a spatial updating task, in a full-cue virtual environment with visual landmarks. Ten- to 12-year-old children may or may not rely on either visual or body-based information for translation, given their difficulty in ignoring conflicting sensory-motor information (Petrini et al., 2016). Furthermore, visual environmental cues may play a role in the relative influence of body-based and visual cues for self-motion for both age groups. We present two experiments that test the role of dynamic self-motion information in two age-defined samples by manipulating the self-motion information available for path integration with three different virtual locomotion methods and the visual landmark cues provided by the environmental context itself.

Spatial updating in VR has traditionally used joystick-controlled or video translation methods that provide visual-only information for spatial updating, but a newly developed interactive technique termed teleporting allows for removal of all dynamic translational information (e.g., Coomer et al., 2018). Teleporting involves pointing a controller to and then selecting a location in a virtual environment to instantaneously arrive without receiving either visual or body-based information for self-motion. While useful as a method for quickly traversing large distances and reducing motion sickness, it negatively affects spatial updating (Cherep et al., 2020), especially in large environments without visual landmark cues. The reason for this deficit is not fully understood, although it is likely that the loss of translational self-motion information

contributes to error (Barhorst-Cates et al., 2020). Paris et al. (2019) have demonstrated that locomotion methods that provide continuous motion information are more effective for accurate spatial updating than those that provide discrete motion (such as in teleporting).

As a secondary analysis, we also assessed two individual differences that may contribute to spatial updating performance that were motivated by the literature. First, we considered the role of motor control, as locomotion status is a strong predictor of navigation ability in children (Foreman et al., 1989; Yan et al., 1998). To operationalize motor control, we included a measure of balance ability that is predictive of performance on spatial-cognitive tasks (Frick & Möhring, 2016; Jansen & Heil, 2010), expecting that better balance ability would relate to reduced errors in spatial updating. Second, we considered the role of small-scale spatial abilities by including a standard mental rotation task, which has been implicated in spatial updating in larger-scale navigation tasks (e.g., Hegarty et al., 2006; Ruginski et al., 2019). Based on the prior work, we predicted that higher mental rotation scores would relate to higher accuracy on the spatial updating task.

General method

Participants

Our sample size goal was based on prior navigation research that has used within-subjects manipulations with samples of 15 children and 18 adults (Petrini et al., 2016) or 14 children and 17 adults (Nardini et al., 2008). Other studies detecting age differences in path integration have used a range of from 18–20 children and up to 40 adults (Smith et al., 2013). As such, we aimed to recruit at least 25 individuals per age group in each study. All participants had normal or corrected-to-normal vision and could walk without impairment. We contacted local after-school programs and camps, and fifth-grade teachers at elementary schools in the Salt Lake Valley, used word-of-mouth recruitment, and posted flyers around campus to recruit a large sample of children aged 10–12 years. These ages can accurately integrate cues, but may be biased by visual information (Petrini et al., 2016). Thirty 10- to 12-year-old participants completed Experiment 1 (M age = 10.97 years, SD = 0.80; 13 female). We also recruited 17 additional children who were outside the target age range and included them in an exploratory analysis. For Experiment 2, the main child sample consisted of 25 children aged 10–12 (seven female) and five 9-year-old children (three males) for the exploratory age analysis. For all child participants, we obtained written informed parental consent and participant assent with procedures approved by the University of Utah Institutional Review Board. Child participants were compensated with $10 for their time.

We recruited young adults from the psychology department participant pool. Forty-one young adults (26 female, M age = 21 years, SD = 5.14) completed Experiment 1. 32 young adults (18 female, M age = 22, SD = 5.7) completed Experiment 2. One adult participant (a male) dropped out of the experiment due to motion sickness. All adult participants gave written informed consent with procedures approved by the University of Utah Institutional Review Board and received partial course credit for participation.

Materials

The virtual space for the virtual point-to-origin task in Experiment 1 was a model of a real lab space at the University of Utah built with Unity (version 2018.2.12f1). The real room was 28 × 38 ft. The environment was the same as that used in Barhorst-Cates et al. (2020). The geometry, coloring, and texturing on the walls and floors of the virtual space matched those of the real lab but the relative horizontal dimension was elongated to allow for more confidence for the participant when walking near walls. The landmark cues in this environment included a door, four corners, two textured walls, and six mounted cameras (see Fig. 1). The virtual environment for Experiment 2 was modified to be a boundless grassy field with blue sky and a visible horizon but no other visual cues (see Experiment 2 below). We also changed the colors of the poles to be black and orange to contrast the grassy plain and to remedy concerns about red-green color blindness. Participants did not view the real room before seeing the virtual room.

The head-mounted display (HMD) was the HTC Vive Pro, which has a field-of-view of 110° and a resolution of 1,440 × 1,600 pixels per eye (www.vive.com/us/product/vive-pro/). Interpupillary distance (IPD) was set to 64 mm for adults and 61 mm (the lowest possible setting) for children. We used four Lighthouse motion trackers positioned in an approximate 4 × 4 m square. All participants used the wireless HMD system, except for six children and two adults in Experiment 1, who had to use a corded HMD due to technical difficulties.

Individual differences measures

Participants balanced as long as possible standing on a balance pad (ProSource, 15.5-in. L × 13-in. W × 2.5-in. H) with one leg, first with eyes closed while wearing a Mindfold blindfold, and then with eyes open (see Frick & Möhring, 2016). If participants lost balance within 1 s of lifting their foot, they tried again. Participants also completed the English version of the short Mental Rotation task developed by De Beni et al. (2014), an abbreviated 10-item version of the paper-and-pencil task developed by Peters et al. (1995). Participants viewed a target figure on the left and four response options on the right, and had to select which of the two items on the
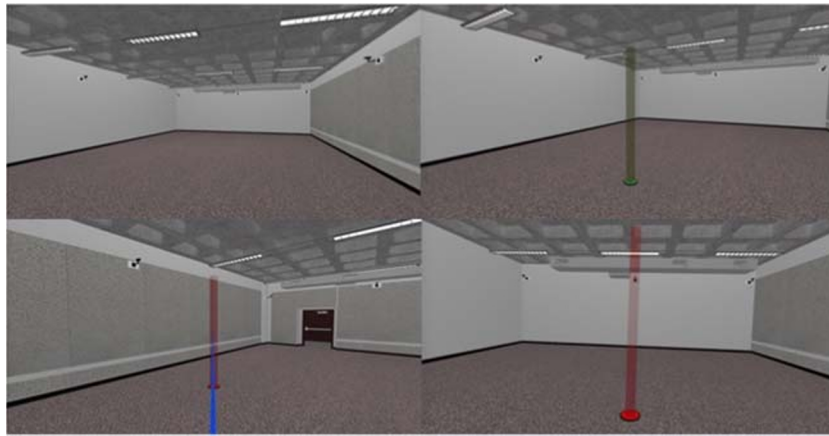
Fig. 1 One trial of the virtual point-to-origin task in Experiment 1. Participants first appeared at a random location in the room (**top left**), then located the green pole (**top right**). They locomoted to the green pole then looked around for the first red pole (**bottom left**), receiving feedback for the first 0.5 m of movement. Finally, they located the second red pole (**bottom right**). Upon reaching the second red pole, the screen turned black, and participants made their response by facing back to the green pole

right represented a rotated version of the target figure. A point was only given when both correct answers were selected, for a maximum possible score of 10.

