Article

PERCEPTION

Spatial Reference Frame but Neither Age nor Gender Predict Performance on a Water-Level Task in 8- to 11-Year-Old Children

Perception
2020, Vol. 49(11) 1200–1212

DOI: 10.1177/0301006620964414
journals.sagepub.com/home/pec
SAGE

Erica M. Barhorst-Cates
Moss Rehabilitation Research Institute, United States

Sarah H. Creem-Regehr and Jeanine K. Stefanucci
University of Utah, United States

Jean Gardner
Salt Lake City School District, United States

Trish Saccomano and Cheryl Wright
University of Utah, United States

Abstract
Successful performance on the water-level task, a common measure of spatial perception, requires adopting an environmental, rather than object-centered, spatial frame of reference. Use of this strategy has not been systematically studied in prepubertal children, a developmental period during which individual differences in spatial abilities start to emerge. In this study, children aged 8 to 11 reported their age and gender, completed a paper-and-pencil water-level task, and drew a map of their neighborhood to assess spontaneous choice of spatial frame of reference. Results showed a surprising lack of age or gender difference in water-level performance, but a significant effect of spatial frame of reference. Although they made up only a small portion of the sample, children who drew allocentric maps had the highest water-level score, with very high accuracy. These results suggest that children who adopt environmental-based reference frames when depicting their familiar environment may also use environmental-based reference frame strategies to solve spatial perception tasks, thereby facilitating highly accurate performance.

Corresponding author:
Erica M. Barhorst-Cates, Moss Rehabilitation Research Institute, Elkins Park, PA, United States.
Email: BarhorsE@einstein.edu

Keywords
water-level task, spatial frames of reference, middle childhood

Date Received: 1 June 2020; accepted: 16 September 2020

Common tests of spatial perception involve determining horizontality or verticality while ignoring potentially biasing visual information. The water-level task, originally described by Piaget and Inhelder (1948), requires participants to draw a line of water in various depicted glasses that are tilted in different ways. While the correct answer is always a perfectly horizontal line, performance on this task has shown marked individual differences in terms of gender (Vasta et al., 1996), age (Thomas & Turner, 1991), and even liquid-filled-container-carrying expertise (i.e., bartending/waitressing; Hecht & Proffitt, 1995). Piaget and Inhelder proposed that accurate water-level performance may be achieved by around 9 years of age. However, research has shown that many adults remain inaccurate (Vasta et al., 1996), and individual differences are extensive even in adults (Hecht & Proffitt, 1995), which has led to numerous questions about the strategies used to solve the task.

Correctly solving the water-level task involves using an environmental reference frame, where axes are determined based on gravitational horizontal and vertical. People who perform worse may instead use an object-relative reference system, where axes are determined based on the object alone, disregarding environmental features (Hecht & Proffitt, 1995). While few methods exist for assessing reference frames for small-scale spatial tasks, reference frames are commonly assessed in the context of larger scale spatial tasks such as navigation. Egocentric or viewer-dependent perspectives (the representation of the environment is from the point of view of someone standing within it) can be distinguished from allocentric or world-dependent (the representation of the environment is from a birds-eye view) perspectives. Allocentric reference frames are more stable because the representation involves a fixed coordinate system that does not rely on the potentially dynamic viewpoint of the viewer. Allocentric reference frames are also generally considered more difficult to compute (Klatzky, 1998) because of the need to imagine a perspective that is different than that of the viewer. Similar to the water-level task described earlier, in which the individuals who make use of environmental reference frames tend to outperform those who make use of object-relative reference frames, some research suggests that allocentric reference frames are more often used by better navigators (Chen et al., 2009). These better navigators may also have better underlying small-scale spatial abilities (Blajenkova et al., 2005; Hegarty et al., 2006; Ruginski et al., 2019), suggesting a relationship between frames of reference in large-scale spatial thinking and performance on small-scale spatial tasks. People who use environmental-dependent rather than egocentric or object-centered frames of reference may succeed because those reference frames are more stable or constant across perspectives (Silverman et al., 2000).

Developmentally, classic theories argue that children progress from egocentric to allocentric representations of environments with development and experience (Piaget & Inhelder, 1967; Siegel & White, 1975) and that allocentric representations may not be used until fairly late in childhood. More recent work has shown that allocentric representations can be used by very young children in certain tasks (2–3 years old, Ribordy et al., 2013;

