Characterizing spatial construction processes: Toward computational tools to understand cognition

Cathryn S. Cortesa (ccortes4@jhu.edu)
Department of Cognitive Science, 3400 N. Charles St. Baltimore, MD 21218

Jonathan D. Jones (jdjones@jhu.edu)
Department of Electrical and Computer Engineering, 3400 N. Charles St. Baltimore, MD 21218

Gregory D. Hager (hager@cs.jhu.edu)
Department of Computer Science, 3400 N. Charles St. Baltimore, MD 21218

Sanjeev Khudanpur (khudanpur@jhu.edu)
Department of Electrical and Computer Engineering, 3400 N. Charles St. Baltimore, MD 21218

Amy L. Shelton (ashelton@jhu.edu)
School of Education, 2800 N. Charles St. Baltimore, MD 21218

Barbara Landau (landau@jhu.edu)
Department of Cognitive Science, 3400 N. Charles St. Baltimore, MD 21218

Abstract

Spatial construction—creating or copying spatial arrangements—is a hallmark of human spatial cognition. Spatial construction appears early in development, predicts later spatial and mathematical skills, and is used throughout life. Despite its importance, we know little about the cognitive processes underlying skilled construction. Construction tasks are highly complex but analyses have tended to focus on broad-stroke measures of end-goal accuracy. In this paper we introduce a novel behavioral coding formalism to characterize an individual's *entire construction process*, examine many individuals' processes in aggregate, and summarize patterns that emerge. The results show high consistency at certain points occurring throughout the construction, but also indicate flexibility in the interim paths that lead to and diverge from these points. Our approach offers a new method that can more precisely describe the behavioral patterns observed during construction in order to reveal the underlying cognitive processes engaged, and capture individual differences in building expertise.

Keywords: spatial skills; spatial cognition; block copying; computational model

Introduction

Spatial construction—the activity of creating novel spatial arrangements or copying existing ones—is a hallmark of human spatial cognition. These activities naturally occur during childhood and adolescence and are related to later achievements in science, technology, engineering, and mathematics (STEM) fields (Hsi, Linn, & Bell, 1997; Kell, Lubinski, Benbow, & Steiger, 2013; Verdine, Golinkoff, Hirsh-Pasek, Newcombe, et al., 2014). Moreover, spatial play during early schooling—including spatial building tasks—contributes to school readiness (Verdine, Golinkoff, Hirsh-Pasek, & Newcombe, 2014; Wai, Lubinski, & Benbow, 2009), developmental of logico-mathematical abilities (Casey et al., 2008; Cheng & Mix, 2012; Nath & Szücs, 2014), and math performance in middle and high school (Stannard, Wolfgang, Jones, & Phelps, 2001; Wolfgang, Stannard, & Jones, 2003).

Despite the importance of spatial construction skills, little is known about the cognitive processes underlying their origins and development. Part of the reason for this is that spatial construction skills are highly complex, yet the cognitive characterization of these skills and their measurement has been quite limited. For example, although evaluation of block construction tasks has long been recognized as an important assessment of spatial skills (Bailey, 1933), most methods of assessment only evaluate the end product (accuracy), and fail to measure the construction process. Studies have generally reported broad stroke outcome measures such as time to complete a structure (Akshoomoff & Stiles, 1996; Frick, Hansen, & Newcombe, 2013), binary measures of block placement as correct or incorrect (Brosnan, 1998; Hoffman, Landau, & Pagani, 2003; Stiles & Stern, 2001), or summary ratings for the complexity, planning, or organization of free-play block designs (Caldera et al., 1999; Casey & Bobb, 2003; Stiles-Davis, 1988; Stiles & Stern, 2001). Even studies that aim to characterize development of construction processes or strategies have used analytic categories that are limited in their generality for understanding construction. For example, some have suggested that children start with simple iterative methods (i.e. stacking blocks on top of one another), then move to sequential combinations of methods (i.e. first creating a line of blocks next to one another, then creating a stack), and finally come to flexibly shift between multiple methods (Stiles-Davis, 1988; Stiles & Stern, 2001; Stiles, Stern, Trauner, & Nass, 1996). These characterizations tell us little about the step-by-step processes that the user takes when carrying out a complex construction, nor how the ever-expanding set of outcomes grows over time.