Child participants in both experiments filled out a spatial activities questionnaire that asked about involvement in video games, dance, carpentry, three-dimensional painting or drawing, and graphic design and a demographics questionnaire. In Experiment 1, adult participants completed an extensive survey including a dance experience questionnaire, spatial activities survey, video game questionnaire, general demographics survey, and the Vividness of Movement Imagery Questionnaire (Roberts et al., 2008). Due to the lack of relationships between these questionnaires and point-to-origin task performance in Experiment 1 (see Results), in Experiment 2 adults completed only the children's abbreviated spatial activities and demographics questionnaire.

Procedure

Upon arriving at the lab, adult participants gave informed consent, parents signed parental permission forms, and children filled out written assent forms. All participants then completed the eyes closed balance task, followed by the mental rotation task, then the eyes open balance task. Then the experimenter explained the point-to-origin task in the real world, demonstrating it with three cones. Finally, the experimenter described motion sickness and encouraged participants to inform the researchers if they felt sick. Participants were then blindfolded and guided into the room where the experimenter placed and adjusted the HMD on the participant's head. Participants held a Vive controller in each hand. The experiment began with practice pointing trials in which participants were instructed to turn their bodies to face toward objects and received feedback via a blue feedback line protruding at face level.

Locomotion condition order was randomized and counterbalanced across participants. Participants were instructed in all conditions to keep their head facing in the same direction of their body when turning, which was emphasized with a blue feedback line that disappeared after 0.5 m of translation. The locomotion conditions were the same as those used in Barhorst-Cates et al. (2020) but renamed to more clearly define the visual and body-based information provided in each condition. The *full-dynamic* condition consisted of real walking at a comfortable speed. The *visual-dynamic* motion was executed by pulling the trigger on the Vive controller with the index finger of the dominant hand. We disabled lateral visual movement during each leg of forward visual translation to the target to ensure that participants took a direct path to the target and to possibly reduce the chances of motion sickness, particularly with the unknown effects on children. Trutoiu et al. (2009) previously suggested that left-right visual motion inducing perception of self-motion with a large screen display was a factor in self-reported simulator sickness. Physical head rotation was the same as in the full-dynamic condition, but if the participant looked away from the target, the visual translation was halted. Movement speed jumped to 0.5 m/s and only progressed if participants were looking directly at the target. There was no smooth acceleration or deceleration and translation halted if participants looked away from the target. In the *no-dynamic* condition, participants "teleported" by using the dominant hand controller to point to the goal position, pressing down the thumbpad to view the "arc" that designated trajectory, and then releasing the thumbpad to be immediately relocated. The direction of teleporting was determined by the pointing direction of the Vive controller. Participants selected a distance by moving the controller closer to or farther away from their torso to extend or shrink the arc before selecting the end location. Teleporting was restricted to target locations only.

In all locomotion conditions, participants physically turned their heads and bodies in the environment. For each trial, participants located the starting position (a green or black circle on the ground with a semi-transparent green or black pole) and locomoted there. The experimenter told the participant to remember this location. Participants then traveled to the first red or orange pole, then located and traveled to a second red or orange pole. After each movement, the poles disappeared and a beep sounded through the headphones upon arrival. After reaching the second red or orange pole, the screen turned black and participants faced back to the starting location as if they were going to walk there and verbalized out loud to the experimenter "ready." The experimenter recorded this position on the computer and then asked the participant to take one step forward and recorded the position again to emphasize the need to estimate the direction back to the start. Angles were based on the orientation of the participant's HMD. We used only the pre-step angle as our measure of pointing. While the screen was still black, the experimenter led the participant on a circuitous random path to a new position before beginning the next trial. The start pole changed locations in the virtual environment on every trial except in the Full-Dynamic condition, where physical room constraints required the pole to be in a central location on all trials (so that participants did not collide with walls or objects when walking). To decrease the likelihood of participants learning the real-world location of the start pole that was always in the same position in the real room, the virtual room was rotated on every trial in all three conditions so that it appeared to be in a different location in the virtual room. We also projected white noise through the headphones built into the HMD. Participants completed three practice trials and eight experimental trials in each locomotion method. Path leg length ranged from 1.5 to 2.5 m and turning angle ranged between 45° and 150°, with varying left and right turns (see Table 1 for

specific trial information). Between each condition, participants were given the opportunity to remove the HMD and take a short break before beginning the next condition. If participants wanted a break, we asked them to close their eyes while we removed the HMD, then placed a blindfold over the participant's eyes and led them out of the room. To ensure safety in the HMD, one experimenter stood near to the participant at all times. Upon completing the three conditions, participants completed the questionnaires.

Design and data analyses

Both experiments used a 2 × 3 mixed factorial design with a between-subjects individual differences factor (age group: child and adult) and a within-subjects repeated-measures manipulated self-motion cue factor (locomotion method: full-dynamic, visual-dynamic, no-dynamic). Linear mixed effects modeling analyses were performed using the *lme4* and *lmerTest* packages in R version 3.6.1. Linear mixed effects modeling is a flexible analysis approach that allows for imbalanced (missing) data and the inclusion of multiple random effects. We included random effects of participant and trial in all models. We ran a series of models with each of the dependent variables to assess changes in model fit (likelihood ratio test) with the addition of locomotion condition, age group, and the condition × age group interactions factors. Because we planned *a priori* to assess differences in the patterns of results for the different age groups, we also tested the effect of condition separately for adults and children and conducted planned post hoc contrasts to examine differences between conditions within each age group. We used the *emmeans* package with a Tukey adjustment for multiple comparisons. There is no consensus on appropriate measures of effect size for mixed effects models (Peugh, 2010). We report

Table 1　Trial information for virtual point-to-origin task

Trial type	Trial number	Leg 1 length	Turning angle between legs 1 and 2	Leg 2 length	Correct response angle
Practice	Practice 1	2 m	90° right	2 m	-
Practice	Practice 2	2 m	90° left	2 m	-
Practice	Practice 3	1.5 m	45° right	2.5 m	-
Experimental	1	1.5 m	50° right	1.5 m	155° right
Experimental	2	1.5 m	50° left	2.5 m	161.65° left
Experimental	3	2.5 m	80° right	1.5 m	128.15° right
Experimental	4	2.5 m	80° left	2.5 m	140° left
Experimental	5	1.5 m	120° left	1.5 m	120° left
Experimental	6	2.5 m	120° right	1.5 m	96.59° right
Experimental	7	2.5 m	150° left	2.5 m	105° left
Experimental	8	1.5 m	150° right	2.5 m	148.02° right

Note. All participants completed all trials in each condition and the order of the experimental trials was randomized in each condition for each participant. Reproduced from Barhorst-Cates, Stefanucci, and Creem-Regehr (2020)

two indices of model fit – the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), where a lower number is better, and report Cohen's d as a measure of effect size based on the estimated means. We also report the standardized regression coefficients for the models with continuous predictors.

Experiment 1

In Experiment 1, we tested 10- to 12-year-old children and young adults' ability to perform spatial updating with and without dynamic body-based and visual information for self-motion in an indoor room virtual environment. We expected both adults and children to perform with the greatest accuracy in the full-dynamic condition because body-based and visual information were provided for both translation and rotation. We expected the second-best performance for adults and children in the visual-dynamic condition, because of the presence of visual self-motion cues. We expected the worst performance for both adults and children in the no-dynamic condition, because of the lack of translational information (both visual and body-based). In comparison to adults, we expected worse performance for children in all conditions. We also expected a greater detriment to performance in children compared to adults when any dynamic self-motion information was removed. Critically, the presence of visual landmarks could compensate for the decreased self-motion information.