5 years old, Nardini et al., 2006; Newcombe & Huttenlocher, 2003), suggesting that more flexible use of different types of representations may be developed quite early in life and can be used when needed. Nevertheless, some research shows a developmental progression from egocentric to allocentric representations from age 5 to 10 on certain tasks (Bullens et al., 2010), with children performing like adults on spatial frames of reference tasks around the age of 10 (Ruggiero et al., 2016). As such, the age at which children adopt allocentric representations is unclear. Gender differences have also long been reported in the water-level task (e.g., Geiringer & Hyde, 1976; Signorella & Jamison, 1978), with males typically outperforming females. It follows that males sometimes use more allocentric spatial strategies (e.g., Dabbs et al., 1998; Sandstrom et al., 1998) on certain tasks, which may facilitate their performance (Boone et al., 2019; Hegarty et al., 2002, 2006; Lawton, 1994; Pazzaglia & De Beni, 2001). The prepubertal age is of particular interest in understanding individual differences in spatial cognition, as some research argues that individual differences (particularly gender differences) may not emerge until puberty (Newhouse et al., 2007) or just before (Johnson & Meade, 1987; Matthews, 1987). Gender differences in spatial abilities (and at what age they can be observed) seem to be task-specific (Coluccia & Louse, 2004) and the prior research is inconclusive. Male and female children perform similarly on some tasks, especially when navigating in a familiar environment (Coluccia & Louse, 2004; Herman et al., 1987), but males outperform females on other tasks such as route finding (Merrill et al., 2016; Newhouse et al., 2007), and this advantage can be seen even early in childhood (6 years old; Merrill et al., 2016). By the age of 11, males draw more detailed spatial maps of routes through their neighborhood than females (Matthews, 1987), which correlates with the extent of exploration of the neighborhood. As such, it is unclear whether and when sex differences emerge in childhood, although the years immediately before puberty seem to be an active developmental period for spatial cognition. Much of the prior research has considered gender differences in spatial abilities in childhood, but there remain open questions about other individual differences such as frame of reference choice (types of which may be more prominent in the different genders). Additional research is needed to understand the development of individual differences by further understanding spatial abilities at prepubertal age (below the age of 12).

Determining an individual's reference frame has been a challenge for researchers from many different fields. One common method involves sketch maps of a familiar environment or route, which can then be classified by the use of particular frames of reference or other key spatial components (Blajenkova et al., 2005). Many have argued that sketch maps provide an accurate assessment of one's spatial representation of their environment, as sketch map variables are strong predictors of performance on actual navigation tasks (Rovine & Weisman, 1989) and verbal environment descriptions (Tversky, 2000). Evidence is mixed regarding whether the reference frame of sketch maps relates to performance on small-scale spatial tasks such as water level. In one large study of children aged 7 to 12 in Germany, Quaiser-Pohl et al. (2004) found no relationships between sketch map frame of reference and performance on mental rotation, water-level, or rod and frame tasks. Moreover, the spatial and mapping tasks loaded onto two distinct and not significantly correlated factors, which the authors termed *spatial ability* and *spatial representations of the environment*. However, Blajenkova et al. (2005) found that participants who drew three-dimensional or two-dimensional sketch maps of a learned environment (i.e., maps that included more spatial-relational features and indicators of different levels in the building, which represented allocentric processing) performed better on a mental rotation task than those who drew one-dimensional maps (which represented egocentric processing). As such, evidence has been mixed regarding the relationship between sketch map factors and small-scale spatial abilities such as spatial perception. However, many of the sketch map paradigms used in prior research include explicit instructions to map a certain route through an environment (Aginsky et al., 1997; Moeser, 1988; Rovine & Weisman, 1989), either familiar or recently learned. As such, there are interesting open questions about categorizations of sketch maps that measure more "default" frames of reference. It is unknown, for instance, how the spontaneous frame of reference (i.e., the one taken without task-specific instructions) one portrays in a drawing of a familiar environment relates to spatial perception. Spontaneous frames of reference reveal important individual differences in spatial thinking, as has been demonstrated with the well-known "animals in a row" task (e.g., Levinson & Schmitt, 1993) that classifies individuals' spatial thinking into default egocentric or allocentric. We aimed to measure spatial thinking about the environment in a similar default way and examine the relationship with water-level performance.

Taken together, prior research has revealed compelling individual differences in water-level task performance related to age and gender, although the age at which gender differences emerge for spatial abilities is debated. Successful performance on the water-level task depends on the reference frame used, with those who use an environmental reference frame outperforming those who use an object-centered reference frame. This study aimed to characterize individual differences in water-level performance in prepubertal 8- to 11-year-old children. We assessed age, gender, and spatial frames of reference, anticipating that males may outperform females, accuracy should increase with increasing age, and that the adoption of an allocentric reference frame in map drawing should be related to higher accuracy in water-level task performance. We used an open-ended sketch map task to measure default frames of reference similar to the spontaneous frame of reference assessed in the animals-in-a-row task (Levinson & Schmitt, 1993).

Method

Participants

Participants included 72 children (35 males, 36 females, and 1 missing) aged 8 to 11 ($M = 9.28$, $SD = .76$). Parents gave consent for their child to participate using an opt-out procedure. Participant assent was given on the day of testing using an opt-out procedure at any time during the study. We had no exclusion criteria and no participant opted out of the procedure. All procedures were approved by the University of Utah institutional review board.