More recent studies have attempted to provide a more precise characterization of the process occurring during construction. Verdine and colleagues characterized children's placement errors, including whether a block was placed in the correct layer, in correct orientation relative to other blocks, and with the correct attachment studs connected (Verdine, Golinkoff, Hirsh-Pasek, Newcombe, et al., 2014; Verdine, Golinkoff, Hirsh-Pasek, & Newcombe, 2016). Researchers in computer science have generated step-by-step instructions for assembling block models based on physical constraints such as avoiding 'floating blocks' not supported from below (Zhang, Igarashi, Kanamori, & Mitani, 2016). Both studies begin to characterize the temporal and incremental nature of block construction.

Each of the approaches discussed above provides a description of the accuracy of a block construction at points intermediate to building or at the end; but none provides a characterization of an individual's complete construction process. Yet, variability and/or consistency across individuals' construction processes may reveal much about the underlying cognitive and perceptual abilities and biases that influence the builder's construction choices.

The incremental process of adding blocks to a structure can unfold in many ways, with different strategies leading to the same successful solution. Some of this variation may be unimportant—merely small tweaks in the options one can use to complete a construction. Other aspects of variation are likely to reflect important cognitive processes. For example, limitations of attention and memory make it likely that certain strategies or processes will be preferred as they may reflect more efficient use of available cognitive resources (Ballard, Hayhoe, Pook, & Rao, 1997). Certain strategies may also reflect the builder's understanding of the physical principles engaged during building. For example, the effect of gravity could bias the builder to construct from the bottom layer upwards (Zhang et al., 2016). Finally, construction strategies may be related to perceptual or semantic groupings of the blocks within the structure. The sub-parts which the builder chooses to construct, and the order in which they are created may be driven by the builder's perceptual parsing of the model being copied. More generally, there may be systematic commonalities in the construction paths that builders use, and these may vary depending on the builder's level of skill.

Understanding the principles underlying construction requires methods that can characterize the builder's full construction process. Ideally, the best analysis would completely describe the entire construction process, capturing any imaginable construction outcome as well as each step of building along the way. This kind of characterization would be as relevant for a simple stack of blocks as it would be for an elaborate castle or an abstract collection of connected pieces.

To our knowledge, such methods have never been reported. Therefore, in this paper, we report a new method for characterizing the precise nature of processes involved when people carry out a relatively simple construction task: using blocks to copy a target model. To do this, we ask adults to carry out a simple building task, using a set of Duplos™. We describe our new method for coding block construction behavior that uses a novel computer interface. Our method characterizes each partial assembly created during the building process as a step taken along a path from the start to the end of construction. We then evaluate common states and predominant path types traversed by adults as they move through the construction process. Finally, we make inferences about the underlying cognitive mechanisms engaged during block construction.

Method

Participants

Twenty-seven healthy adults 18-53 years old (*M* = 21;4, *SD* = 6;6) participated in the study. A university ethical review board approved the study's procedures, and all participants provided informed consent.

Materials

Participants were asked to copy six different block models of varying size, each consisting of 4, 6, or 8 blocks. Each participant copied each of the six models in randomized order, but always began with the two smallest models (models 1 and 2). Figure 1 shows each of the six models.

We used Duplo™ blocks for the construction copy task. These blocks were chosen for several reasons. First, the attachment mechanism allowed the blocks to be connected to each other in fixed ways. The attachment studs permitted precise specification of the relationships between blocks above and beside one another. In addition, the limited set of colors (red, yellow, green, blue) of each shape (2x2 square, 4x2 rectangle) were ideal for the precise measurement system we developed.

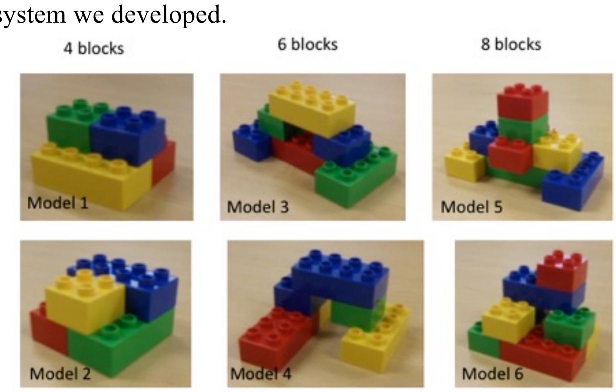

Figure 1: The block models used in this study. Models 1 and 2 contain four blocks, models 3 and 4 contain six blocks, and models 5 and 6 contain 8 blocks.