Results

Angular error

Absolute (unsigned) angular error was calculated as the smallest difference between the participant's heading direction at the final pole and their heading direction after making the response turn, but before taking the step. We performed a square-root transformation on the data to account for non-normality before modeling. Across the sample from age 10 years to adulthood, we observed significant effects of Condition ($\chi^2(1) = 37.30$, $p < .001$, $AIC = 7639.4$, $BIC = 7672.0$) and Age Group ($\chi^2(1) = 25.78$, $p < .001$, $AIC = 7615.6$, $BIC = 7653.7$). The addition of these factors improved model fit compared to a baseline intercept-only model ($AIC = 7672.7$, $BIC = 7694.4$). The Condition × Age Group interaction was not significant ($\chi^2(2) = 2.70$, $p = .3$). Post hoc contrasts comparing the three conditions revealed that, surprisingly, error in Full-Dynamic ($M = 28.1$, $SE = 2.0$) was significantly higher than error in Visual-Dynamic ($M = 23.5$, $SE = 1.8$, $t = 3.47$, $p = .002$, $d = 0.10$) but lower than error in No-Dynamic ($M = 31.9$, $SE = 2.11$, $t = -2.64$, $p = .02$, $d = 0.08$). Error in Visual-Dynamic was significantly lower than error in No-Dynamic ($t = -6.12$, $p < .0001$, $d = 0.18$). Across all conditions, adults ($M =$

21.6, $SE = 1.81$) significantly outperformed children ($M = 34.6$, $SE = 2.51$, $d = 1.02$, see Fig. 2).

For adults only, there was a significant effect of Condition ($\chi^2(2) = 17.31$, $p < .0002$, $AIC = 4199.6$, $BIC = 4229.0$). Post hoc pairwise contrasts revealed that error in Full-Dynamic ($M = 21.4$, $SE = 1.86$) did not significantly differ from error in Visual-Dynamic ($M = 18.8$, $SE = 1.74$, $t = 1.88$, $p = .1$, $d = 0.08$) for adults. No-Dynamic error ($M = 24.8$, $SE = 2.0$) was marginally higher than Full-Dynamic ($t = -2.27$, $p = .06$, $d = 0.10$) and significantly higher than Visual-Dynamic ($t = -4.17$, $p = .0001$, $d = 0.18$). For children only, there was a significant effect of Condition ($\chi^2(2) = 20.82$, $p < .0002$, $AIC = 3383.0$, $BIC = 3410.5$). Surprisingly, errors in Full-Dynamic ($M = 35.9$, $SE = 3.33$) were significantly higher than errors in Visual-Dynamic ($M = 28.3$, $SE = 2.96$, $t = 3.01$, $p = .008$, $d = 0.16$), but did not differ significantly from No-Dynamic ($M = 40.0$, $SE = 3.51$, $t = -1.50$, $p = .3$, $d = 0.08$). No-Dynamic errors were higher than Visual-Dynamic ($t = -4.51$, $p < .0001$, $d = 0.23$) for children. Finally, we examined the age-group effect separately for each of the three conditions. Adults outperformed children in Full-Dynamic ($\chi^2(1) = 33.39$, $p < .0001$, $AIC = 2407.4$, $BIC = 2429.1$), Visual-Dynamic ($\chi^2(1) = 10.33$, $p = .001$, $AIC = 2523.8$, $BIC = 2545.5$), and No-Dynamic ($\chi^2(1) = 14.41$, $p = .0001$, $AIC = 2681.4$, $BIC = 2703.1$).

Together, these results suggest that locomotion method affects point-to-origin accuracy with significant improvements in accuracy from the age of 10–12 years to adulthood.[1] As predicted, both adults and children performed the worst with no-dynamic self-motion cues. Moreover, contrary to adults, who showed similar performance in Full-Dynamic and Visual-Dynamic, children performed with the highest accuracy in the Visual-Dynamic condition.

Response time

We also examined the time taken to make the final turn response. We expected that longer response time would reflect more cognitive processing, which could indicate difficulty. We log-transformed the response-time data. Across the full dataset, there was a significant effect of Condition ($\chi^2(2) = 30.32$, $p < .001$, $AIC = 7315.4$, $BIC = 7348.0$) but no effect of Age Group ($\chi^2(1) = 1.01$, $p = .3$) and no interaction ($\chi^2(2) = 2.81$, $p = .2$). The Condition factor improved model fit compared a baseline intercept-only model ($AIC = 7341.7$, $BIC = 7363.5$). We performed post hoc contrasts to query the Condition effect. RT

---

[1] We also measured age differences continuously across our full sample, including the children outside of our target age range of 10–12 years. Specifically, we recruited three 8-year-olds, ten 9-year-olds, eleven 11-year-olds, nine 12-year-olds, three 13-year-olds, and one 14-year-old. We tested the effect of Age by including Age as a continuous variable instead of Age Group in the mixed effects model. There was a significant effect of Age ($\chi^2(1) = 20.02$, $p < .001$), with error decreasing across conditions with increasing age ($B = -.08$, $\beta = -.08$, $p < .001$).
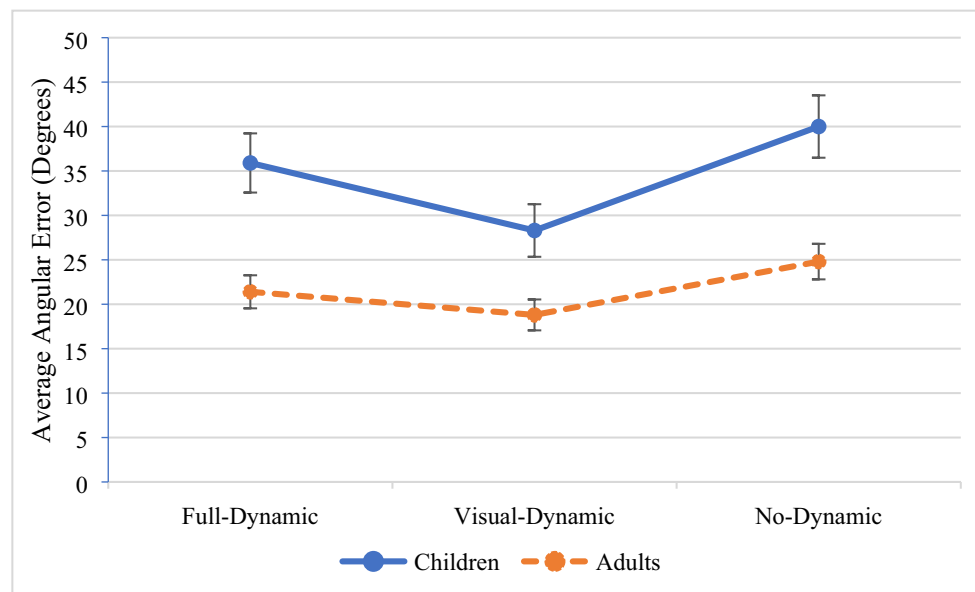
Fig. 2 Average error for adults and children in each locomotion condition. Error bars represent ± 1 standard error of the mean

after Full-Dynamic (M = 7.01, SE = .35) did not differ significantly from RT after Visual-Dynamic (M = 7.16, SE = .36, t = -1.39, p = .3, d = 0.02) but was significantly quicker than RT after No-Dynamic (M = 7.56, SE = .38, t = -5.01, p < .0001, d = 0.06). Visual-Dynamic RT was also significantly quicker than No-Dynamic (t = -3.62, p = .0008, d =0.05).