Materials and Procedure

The results described here were part of a larger study aimed at assessing spatial abilities in children in middle childhood, including measures of mental rotation, spatial updating, and spatial working memory. We conducted this study at the participants' elementary schools in a high socioeconomic status neighborhood in Salt Lake City. Participants completed the battery of tasks on a rolling basis during the allotted time, rotating between specified stations set up in two classrooms. The water-level task and Neighborhood Map Drawing Task were administered in groups of three to four students with one instructor. Task order was not controlled, as participants were rotating through all tasks as they became available. As such, task order was quasi-randomized.

Open-Ended Neighborhood Map Drawing Task. Experimenters provided white paper and markers to each participant and instructed each participant to "draw your neighborhood." There was no further instruction and no time limit. Three separate coders classified the frame of reference of each neighborhood map as being fully egocentric, fully allocentric, or mixed. Maps were coded as fully egocentric following these instructions: "Only includes views from an egocentric perspective as if a viewer was standing in the environment and viewing a scene from a first-person perspective, no top-down or birds-eye perspective components." See Figure 1. Maps were coded as fully allocentric following these instructions: "Includes only views from a birds-eye or top-down perspective, no components of being in front of the houses, etc. If buildings are included, must portray only the tops of the buildings (not a front view)." See Figure 2. Finally, maps were coded as mixed if they contained both allocentric and egocentric components with the following instructions: "Includes both egocentric and allocentric views, front of buildings or houses portrayed in addition to map components, such as streets overviewed or other things that couldn't be seen from a view standing within the environment." See Figure 3. Interrater reliability was computed using Fleiss' Kappa for more than two raters in the R package *irr*. As data collection took place over the span of 3 years, the first 36 participants were rated by 3 raters soon after data collection with a Fleiss' Kappa of 0.72. The remaining 36 participants were rated separately by 3 raters (including 1 of the same raters from the first group) with a Fleiss' Kappa of 0.83. Cases that showed discrepancies among raters received the rating given by the majority of raters (two out of three raters). A total of 13 cases showed discrepancies (18.1% of the data).

Water-Level Task. Participants completed a paper and pencil water-level task with eight items that were tilted to match eight different angles: 0°, 90° to the right, 90° to the left, 180°, 45° to the right, 45° to the left, 135° to the right, and 135° to the left. Experimenters showed participants an image of a level bottle with a line of water drawn straight across.

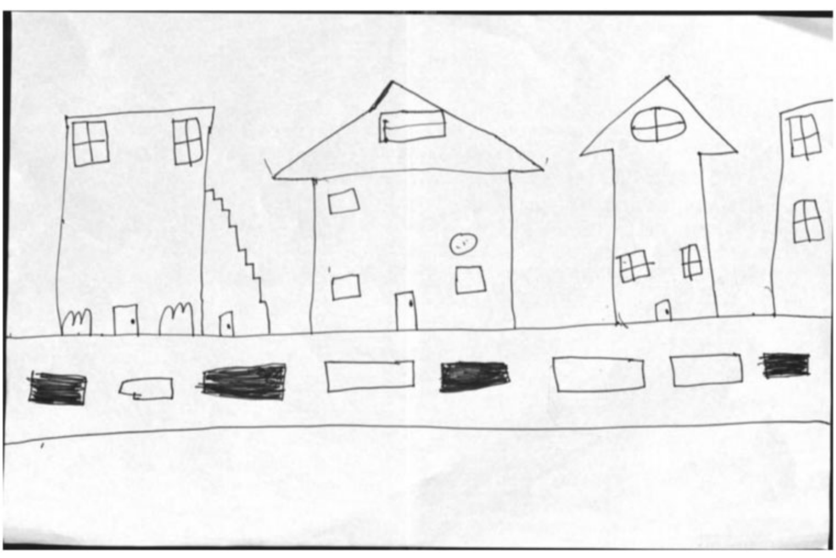

Figure 1. Example of an Egocentric-Only Perspective of a Child's Neighborhood.
The perspective is that of an individual standing in front of the houses and contains no birds-eye perspective components.