We mounted a PrimeSense Carmine RGBD camera in an overhead configuration to record participants' behaviors as they carried out the construction task, at a rate of 30 frames per second. All videos were coded using our annotation interface. The coder viewed the video recording frame-by-frame on a desk-top computer.

Procedures

Participants were seated at a table marked with an outline of a rectangular area (14.75 x 24.00 in.) in which they completed their block construction copy. During data collection, the experimenter observed participants in real time on a video display monitor. A vertical black barrier was placed on the table behind the construction area to obscure the video display monitors from the participant and to avoid distraction. Figure 2 shows the testing equipment setup used for the study.

In the procedure, the experimenter first placed the model at a 45° angle in the rear left corner of the marked construction area on the table. Each model was presented in a standardized orientation so that the greatest number of model surfaces were visible to the participant. Then, the experimenter placed the corresponding loose blocks on the table in the center of the construction area by emptying them from a small bag. This ensured random starting positions for each of the blocks used to construct the copy. Participants were instructed to take their time and to copy the model, building as efficiently and accurately as possible.

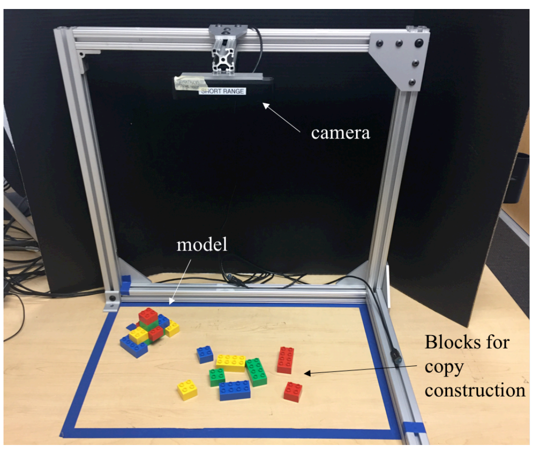

Figure 2: Overhead camera and blocks set up for the block copying task. The model was placed at the rear left of the table, and blocks for the copy construction were placed at the center.

Analytic Rationale

To account for the broad range of construction behaviors and resulting complex patterns in the copy, we developed a new behavioral coding system, executed in a custom designed computer interface. The video frames for each trial from each participant were coded as a series of *actions*, each of which culminated in a *state*. Each action captured the start and end time of a change made to the copy as it was being constructed. Actions could be comprised of a single relationship, such as placing two blocks adjacent on the table. Other actions included a complex set of simultaneous relationships such as adding a single block in a location that was both above and beside other blocks. Actions could be constructive, such as adding a block or connecting multiple parts, or deconstructive, such as removing a block or separating a structure into two parts.

Each relationship was defined specifically by the set of attachment studs involved. For example, if two rectangle blocks were placed horizontally adjacent to each other along the principal (long) axis, then four studs on each block would be involved in the adjacency. Alternatively, if they were attached along the secondary (short) axis, only two studs on each block would be implicated. Block studs were identified according to the column and row on each block, so the coded data specified the exact relationship between sets of two blocks.

Each action modified the environment to result in a new block *state*, defined as a set of block attachments present in the copy. Since the construction process occurred over time, each action included a time stamp that allowed block states to be ordered. Here, we refer to the ordered sequence of states over time as a *construction path*, where actions represent transitions connecting one state to the next. To illustrate, Figure 3a shows six observed states (illustrated as images of block configurations) and 11 observed actions (directed arrows). Any set of arrows that lead from the first null state to the final correct copy state could comprise a construction path.

Data analysis

One researcher coded all videos. The initial state of the model was always a null state in which no blocks were connected. Each other state along the path to the final copy was attained via a constructive or deconstructive action taken at the preceding state. Since each participant could take any number of actions, construction path length was not balanced across individuals. We also used the coded data to count the number of unique state transitions for all participants in aggregate. Results of the analysis are described below. Principles of the results are true across all six models, but we illustrate using two models as examples.