For adults only, there was a significant effect of Condition (χ^2(2) = 27.17, p < .001, AIC = 4138.8, BIC = 4168.1): Full-Dynamic response time (M = 7.05, SE = .45) was significantly faster than both Visual-Dynamic (M = 7.38, SE = .47, z = -2.46, p = .04, d = 0.04) and No-Dynamic (M = 7.75, SE = .49, z = -5.21, p < .001, d = 0.08). Visual-Dynamic was significantly faster than No-Dynamic (z = -2.77, p = .02, d = 0.04). For children only, there was a significant effect of Condition (χ^2(2) = 7.25, p = .03, AIC = 3185.0, BIC = 3212.5). Response time in Full-Dynamic (M = 6.97, SE = .34) was not significantly different than response time in Visual-Dynamic (M = 6.96, SE = .34, z = .07, p = .9, d = 0.002) but was marginally faster than No-Dynamic (M = 7.38, SE = .36, z = -2.28, p = .06, d = 0.08). Visual-Dynamic RT was faster than No-Dynamic (z = -2.35, p = .049, d = 0.08). Lastly, there were no age differences in RT for Full-Dynamic (χ^2(1) = .01, p = .9), Visual-Dynamic (χ^2(1) = 2.19, p = .1), or No-Dynamic (χ^2(1) = 1.44, p = .2). Overall, children took similar amounts of time to respond in the different conditions compared to adults, who showed increases in response time corresponding to reductions in self-motion translation information (see Fig. 3).

Balance and mental rotation

Because we observed no significant Condition × Age Group interaction, we dropped that term from the model in the

following analyses. Model comparison testing confirmed that the angular error model with balance time (eyes open) was a significantly better fit to the data than the model without (χ^2(1) = 5.77, p = .02, AIC = 7611.8 , BIC = 7655.3). Longer balance time predicted a decrease in error across all participants and conditions (B = -.002, β = -.05, p = .02). Balance time (eyes closed) did not improve model fit (χ^2(1) = 1.88, p = .2). The angular error model with mental rotation was also a better fit to the data than the model without (χ^2(1) = 6.81, p = .01, AIC = 7610.8, BIC = 7654.3), and higher mental rotation score predicted a decrease in average error across all participants and conditions (B = -.14, β = -.05, p = .009). These results suggest that better balance ability (at least with eyes open) and better mental rotation ability may contribute to higher accuracy in spatial updating. See Table 2 for means in each age group.

Surveys

Because our adults and children completed different surveys, we assessed the effects separately for each group. For adults, none of the Vividness of Movement Imagery subsections significantly improved the model fit (χ^2(1)s < 3, ps > .08). Spatial activities also did not improve model fit (χ^2(1) = .001, p = .9). Twenty-two of the adults reported video game play. For those individuals, gaming experience did not improve model fit (χ^2(1) = .03, p = .9). For children, there was no significant effect of spatial activities (χ^2(1) = 2.78, p = .1) or gaming (χ^2(1) = .58, p = .4). To look at differences between age groups in video game experience, we ran a one-way analysis of variance with age group predicting hours per week and found a significant effect F(1,69) = 16.86, p < .001. Children (M = 4.47, SD = 4.13) reported a greater number
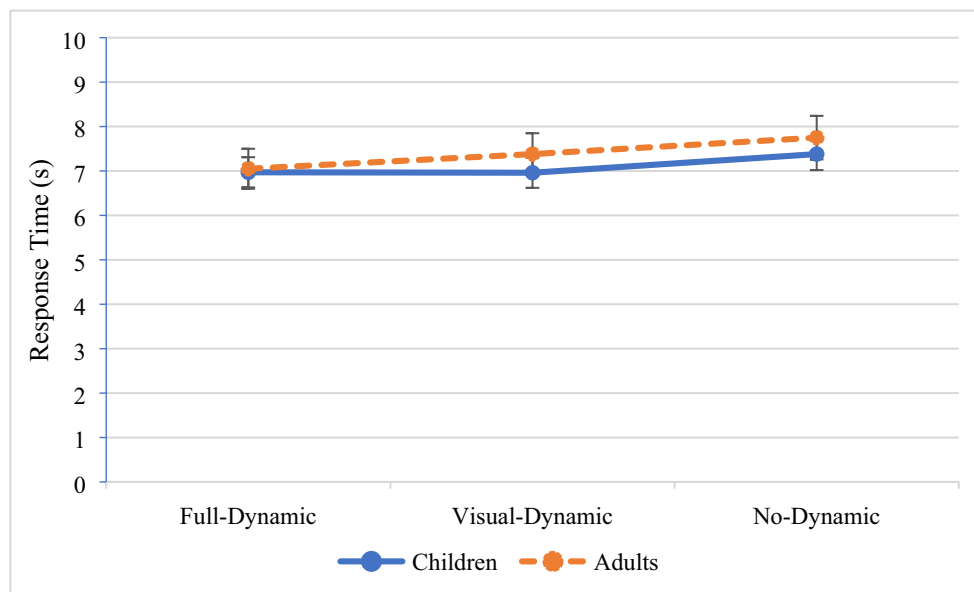
Fig. 3 Average response time for adults and children in each locomotion condition. Error bars represent ± 1 standard error of the mean

of hours of videogame play per week than adults (M = 1.39, SD = 2.1).

Taken together, the results partially supported our hypotheses for adults; performance with no dynamic information was worse than performance with visual dynamic or full dynamic information, which did not differ from each other. For children, error was highest in the condition with no dynamic information. However, children showed larger errors when both visual and body-based translation information was present (full dynamic) than in the visual-dynamic only condition. Overall, adults performed better than children in all conditions.

Experiment 2

In Experiment 1, the presence of visual room cues could have supported a reliance on piloting (use of visual landmark cues) that reduced the need for the use of self-motion information for spatial updating. Indeed, a large body of work reveals the importance of visual landmarks in spatial updating (e.g., Kalia et al., 2013; Kelly et al., 2008; Zhao & Warren, 2015),

potentially because of their greater reliability compared to path integration cues, which tend to be more error-prone. To further understand reliance on dynamic self-motion information for spatial updating in adults and children, we manipulated the environment in Experiment 2 to remove all visual landmark cues (similar to the environment used in Cherep et al., 2020) in order to eliminate piloting (see Fig. 4). The virtual environment for Experiment 2 was modified to be a boundless grassy field with blue sky and a visible horizon but no other visual cues. We used the same point-to-origin task with the same three locomotion methods in Experiment 1. We expected that both adults and children would now show the best performance in the full-dynamic condition, because the lack of visual landmark cues should shift their strategy to be more body-based (Zhao & Warren, 2015) and they would continue to show the greatest decrement in the no-dynamic (teleporting) condition. We expected to replicate the finding that adults would outperform children on all conditions. Although we did not compare directly across studies, we predicted higher mean errors in this environment compared to the room environment of Experiment 1 for both adults and children (Cherep et al., 2020). Finally, we again included

Table 2 Balance and mental rotation task averages

Experiment	Age group	Balance (eyes open) M (SD)	Balance (eyes closed) M (SD)	Mental rotation M (SD)
1	Adults	115 s (179.0)	5.98 s (4.5)	3.66 (2.3)
	Children	70.5 s (82.4)	4.20 s (2.9)	1.90 (1.4)
2	Adults	105 s (172.0)	4.80 s (4.3)	3.67 (2.0)
	Children	80.3 s (97.0)	5.20 s (5.0)	2.84 (2.1)