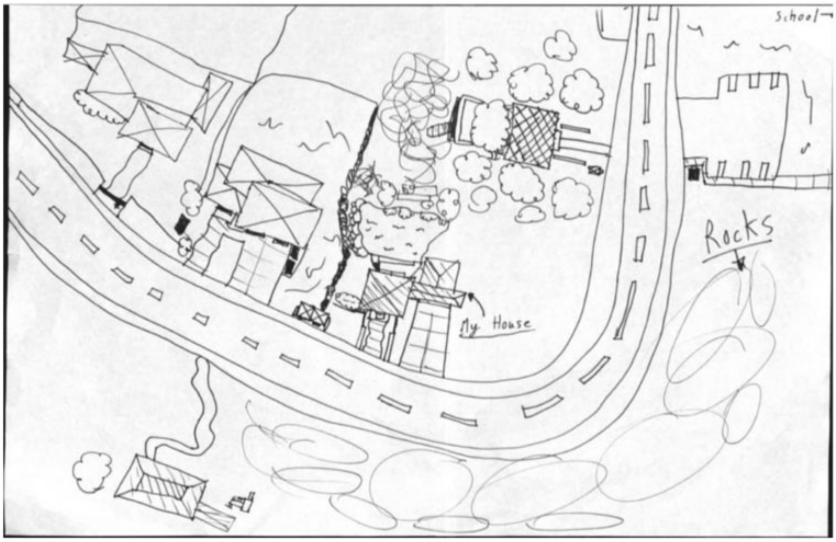

Figure 2. Example of an Allocentric-Only Perspective of a Child's Neighborhood.
Houses, roads, and trees are all represented from a birds-eye perspective and the drawing only contains depictions from earlier.

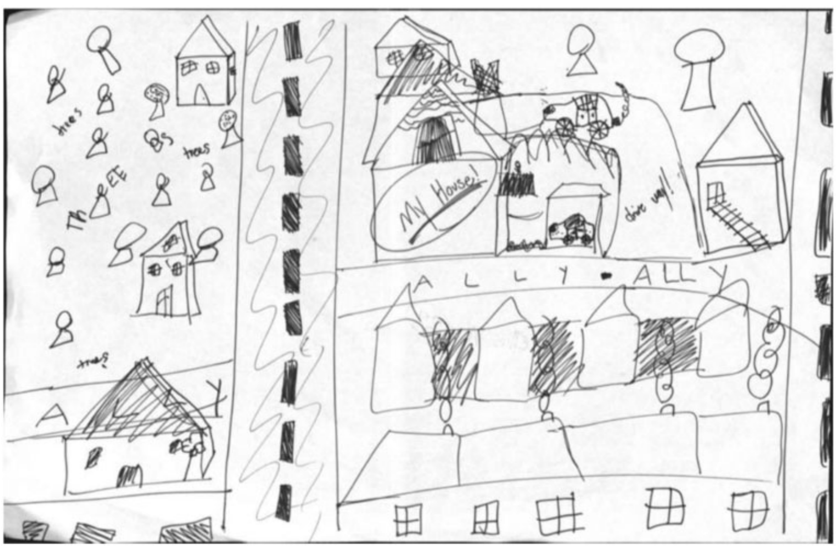

Figure 3. Example of a Mixed Egocentric and Allocentric Perspective.
The roads are depicted from a birds-eye perspective and overview the layout of the neighborhood, while the houses are depicted from an egocentric perspective (the participant drew the fronts of the houses).

They then instructed participants to draw the line of water where it should be in the eight images. Each image was presented individually in a stapled packet with a sheet of paper between trials. There was no time limit. We graded responses by first drawing the correct response (a perfectly horizontal line from one of the intersecting points determined by the participant). When needed, we drew a straight line between the two intersection points

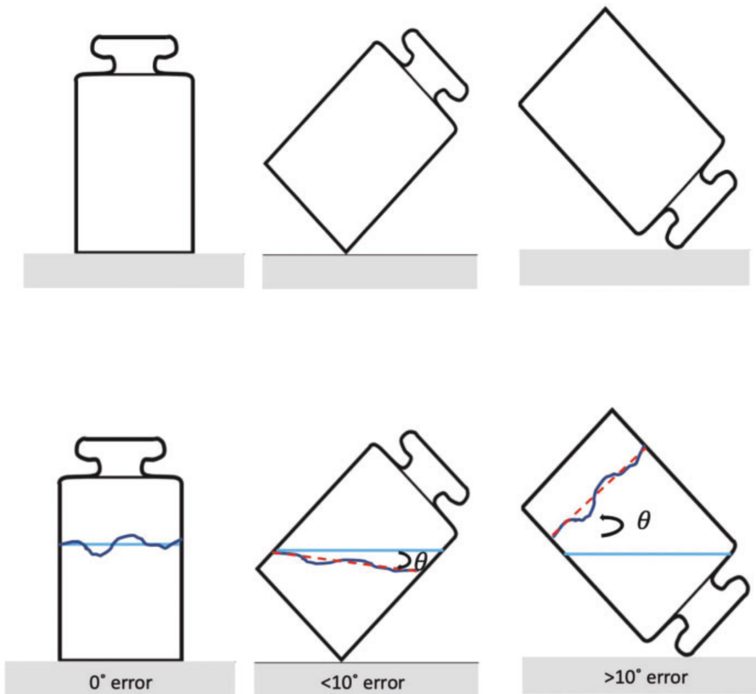

Figure 4. Three Trials of the Water-Level Task With Angles of 0°, 45° to the Right, and 135° to the Right. The bottom panel demonstrates potential responses. The light blue line indicates the correct response, the dark blue squiggly line represents the participant's indicated water level line, and the dotted red line represents the correction method applied by the researchers prior to calculating the angle. Trials were counted as correct if the straight-line-corrected response was less than 10° from the correct response (a perfectly horizontal line). Trials with greater than 10° deviation were counted incorrect.