Results

Overwhelmingly, the most common actions were correct single-block placements over time. Participants tended to take efficient paths that traversed an average of just over n-1 states for a model that contains n blocks. This held true for all six of the models, including the four-block models 1 and 2 (M = 3.3, SD = 0.9 and M = 3.1, SD = 0.6, respectively), the six-block-models 3 and 4 (M = 5.0, SD = 0.7, and M = 5.5, SD = 1.8, respectively), and the eight-block-models 5 and 6 (M = 7.4, SD = 0.9, and M = 7.9, SD = 1.7, respectively). Strikingly, the observed correct states represented only a small proportion of total *possible* correct states. For example, for the six-block model 3 (shown in

Figure 3b), 79 possible correct states exist[1]. In aggregate, our sample executed a total of 136 actions, but only created 16 different correct states (27% of all possible correct states). An additional three erroneous states were observed in model 3; these will be discussed later.

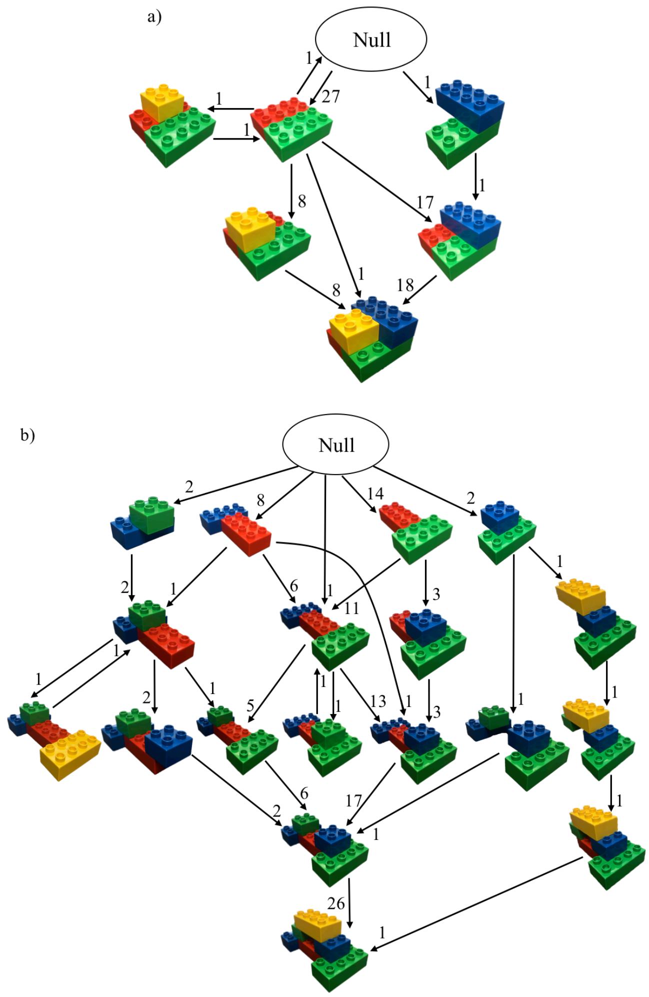

Figure 3: Observed paths for models 2 (a) and 3 (b). Paths begin at the top, where the null state represents no blocks connected. Images represent block states; arrows connecting images represent actions. Numbers adjacent to arrows represent the number of times that action was executed.

Of the 16 observed states for model 3 (Fig 3b), some were created by a majority of participants, while others were rare. We found the same pattern in the observed paths, that is, the actions moving from state to state. Though all observed paths led to a correct copy construction, some were highly likely, while others were highly unlikely. For example, the first image in the fourth row in Figure 3b was created by a great majority of participants (96.3%). The most common path to achieve this state led from the fifth image in the third row of Figure 3b, such that 17 of the 26 (65.4%) people who created the same penultimate state achieved it by placing the green square on the second layer.

The most commonly traversed states created points of convergence. Convergence points represented a single subassembly that results from several different preceding actions. We observe that convergence points tend to be highly likely states, which most or all participants created along the way to a complete construction. As shown in Figure 3b, most participants (66%) passed through the state in which three blocks are joined with horizontal adjacency to create the base of the copy (second image, second row). The observed frequencies of convergence points are remarkable when one considers that only a fraction of all possible efficient paths go through these states.

We also observed points of divergence. Divergence points represented cases where participants, when presented with identical partial assembly states, chose to proceed with several different actions. For example, about 70% of those who created the base of model 3b proceeded to place the green then the blue square in the second layer. The other 30% instead placed the same two blocks in the opposite order, first placing the blue and then the green square in the second layer. This is illustrated in Figure 3b, in the third and fifth images of the third row.