Note. Group means for balance and mental rotation scores in each experiment. The maximum possible score on the mental rotation task was 10
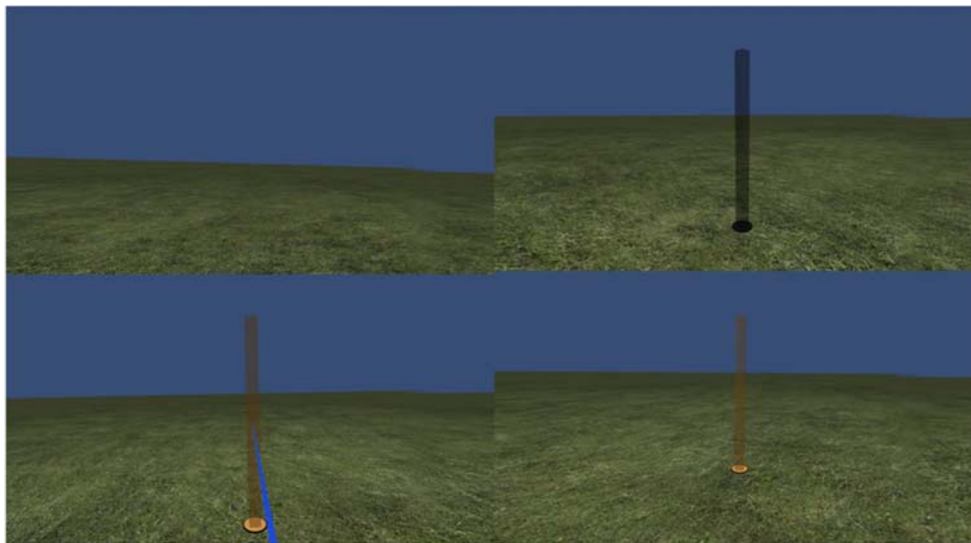
Fig. 4 Outdoor virtual environment and target poles used in Experiment 2

measures of balance ability and mental rotation, expecting for each task that better performance would relate to lower angular pointing errors. (e.g., Frick & Möhring, 2016; Ruginski et al., 2019), similar to Experiment 1.

Results

Angular error

Using the same mixed-effects models from Experiment 1, we tested effects of Condition, Age Group, and the Condition × Age Group interaction. There was a significant effect of Condition ($\chi^2(2) = 21.41$, $p < .001$, $AIC = 6031.2$, $BIC = 6062.3$) but not of Age Group ($\chi^2(1) = .23$, $p = .6$). There was also no Condition × Age Group interaction ($\chi^2(2) = 4.01$, $p = .1$). Condition improved the model fit compared to a baseline intercept-only model ($AIC = 6048.6$, $BIC = 6069.4$). Based on this significant effect of Condition, we performed post hoc contrasts and observed that error in Full-Dynamic ($M = 26.2$, $SE = 2.58$) did not differ from error in Visual-Dynamic ($M = 24.7$, $SE = 2.5$, $t = .93$, $p = .6$, $d = 0.03$) but was significantly lower than error in No-Dynamic ($M = 31.8$, $SE = 2.83$, $t = -3.36$, $p = .002$, $d = 0.10$). Visual-Dynamic error was also significantly lower than No-Dynamic ($t = -4.33$, $p < .0001$, $d = 0.12$) across the sample.

For adults only, there was a significant effect of Condition ($\chi^2(2) = 16.98$, $p = .0002$, $AIC = 3365.1$, $BIC = 3392.7$). Error in Full-Dynamic ($M = 23.7$, $SE = 2.69$) did not differ from error in Visual-Dynamic ($M = 24.9$, $SE = 2.72$, $t = -.58$, $p = .8$, $d = 0.03$) but was significantly lower than error in No-Dynamic ($M = 32.0$, $SE = 3.10$, $t = -3.81$, $p < .001$, $d = 0.18$). Visual-Dynamic error was also significantly lower than No-Dynamic ($t = -3.29$, $p = .003$, $d = 0.15$). For children only, there was a significant effect of Condition ($\chi^2(2) = 8.54$, $p = $

.01, $AIC = 2673.9$, $BIC = 2700.2$). Error in Full-Dynamic ($M = 28.7$, $SE = 3.61$) did not differ from error in Visual-Dynamic ($M = 24.5$, $SE = 3.32$, $t = 1.78$, $p = .2$, $d = 0.08$) or No-Dynamic ($M = 31.5$, $SE = 3.75$, $t = -1.10$, $p = .5$, $d = 0.05$). Visual-Dynamic error was significantly lower than No-Dynamic ($t = -2.91$, $p = .01$, $d = 0.13$). Finally, we wanted to test the effect of age group for each locomotion condition and observed no effect of Age Group for any of the conditions ($\chi^2(1)s < 1$, $ps > .09$).

These results replicate our finding from Experiment 1 for adults, such that adults performed worse when there was no self-motion information available compared to when there was visual only or both visual and body-based, even in an environment with no visual landmark cues. Our results also suggest that children's performance in this environment now more closely resembles adult performance. Children no longer show a statistically significant advantage for visual-only translation over both visual and body-based translation information. See Fig. 5 for the means for each age group in each condition.[2]

Response time

For response time, there was a significant effect of Condition ($\chi^2(2) = 18.37$, $p = .0001$, $AIC = 6103.9$, $BIC = 6135.0$), but no effect of Age Group ($\chi^2(1) = 1.15$, $p = .3$), and no interaction ($\chi^2(2) = 4.87$, $p = .09$). The Condition factor improved the model fit compared to a baseline intercept-only model ($AIC = 6118.3$, $BIC = 6139.0$). Post hoc contrasts revealed that RT in Full-Dynamic ($M = 8.07$, $SE = .5$) was significantly quicker

---

[2] We also wanted to examine age effects across the full sample from age 9 to 43 years, including the 9-year-olds' data that we collected as an exploratory measure. We ran the same mixed-effects model but with Age as a continuous variable instead of Age Group. Counter to our predictions, age did not predict performance.
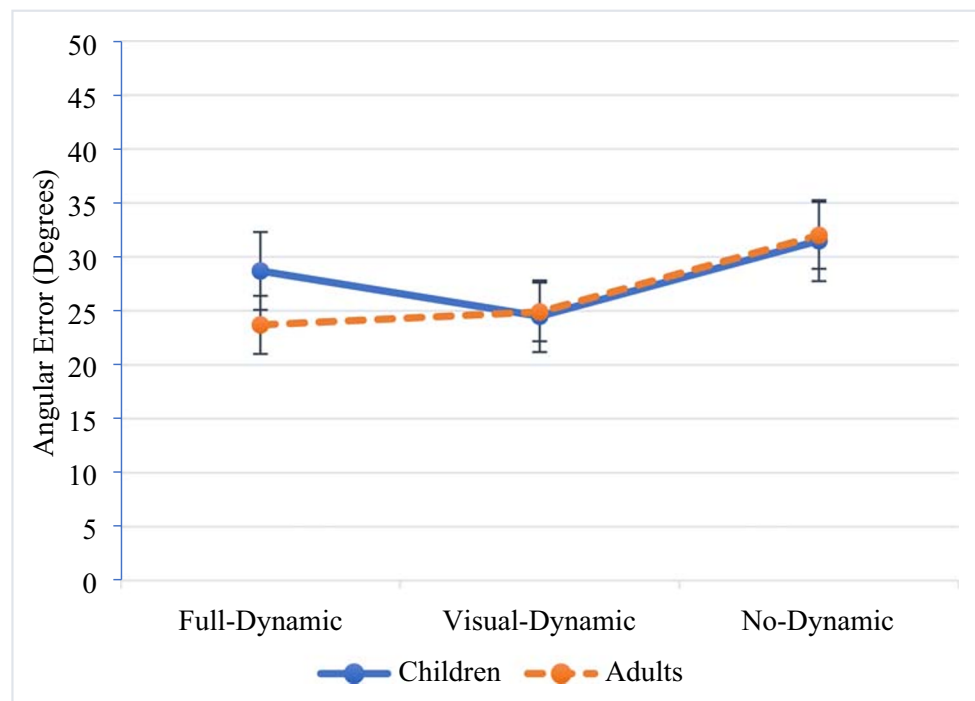
Fig. 5 Average angular error for children and adults in each condition. Error bars represent ± 1 standard error

than RT in Visual-Dynamic ($M = 8.60$, $SE = .52$, $z = -3.68$, $p = .0007$, $d = 0.05$) and No-Dynamic ($M = 8.57$, $SE = .52$, $z = -3.47$, $p = .002$, $d = 0.05$). Visual-Dynamic and No-Dynamic did not differ ($z = .24$, $p = .9$, $d = 0.003$). See Fig. 6 for the mean response times for each age group in each condition.