designated by the participant (e.g., when the line was too squiggly or curved to be able to measure). Then, we measured deviation of the participant's response from the correct water line using a standard protractor. Water level line was counted as correct if it fell within 10° of perfectly horizontal (similar to the methods used by Quaiser-Pohl et al., 2004 and Hecht & Proffitt, 1995 but with a more liberal correct answer range [10° compared with 8° and 5°, respectively]). We graded each trial using this method and calculated a total number of correct items for each individual. We calculated a proportion correct out of eight for each participant. See Figure 4 for a depiction of the water-level task and example responses.

Design and Data Analysis

This correlational study aimed at assessing natural variation in the spontaneous frame of reference used by children in this age-group. We targeted our sampling at prepubescent children and formed post hoc groups according to the reference frame used in the Neighborhood Map task. These groups were highly uneven, justifying the use of nonparametric statistical tests to examine how the reference frame groups predicted accuracy on the Water-Level task. We chose to use the Kruskal–Wallis test, which is analogous to a one-way analysis of variance (Theodorsson-Norheim, 1986). We used multiple linear regression to assess age and gender differences on Water Level. All analyses were performed in *R* (R Core Team, 2017).

Results

The average proportion accuracy on the water level was 0.70 ($SD = .25$) with a range from 0.125 to 1.00. We first assessed the effects of gender and age on water-level performance using multiple linear regression. Surprisingly, neither gender ($B = .06$, $p = .3$) nor age ($B = .03$, $p = .4$) were significantly related to water-level proportion correct, $F(2, 68) = 1.12$, $R^2 = .03$, $p = .3$. Males' proportion correct on water level ($M = .74$, $SD = .22$) did not differ from females' ($M = .66$, $SD = .27$) and there was no significant change with age.

We next assessed the effect of neighborhood map reference frame on water-level proportion correct. A total of 38 maps were coded as fully egocentric, 28 as mixed, and 6 as fully allocentric. Because of the large discrepancy between groups in sample size, we used the Kruskal–Wallis nonparametric test, which revealed a significant effect of group, $\chi^2(2)=8.06$, $p = .02$. A post hoc pairwise Wilcox test with a Bonferroni–Holm adjustment for multiple comparisons revealed that this effect was driven by the allocentric-only group ($M = .94$, $SD = .07$, range: .875–1.00) performing significantly better on water level than the egocentric-only ($M = .66$, $SD = .24$, range: .125–1.00) group ($p = .01$). Performance did not differ significantly between allocentric-only and mixed ($M = .71$, $SD = .24$, range: .125–1.00) groups ($p = .07$), although the difference approached significance. See Figure 5 for a boxplot of results. We provide descriptive information about the participants who fell into each category in Table 1.

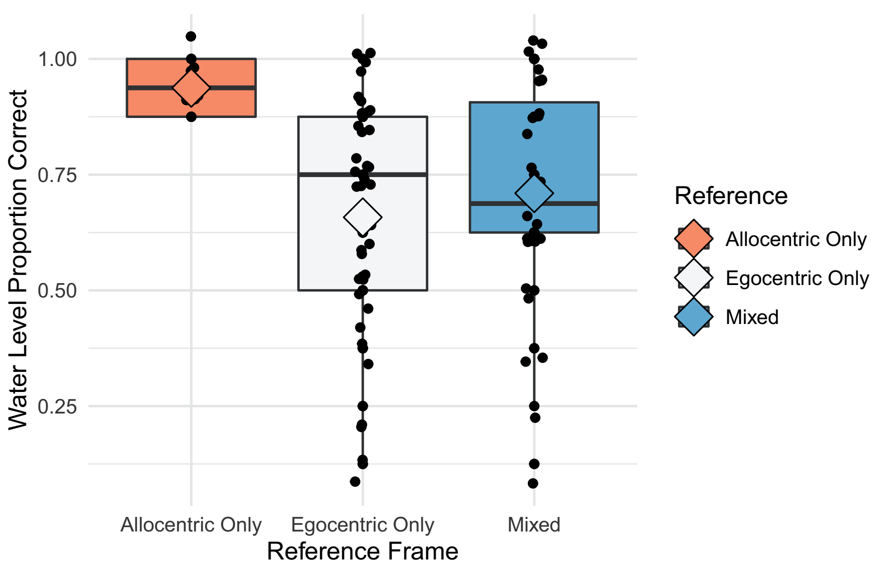

Figure 5. Boxplot of Results. The line indicates the median for each group and the diamond indicates the mean.