For adults only, there was a significant effect of Condition ($\chi^2(2) = 21.96$, $p < .001$, $AIC = 3205.4$, $BIC = 3233.0$). Full-Dynamic RT ($M = 8.23$, $SE = .56$) was significantly quicker than Visual-Dynamic ($M = 8.76$, $SE = .59$, $z = -3.07$, $p = .006$, $d = 0.06$) and No-Dynamic ($M = 9.02$, $SE = .61$, $z = -4.60$, $p < .0001$, $d = 0.08$). Visual-Dynamic and No-Dynamic did not differ ($z = -1.57$, $p = .3$, $d = 0.03$). For children only, there was not a significant effect of Condition ($\chi^2(2) = 4.85$, $p = .09$). Lastly, we tested the effect of age group for each of the three conditions. There was no age group effect for any of the conditions ($\chi^2(1)s < 1$, $ps > .8$).

Individual differences measures

Because of the non-significant effects of Age Group and the Age Group × Condition interaction, we dropped those terms from the model. Contrary to Experiment 1, there was no effect of eyes open ($\chi^2(1) = .54$, $p = .5$) or eyes closed ($\chi^2(1) = .61$, $p = .4$) balance ability. However, there was a significant effect of mental rotation ($\chi^2(1) = 5.06$, $p = .02$, $AIC = 6028.1$, $BIC = 6064.3$), with an increase in MRT score relating to decreases in error across conditions ($B = -.14$, $\beta = -.05$, $p = .02$), similar to Experiment 1.

Finally, we tested effects of spatial activities and gaming (adults and children completed the same questionnaire). There

was no effect of spatial activities participation ($\chi^2(1) = 1.91$, $p = .2$) or gaming ($\chi^2(1) = .15$, $p = .7$). Children ($M = 5.47$, $SE = 4.39$) reported more video game play than adults ($M = 2.13$, $SE = 3.54$, $F(1,59) = 10.72$, $p = .002$).

Taken together, results from Experiment 2 revealed largely similar patterns of performance between conditions as in Experiment 1, although errors were overall higher, at least for adults. Adult mean error was greater in all conditions in this environment compared to Experiment 1, by at least 5°, which replicates prior work arguing for the importance of visual landmark cues in spatial updating (Kelly et al., 2008; Zhao & Warren, 2015). We were surprised to observe similar levels of performance in adults and children in all conditions. Contrary to our expectations, removing the room cues did not appear to hurt performance in children, and while difficult to directly compare across participant samples,[3] overall mean accuracy was higher for children. In fact, adults' errors increased about as much as children's errors decreased between the two experiments.

General discussion

In two experiments, we tested the use of dynamic self-motion information in a point-to-origin task for adults and children in an environment with visual landmark cues (Experiment 1) and

---

[3] Many of the child participants were recruited from a summer coding camp on the University campus and could have had superior spatial abilities to our prior Experiment 1 sample. A t-test comparing the two child groups' mental rotation scores revealed a trending effect ($t = 1.98$, $p = .053$) with the children in Experiment 2 having higher MRT.
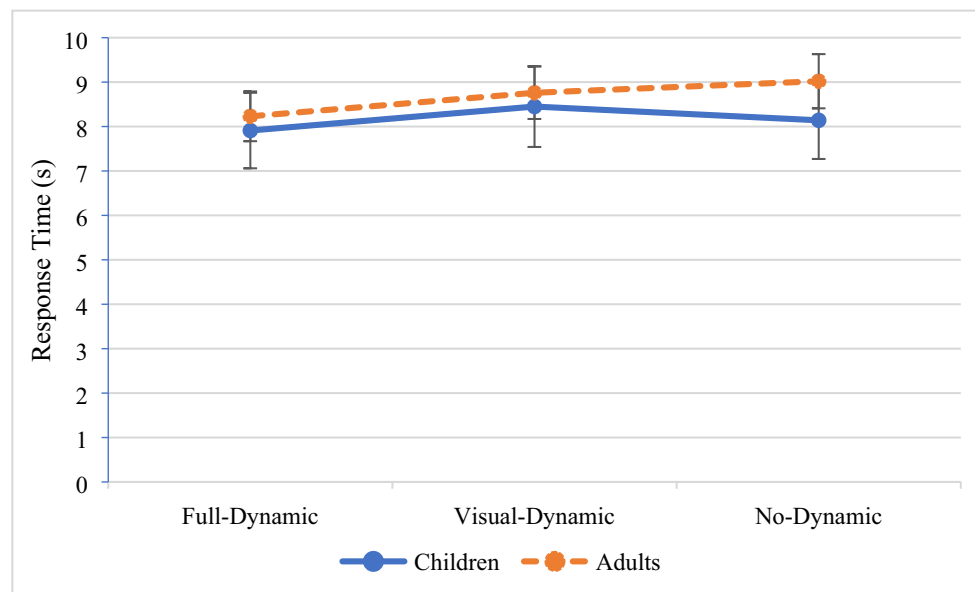
Fig. 6 Average response times for children and adults in each locomotion condition. Error bars represent ± 1 standard error

one with no landmark cues (Experiment 2). In Experiment 1, adults were impaired when there was no dynamic self-motion information (teleporting), but were similarly accurate when translation was performed with full or visual-only self-motion information. Children were also impaired by a lack of dynamic self-motion information, but showed higher accuracy in the visual-only condition compared to using full self-motion information, counter to our predictions. In Experiment 2, we eliminated visual room cues to test whether these effects would replicate in a landmark-free environment. Adults had higher mean errors compared to Experiment 1, but showed the same pattern of performance between conditions, with performance only being impaired when no dynamic self-motion information was present. In Experiment 2, children also showed this pattern among the locomotion conditions, and did not differ significantly from adults.

Overall, these results suggest that spatial updating is impaired when dynamic self-motion information is absent, especially for adults. Other recent work has shown decrements for spatial updating in adults when using teleporting methods compared to walking (Barhorst-Cates et al., 2020; Cherep et al., 2020). Here, we show that teleporting is worse than having only visual information for translation for spatial updating in both adults and children. We were again surprised to observe similar performance when both body-based and visual information was available, as we expected individuals to benefit from the multiple cues available in the full-dynamic condition (Chrastil et al., 2019; Sjolund et al., 2018). However, it is possible that the rotational movements alone were sufficient to elicit automatic spatial updating in this task (Chance et al., 1998; Klatzky et al., 1998; Rieser, 1989; Wraga et al., 2004). We expect that eliminating real rotations (Cherep et al., 2020) or using more complex spatial updating

tasks (Loomis et al., 1993; Ruddle et al., 2011) would show larger decrements in performance.

However, in both experiments we observed slower response times for adults when only dynamic visual information was compared to both visual and body-based information, which replicates the response-time effects observed in Barhorst-Cates et al. (2020). This suggests that although accuracy is the same between these dynamic self-motion conditions, having both visual and body-based cues for translation may result in easier computations of point-to-origin estimates. This advantage for visual and body-based information together is likely because walking elicits automatic spatial updating (Rieser, 1989), which may provide more direct access to the spatial knowledge required to compute the estimated heading.

For children, we observed different patterns of performance in the two environments. In both experiments, children were impaired in the no-dynamic condition, demonstrating the benefit of dynamic self-motion information for translation, similar to adults. We were surprised in Experiment 1 to observe better performance for children with visual only information compared to visual and body-based information. Worse performance with availability of both cues may be related to deficits in sensorimotor calibration (O'Neal et al., 2018; Petrini et al., 2016, Nardini et al., 2008; Newell & Wade, 2018). While walking provides multiple self-motion cues, these cues may also serve as multiple sources of potential noise, which could lead to greater error accumulation. Children's poorer motor control, more variable walking patterns, and rapidly changing body sizes may exacerbate the noise of the body-based signal (Newell & Wade, 2018). In contrast, the visual-dynamic condition is a single signal, which may be more "pure" and less prone to error accumulation for children. However, this enhanced performance with

only visual information for translation was not observed for children in Experiment 2, which suggests that the effect may be present only in situations where salient landmark cues are available. Children also showed response-time costs associated with any reduction in self-motion information in Experiment 1, but this effect was not observed in Experiment 2, where children responded in about the same amount of time regardless of condition. This difference could be due to children's processing and decision time to make a heading estimate being less sensitive to manipulations of translation information than adults. It is also possible that the visual context in Experiment 1 was distracting and the lack of visual landmarks in Experiment 2 may have made it easier for children to process heading estimates across all locomotion conditions. These reasons are speculative, though, and the nuanced relationship between locomotion method and environment type for different individuals and ages is an area in need of more research.

We did not compare across experiments due to the use of separate samples, but we observed interesting overall performance differences between experiments that warrant discussion. Adults' error increased in the outdoor environment in Experiment 2 in all conditions and children's error decreased about the same amount, making the difference between the age groups overall smaller. We postulate that children may have improved in Experiment 2 because they no longer had visual landmark cues to potentially bias their performance (Petrini et al., 2016), and adult performance may have worsened due to lack of visual landmark cues (Cherep et al., 2020; Kelly et al., 2008; Zhao & Warren, 2015). Children may have also been more negatively affected by the rotation of the virtual room on each trial (which was done to increase variability in the perceived home target location). Even though we took precautions to reduce attention to real room locations (e.g., white noise), it is possible that the changing perspective of the landmarks in the virtual room influenced children more than adults. The higher errors for children in Experiment 1 are consistent with prior research arguing that children cannot ignore visual cues because they rely on those cues to calibrate other sensory modalities (Gori et al., 2008; Petrini et al., 2015). Future research should test the same individuals in the two environments with the different locomotion methods to explore this idea.

For our individual differences measures, we provide evidence for the role of mental rotation in spatial updating, which we observed in both experiments. We found that better mental rotation is related to lower angular errors in the point-to-origin task, consistent with prior research demonstrating a relationship between small- and large-scale spatial abilities (e.g., Hegarty et al., 2006; Ruginski et al., 2019). Mental rotation may underlie larger scale spatial abilities, such as spatial updating, because it requires the ability to spatially transform objects, an important activity involved in navigation, for instance when processing a map presented at a different perspective or when considering the relative positioning of objects that are out of sight and at a different angle. Better mental rotation ability may improve spatial updating because it allows a more accurate (re)assessment of where one has been in space and how one may compute an estimate of return-to-home heading while taking into account the position of landmark objects.

We also observed weak evidence for a role of balance ability in spatial updating. Consistent with prior small-scale spatial studies (Frick & Möhring, 2016), improved balance related to decreased pointing error but only in the first experiment. As Frick and Möhring (2016) discuss, better balance is a foundation for more complex motor skills, such as locomotion, that may allow an individual to better explore and develop understanding of the spatial world. Balance may also reflect an optimal coordination of visual, proprioceptive, and vestibular information that could help to build stable and reliable spatial representations, which could then translate into better performance (Frick & Möhring, 2016). We are unsure why the effect was only observed in Experiment 1, although it is possible that the lack of landmarks in Experiment 2 may have made it so that this optimal coordination as indexed by balance ability was not as important to the task. It is possible that other, more sensitive measures of motor control would be better predictors of performance across different samples and environments, which should be investigated in future research. Although beyond the scope of this study, improvements in motor control and mental rotation during adolescence may be contributors to the development of spatial updating skill. It is also possible that videogame experience could influence spatial updating in VR, and that this may have had different effects on children's and adults' performance (i.e., more or less familiarity with the locomotion methods).

This study had several limitations, including sample size. While our sample size within each age group was likely sufficient for detecting within-subjects effects, the between-subjects age group effects should be interpreted with caution due to the relatively small samples. Additionally, our sample size may have been too small to observe strong individual differences effects, although we were able to detect effects of mental rotation. The effect sizes were quite small as well, which is apparent when examining the mean differences between conditions – performance was quite similar in some comparisons. However, there is debate about proper effect size reporting in mixed effects models (Peugh, 2010), so these values should be interpreted with caution. Another limitation involves a technical component that could potentially explain why real translation in the full-dynamic condition was not better than visual-dynamic translation. In the visual-dynamic condition, translation was locked to the heading direction of the HMD, forcing a straight-line path for navigation without deviation. Speed was constant and not controlled by the participant. With real walking in the full-dynamic condition, this

inherent straight-line path was absent and participants were able to move between points at their own speed and trajectory. It is possible that potential deviations from a straight-line path when walking could have made the integration of the path "leakier" (Lappe et al., 2007) compared to the visual-dynamic translation, resulting in more accumulated error. Our visual-dynamic condition also differs from traditional joystick virtual locomotion methods, in that participants pulled the trigger on the back of the controller and were only able to translate directly toward a target. This limited the number of degrees of freedom for movement in visual-dynamic compared to full-dynamic. Natural lateral oscillations were also disabled in the dynamic-visual condition, which makes direct comparison between these conditions difficult. Future research should compare full-dynamic to true joystick locomotion with controllable speeds and allow deviations from a straight-line path.

Overall, we replicate and extend previous findings that spatial updating is impaired with locomotion methods that do not allow for any dynamic self-motion information, in both children and adults. We also provide evidence that in a point-to-origin task that allows for physical rotation, translation specified only by vision (using the constrained path control implemented here) leads to comparable performance to real walking. Further, children show some advantage with this visual translation, at least in visually rich environments. These results advance our understanding of the sensory-motor information used in spatial updating and suggest that the choice of virtual locomotion methods used in applications should consider both the visual context of the environment and the age of the user.

References

Barhorst-Cates, E.M., Stefanucci, J.K., & Creem-Regehr, S.H. (2020). A comparison of virtual locomotion methods in movement experts and non-experts: Testing the contributions of body-based and visual translation for spatial updating. *Experimental Brain Research, 1–13.*

Bremmer, F., & Lappe, M. (1999). The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. *Experimental Brain Research 127*, 33–42.

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence, 7*(2), 168–178.

Cherep, L. A., Lim, A. F., Kelly, J. W., Acharya, D., Velasco, A., Bustamante, E., Ostrander, A. G., & Gilbert, S. B. (2020). Spatial cognitive implications of teleporting through virtual environments. *Journal of Experimental Psychology: Applied.* Advance online publication.

Chrastil, E. R., Nicora, G. L., & Huang, A. (2019). Vision and proprioception make equal contributions to path integration in a novel homing task. *Cognition, 192*, 103998.

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic Bulletin & Review, 19*(1), 1–23.

Chrastil, E. R., & Warren, W. H. (2013). Active and passive spatial learning in human navigation: Acquisition of survey knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(5), 1520.