Table 1. Participant Data for Individuals Within Each Map Category.

Map category		Gender		Age		
	n	*n* male	*n* females	Mean (*SD*)	Mode	Count
Egocentric only	38	18	20	9.3 (.66)	9	3 8-year-olds 21 9-year-olds 12 10-year-olds 1 11-year-old
Mixed	28	12	16	9.14 (.89)	9	7 8-year-olds 12 9-year-olds 7 10-year-olds 2 11-year-olds
Allocentric only	6	5	1	9.8 (.41)	10	5 10-year-olds 1 9-year-old

Discussion

Children aged 8 to 11 completed a water-level task and a neighborhood sketch map task. Counter to our expectations, we observed no effects of age or gender. However, children who drew neighborhood map representations that contained allocentric-only components performed better on a water-level task than children who included egocentric components. The lack of a gender difference in this study suggests that the oft-observed male advantage in spatial perception (e.g., Geiringer & Hyde, 1976) may not emerge until later in development than the ages tested here (age 8–11). Moreover, the lack of an age effect in water-level performance suggests that spatial perception ability is relatively stable through the middle childhood years and does not improve substantially from age 8 to 11. Prior research has been mixed regarding gender differences in spatial cognition and at what age and for what tasks those differences may be observed (e.g., Coluccia & Louse, 2004). While we had an even number of males and females across the sample, the relatively low number of individuals at lower and higher ages may have minimized the potential to detect gender differences at different ages (i.e., the difference may be larger for those individuals closer to puberty). Our lack of an age effect is also counter to early research that has demonstrated improvements on water level around the age of 9 or 10 years (Piaget & Inhelder, 1948). Again, these results should be viewed with caution because of the relatively fewer individuals at younger and older ages. A limitation of our study was the lack of diversity of the children in our sample. They attended one of two schools in the same neighborhood and likely had similar opportunities for education and extracurricular activities that support spatial reasoning. It is possible that expanding the sample across different schools and socioeconomic status (SES) would have led to age or gender differences. Instead of the commonly studied factors of age or gender, we observed a significant relationship between spatial frame of reference and performance on the water-level task. Spatial frame of reference may be a unifying predictor that explains individual variation across age and gender in success at the water-level task. Indeed, those participants who drew allocentric-only neighborhood maps were disproportionately males (83.3%) and were almost all 10 years old. While these individuals made up only 8.3% of the sample, this finding sheds some light on who may be most likely to use allocentric representations.

This study suggests that the default reference frame one uses to represent a familiar environment is related to how one performs on a small-scale spatial perception task.

Namely, those who adopt an environmental (allocentric) reference frame in one task may also adopt that reference frame in the other task, which may facilitate performance. Prior research has shown that allocentric spatial representations are more often observed in good navigators (Chen et al., 2009), who also may have better small-scale spatial abilities (Blajenkova et al., 2005; Ruginski et al., 2019; Vashro & Cashdan, 2015). These data demonstrate that even in children aged 8 to 11 there appears to be a relationship between larger scale spontaneous spatial frames of reference and small-scale spatial perception abilities. This relationship may suggest a ubiquitous component of spatial thinking (i.e., adopting a certain frame of reference) that is present across different spatial task types.

These results suggest that the ability to ignore object-specific properties in a spatial perception task may also be reflected in how one mentally represents a familiar space in a default manner. This result is particularly interesting because it suggests a ubiquitous property of frames of reference and spatial perception that is present in middle childhood. Namely, the relationship between allocentric map drawing and better water-level task performance may indicate a general ability to flexibly represent space. To succeed at both the water-level task and drawing allocentric mental maps, one must be able to ignore or extrapolate information beyond what is immediately visible (Blajenkova et al., 2005). In a sense, both of these tasks represent an ability to ignore some visual information (current or imagined), which may be related to executive functioning components such as inhibition. Although not tested here, it is possible that individuals with poorer inhibition may be less likely to ignore object-based or visual information, which may relate to poorer performance on the water-level task or encourage defaulting to simpler frames of reference (i.e., egocentric) when drawing a map.

One important limitation in this study is that the directionality of the relationship between neighborhood map reference frame and water-level task performance cannot be assumed based on the correlational design. It is unknown whether better spatial perception abilities facilitate allocentric default representations or vice versa. Future research should further explore this relationship in a longitudinal manner and with training interventions. It would be particularly interesting to assess how spatial frames of reference and spatial perception abilities in middle childhood predict small- and large-scale spatial abilities in adulthood. If training on spatial perception improved allocentric representations or vice versa, this could have implications for education and training in spatial thinking. In addition, we also did not include an explicit assessment of strategy in the water-level task. As such, we cannot say for certain that those who performed with high accuracy were using a specific strategy (such as adopting an environmental reference frame; Hecht & Proffitt, 1995). Strategy use should be tested and manipulated in future research on the water-level task in this age-group.