Coomer, N., Bullard, S., Clinton, W., & Williams-Sanders, B. (2018). Evaluating the effects of four VR locomotion methods: Joystick, arm-cycling, point-tugging, and teleporting. In *Proceedings of the 15th ACM Symposium on Applied Perception* (p. 7). ACM.

De Beni, R. C. Meneghetti, C. Fiore, F., Gava, L., & Borella, E. Da Batteria VS. (2014). Valutazione delle abilità e delle autovalutazioni visuo-spaziali nell'arco di vita adulta, di Rielaborato da Vandenberg e Kuse (1978).

Downing, H. C., Barutchu, A., & Crewther, S. G. (2015). Developmental trends in the facilitation of multisensory objects with distractors. *Frontiers in Psychology, 5*, 1559.

Farrell, M. J., & Thomson, J. A. (1998). Automatic spatial updating during locomotion without vision. *The Quarterly Journal of Experimental Psychology: Section A, 51*(3), 6377–654.

Foreman, N., Orencas, C., Nicholas, E., Morton, P., & Gell, M. (1989). Spatial awareness in seven to 11-year-old physically handicapped children in mainstream schools. *European Journal of Special Needs Education, 4*(3), 171–179.

Frick, A., & Möhring, W. (2016). A matter of balance: Motor control is related to children's spatial and proportional reasoning skills. *Frontiers in Psychology, 6*, 2049.

Fujita, N., Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (1990). A minimal representation for dead-reckoning navigation: Updating the homing vector. *Geographical Analysis, 22*(4), 324–335.

Gabbard, C., Caçola, P., & Bobbio, T. (2012). The ability to mentally represent action is associated with low motor ability in children: A preliminary investigation. *Child: Care, Health and Development, 38*(3), 390–393.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology, 18*(9), 694–698.

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence, 34*(2), 151–176.

Hettinger, L. J. (2002). Illusory self-motion in virtual environments. In K. M. Stanney (Ed.), *Handbook of Virtual Environments* (pp. 471–492). Hillsdale: Lawrence Erlbaum.

Huffman, D.J., & Ekstrom, A.D. (2019). A modality-independent network underlies the retrieval of large-scale spatial environments in the human brain. *Neuron 104*(3), 611–622.

Jansen, P., & Heil, M. (2010). The relation between motor development and mental rotation ability in 5-to 6-year-old children. *European Journal of Developmental Science,* 4, 66–74.

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, doi: https://doi.org/10.1371/journal.pone.0072170

Kelly, J. W., McNamara, T. P., Bodenheimer, B., Carr, T. H., & Rieser, J. J. (2008). The shape of human navigation: How environmental

geometry is used in maintenance of spatial orientation. *Cognition*, *109*(2), 281–286.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, *9*(4), 293–298.

Lappe, M., Jenkin, M., & Harris, L. R. (2007). Travel distance estimation from visual motion by leaky path integration. *Experimental Brain Research*, *180*(1), 35–48.

Lehnung, M., Leplow, B., Ekroll, V., Herzog, A., Mehdorn, M., & Ferstl, R. (2003). The role of locomotion in the acquisition and transfer of spatial knowledge in children. *Scandinavian Journal of Psychology*, *44*(2), 79–86.

Lehnung, M., Leplow, B., Friege, L., Herzog, A., Ferstl, R., & Mehdorn, M. (1998). Development of spatial memory and spatial orientation in preschoolers and primary school children. *British Journal of Psychology*, *89*(3), 463–480.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General*, *122*(1), 73.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. *Wayfinding behavior: Cognitive mapping and other spatial processes*, 125–151.

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology*, *18*(9), 689–693.

Newell, K. M., & Wade, M. G. (2018). Physical growth, body scale, and perceptual-motor development. In *Advances in child development and behavior* (Vol. 55, pp. 205–243). JAI.

Nguyen-Vo, T., Riecke, B. E., Stuerzlinger, W., Pham, D. M., & Kruijff, E. (2019). NaviBoard and NaviChair: Limited Translation Combined with Full Rotation for Efficient Virtual Locomotion. *IEEE transactions on visualization and computer graphics*.

O'Neal, E. E., Jiang, Y., Franzen, L. J., Rahimian, P., Yon, J. P., Kearney, J. K., & Plumert, J. M. (2018). Changes in perception–action tuning over long time scales: How children and adults perceive and act on dynamic affordances when crossing roads. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 18.

Paris, R., Klag, J., Rajan, P., Buck, L., McNamara, T. P., & Bodenheimer, B. (2019). How video game locomotion methods affect navigation in virtual environments. In *ACM Symposium on Applied Perception 2019*, 1–7.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test-different versions and factors that affect performance. *Brain and Cognition, 28*(1), 39–58.

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.

Petrini, K., Jones, P. R., Smith, L., & Nardini, M. (2015). Hearing where the eyes see: Children use an irrelevant visual cue when localizing sounds. *Child development*, *86*(5), 1449–1457.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of school psychology*, *48*(1), 85–112.

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

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception, 23*(12), 1447–1455.

Riecke, B. E., Cunningham, D.W., & Bülthoff, H.H. (2007). Spatial updating in virtual reality: The sufficiency of visual information. *Psychological Research, 71*(3), 298–313.

Riecke, B. E., Heyde, M. V. D., & Bülthoff, H. H. (2005). Visual cues can be sufficient for triggering automatic, reflexlike spatial updating. *ACM Transactions on Applied Perception (TAP)*, *2*(3), 183–215.

Riecke, B. E., Schulte-Pelkum, J., Caniard, F., & Bulthoff, H. H. (2005, March). Towards lean and elegant self-motion simulation in virtual reality. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* (pp. 131–138). IEEE.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1157.

Rieser, J. J., Guth, D. A., & Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision. *Perception*, *15*(2), 173–188.

Roberts, R., Callow, N., Hardy, L., Markland, D., & Bringer, J. (2008). Movement imagery ability: Development and assessment of a revised version of the vividness of movement imagery questionnaire. *Journal of Sport and Exercise Psychology*, *30*(2), 200–221.

Ruddle, R. A., & Lessels, S. (2006). For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science*, *17*(6), 460–465.

Ruddle, R. A., Volkova, E., & Bülthoff, H. H. (2011). Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(2), 10.

Ruginski, I. T., Creem-Regehr, S. H., Stefanucci, J. K., & Cashdan, E. (2019). GPS use negatively affects environmental learning through spatial transformation abilities. *Journal of Environmental Psychology*, *64*, 12–20.

Sjolund, L. A., Kelly, J. W., & McNamara, T. P. (2018). Optimal combination of environmental cues and path integration during navigation. *Memory & Cognition*, *46*(1), 89–99.

Smith, A. D., McKeith, L., & Howard, C. J. (2013). The development of path integration: combining estimations of distance and heading. *Experimental Brain Research, 231*(4), 445–455.

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

Trutoiu, L. C., Mohler, B. J., Schulte-Pelkum, J., & Bülthoff, H. H. (2009). Circular, linear, and curvilinear vection in a large-screen virtual environment with floor projection. *Computers & Graphics*, *33*(1), 47–58.

Warren Jr, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience*, *4*(2), 213.

Wraga, M., Creem-Regehr, S. H., & Proffitt, D. R. (2004). Spatial updating of virtual displays. *Memory & Cognition*, *32*(3), 399–415.

Yan, J. H., Thomas, J. R., & Downing, J. H. (1998). Locomotion improves children's spatial search: A meta-analytic review. *Perceptual and Motor Skills*, *87*(1), 67–82.

Zhao, M., & Warren, W.H. (2015). How you get there from here: Interaction of visual landmarks and path integration in human navigation. *Psychological Science, 26*(6), 915–924.