Finally, we posit that the neighborhood map drawing task is a novel contribution of this work, in that it serves as an easy-to-administer, fun, instructionally simple assessment of frames of reference. It can be used in a variety of experimental settings such as a classroom or art space. We did not provide explicit instructions to participants other than to "draw your neighborhood." This design differs from that used by Quaiser-Pohl et al. (2004), in which researchers instructed child participants to draw a map explicitly regarding the places they often visit including the routes taken to get there. We chose to use vague instructions in an attempt to understand how children think about their neighborhoods without considering the task-specific requirement of wayfinding, similar to the animals in a row measure of spontaneous frame of reference (Levinson & Schmitt, 1993). Instead, these maps may reflect the more spontaneous frame of reference with which children represent their environment. The use of this method could be considered a limitation, because the measure is less controlled than map tasks that involve specific navigation-related instructions. Our measure may be capturing noise as well as other useful individual differences factors (such as creativity or artistic skill) that could be studied in future research. The vague instructions for the map task are similar in a way to the vague instructions in the water-level task (participants are not told to think about how the cup is tilted relative to the ground plane or provided any other instructions about potential strategies). As such, both tasks reflect the inferred frame of reference used when participants are not given instructions about the frame of reference and are instead free to choose to perform the task using whichever strategy they want. While purely allocentric frames of reference were relatively rare in our sample, which is one limitation of this study, those children who portrayed them had very accurate spatial perception. Moreover, we also observed that the majority of children defaulted to portraying their neighborhoods with egocentric-only reference fames. It is perhaps unsurprising that purely allocentric frames of reference were rare in our sample, given that people often use both egocentric and allocentric frames of reference in navigation tasks (e.g., Burgess, 2006). Future research should incorporate a variety of instructional manipulations in map-drawing tasks to assess the nature of the relationship between familiar environment representation and spatial perception abilities and to measure the commonality of these types of representations in various populations.

Authors' contributions

E. B., C. W., S. C., T. S., and J. G. designed the experiment. E. B., C. W., S. C., T. S., J. S., and J. G. collected the data. E. B. and J. S. processed the data and conducted the analyses. E. B. wrote the manuscript with feedback from J. S., S. C., and C. W.

Acknowledgements

The authors would like to thank Andrea Frick for providing the water-level task and helpful testing tips. The authors thank Padmapriya Muralidharan, Holly Gagnon, Morgan Saxon, Jessica Stoker, Boris Tirado, Mirinda Whitaker, and Grant Pointon for helping with data collection and are grateful to Padmapriya Muralidharan, Leigh Taylor, Taylor Larsen, and Makayla Luo for aiding with processing the data. The authors thank the Salt Lake City School District and teachers for working with us on this project, especially Erica Wangsgard who was an early collaborator.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by the Beverly Taylor Sorenson Arts Learning Program and the National Science Foundation, grant no. 1763254. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ORCID iD

Erica M. Barhorst-Cates https://orcid.org/0000-0001-9007-5652

References

Aginsky, V., Harris, C., Rensink, R., & Beusmans, J. (1997). Two strategies for learning a route in a driving simulator. *Journal of Environmental Psychology*, *17*, 317–331.

Blajenkova, O., Motes, M. A., & Kozhevnikov, M. (2005). Individual differences in the representations of novel environments. *Journal of Environmental Psychology*, *25*, 97–109.

Boone, A. P., Maghen, B., & Hegarty, M. (2019). Instructions matter: Individual differences in navigation strategy and ability. *Memory & Cognition*, *47*, 1401–1414.

Bullens, J., Iglói, K., Berthoz, A., Postma, A., & Rondi-Reig, L. (2010). Developmental time course of the acquisition of sequential egocentric and allocentric navigation strategies. *Journal of Experimental Child Psychology*, *107*, 337–350.

Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences*, *10*, 551–557.

Chen, C. H., Chang, W. C., & Chang, W. T. (2009). Gender differences in relation to wayfinding strategies, navigational support design, and wayfinding task difficulty. *Journal of Environmental Psychology*, *29*, 220–226.

Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of Environmental Psychology*, *24*, 329–340.

Dabbs, J. M., Jr, Chang, E. L., Strong, R. A., & Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evolution and Human Behavior*, *19*, 89–98.

Geiringer, E. R., & Hyde, J. S. (1976). Sex differences on Piaget's water-level task: Spatial ability incognito. *Perceptual and Motor Skills*, *42*, 1323–1328.

Hecht, H., & Proffitt, D. R. (1995). The price of expertise: Effects of experience on the water-level task. *Psychological Science*, *6*, 90–95.

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*, 151–176.

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*, 425–447.

Herman, J. F., Heins, J. A., & Cohen, D. S. (1987). Children's spatial knowledge of their neighborhood environment. *Journal of Applied Developmental Psychology*, *8*, 1–15.

Johnson, E. S., & Meade, A. C. (1987). Developmental patterns of spatial ability: An early sex difference. *Child Development*, *58*, 725–740.

Klatzky, R. L.(1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition* (pp. 1–17). Berlin: Springer.

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles*, *30*, 765–779.

Levinson, S. C., & Schmitt, B. (1993). Animals in a row. *Cognition and Space Kit Version*, *1*, 65–69.

Matthews, M. H. (1987). Sex differences in spatial competence: The ability of young children to map 'primed' unfamiliar environments. *Educational Psychology*, *7*, 77–90.

Merrill, E. C., Yang, Y., Roskos, B., & Steele, S. (2016). Sex differences in using spatial and verbal abilities influence route learning performance in a virtual environment: A comparison of 6-to 12-year old boys and girls. *Frontiers in Psychology*, *7*, 258.

Moeser, S. D. (1988). Cognitive mapping in a complex building. *Environment and Behavior*, *20*, 21–49.

Nardini, M., Burgess, N., Breckenridge, K., & Atkinson, J. (2006). Differential developmental trajectories for egocentric, environmental and intrinsic frames of reference in spatial memory. *Cognition*, *101*, 153–172.

Newcombe, N. S., & Huttenlocher, J. (2003). *Making space: The development of spatial representation and reasoning*. MIT Press.

Newhouse, P., Newhouse, C., & Astur, R. S. (2007). Sex differences in visual-spatial learning using a virtual water maze in pre-pubertal children. *Behavioural Brain Research*, *183*, 1–7.

Pazzaglia, F., & De Beni, R. (2001). Strategies of processing spatial information in survey and landmark-centred individuals. *European Journal of Cognitive Psychology*, *13*, 493–508.

Piaget, J., & Inhelder, B. (1948). *La representation de l'espace chez l'enfant* [The child's conception of space]. Paris: Presses Universitaires de France.

Piaget, J., & Inhelder, B. (1967). *The coordination of perspectives. In the child's concept of space* (pp. 209–246). New York: Norton.

Quaiser-Pohl, C., Lehmann, W., & Eid, M. (2004). The relationship between spatial abilities and representations of large-scale space in children—A structural equation modeling analysis. *Personality and Individual Differences*, *36*, 95–107.

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ribordy, F., Jabès, A., Lavenex, P. B., & Lavenex, P. (2013). Development of allocentric spatial memory abilities in children from 18 months to 5 years of age. *Cognitive Psychology*, *66*, 1–29.

Rovine, M. J., & Weisman, G. D. (1989). Sketch-map variables as predictors of way-finding performance. *Journal of Environmental Psychology*, *9*, 217–232.

Ruggiero, G., D'Errico, O., & Iachini, T. (2016). Development of egocentric and allocentric spatial representations from childhood to elderly age. *Psychological Research*, *80*, 259–272.

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.

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research*, *6*, 351–360.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior*, *10*, 9–55.

Signorella, M. L., & Jamison, W. (1978). Sex differences in the correlations among field dependence, spatial ability, sex role orientation, and performance on Piaget's water-level task. *Developmental Psychology*, *14*, 689.

Silverman, I., Choi, J., Mackewn, A., Fisher, M., Moro, J., & Olshansky, E. (2000). Evolved mechanisms underlying wayfinding: Further studies on the hunter-gatherer theory of spatial sex differences. *Evolution and Human Behavior*, *21*, 201–213.

Theodorsson-Norheim, E. (1986). Kruskal-Wallis test: BASIC computer program to perform nonparametric one-way analysis of variance and multiple comparisons on ranks of several independent samples. *Computer Methods and Programs in Biomedicine*, *23*, 57–62.

Thomas, H., & Turner, G. F. (1991). Individual differences and development in water-level task performance. *Journal of Experimental Child Psychology*, *51*, 171–194.

Tversky, B. (2000). Levels and structure of spatial knowledge. In R. Kitchin and S. Freundschuh (eds.), *Cognitive Mapping – Past, present and future* (pp. 24–43). London: Routledge.

Vashro, L., & Cashdan, E. (2015). Spatial cognition, mobility, and reproductive success in northwestern Namibia. *Evolution and Human Behavior*, *36*, 123–129.

Vasta, R., Knott, J. A., & Gaze, C. E. (1996). Can spatial training erase the gender differences on the water-level task? *Psychology of Women Quarterly*, *20*, 549–567.