Our results demonstrated some commonalities across the six different models. The most frequently constructed partial assembly states for all six of the models represented a complete layer. Across all six models, 83.6% of participants began their copy construction by creating the base layer. For the models with six or eight blocks, 77.4% and 86.5% of participants created the complete second layer as a partial assembly, respectively. Across all six models, each complete layer state is visited more frequently than would be expected by chance, even with the most conservative comparison against only other observed states with the same number of blocks (all *p's* $< .001$).

Many participants' construction paths (75.5%) traversed all complete layer states, although this is by no means necessary in order to achieve a correct copy. Specifically for the most complex model, model 6, those participants who traversed each complete layer partial assembly state in their individual construction path also tended to have the shorter path lengths ($t(24) = -2.57$, $p = .017$). In other words, when faced with a complex block copying task, building layer by layer appears to be both highly likely and highly efficient. These observations provide insight into the importance of layers, which may be driven by the builder's understanding of physical properties such as gravity and/or perceptual biases that suggest a natural parse in terms of layers.

Although most block placements were correct (i.e. replicated part of the model in the copy), there were some errors—that is, states that did not represent a correct part of the model. These errors contributed to deviations from the main construction paths. If erroneous states are included in our calculation, the number of possible block states in given

[1] For a given model with *n* blocks, we define the set of possible correct states as the set of unique states traversed along any path which begins at the null state, ends at the goal state, and contains *n* states. These states can be enumerated computationally.

a starting set of 4, 6, or 8 blocks is vast, but finite. For example, a mathematician recently estimated that there are nearly a billion possible ways to connect six uniform rectangular Lego™ blocks contiguously (Abrahamsen & Eilers, 2011). The model and our instructions to participants constrained their behavior such that even though errors occurred, only an extremely small proportion of all possible states were observed.

The errors observed in this sample also provided insight into the cognitive limitations of our adult participants. Two categories of errors were observed. First, spatial errors occurred when a participant utilized the correct block, but placed it in incorrect orientation relative to the rest of the copy. For example, in model 2, one participant placed the yellow square block with incorrect relationship relative to the red and green rectangles underneath, shown in the first image of the first row in Figure 3a. Second, block identity errors occurred when a participant created the correct form in their copy, but used the wrong color block relative to the model. For example, for model 3, one participant used the yellow rectangle to create the base of their copy instead of the green rectangle (first image, first row of Figure 3b).

Overall, our results provided rich detail about the step-by-step process undertaken by our adult participants in the block copying task. We observed only a small portion of all possible correct states, and a yet smaller portion of all possible states including erroneous ones. The distribution of the sample across different construction paths was not uniform, but rather demonstrated commonalities across the six models. Specifically, convergence points were observed corresponding to completed copy layers, and divergence was observed in the order of block placement within a single layer. Most common construction paths involved the sequential construction of horizontal layers, beginning with the base and building upward.

Discussion

Our study presented a precise, quantitative method for understanding how people carry out a simple block construction task. Using a novel behavioral coding method together with computational modelling, we precisely described the block construction process as a temporal sequence of states. This approach shed light on the cognitive processes that support spatial construction tasks.

A description of state transitions illustrated commonalities among the construction paths that participants used for each of the six models. Convergence points tended to correspond to the completion of a horizontal layer in the model, while divergence points tended to correspond to various orders of placing blocks within a layer. We hypothesize that convergence points can be interpreted as boundaries between perceptual or semantic chunks—that is, they represent sub-goals that builders had in mind as they approached and carried out the task. Although we did not provide any pre-determined conceptual units or clear perceptually-based chunks (such as sub-parts built from same-colored blocks), participants nonetheless created these chunks in systematic ways. The location of convergence points, for example, at the completion of a horizontal layer, may indicate that participants grouped or chunked the models principally into horizontal layers or “floors”.

It is likely that the underlying structure of sub-goals will vary substantially, depending on a variety of factors. For example, a model that is organized to highlight salient perceptual units, such as multiple vertically adjacent blocks of the same color, could induce a construction path that would take most builders through a convergence point organized as a vertical chunk, and not the horizontal layers observed in the present study. In this case, we anticipate that adults would attend to the imposed perceptual units and change their construction strategy to build using sub-goals defined by these color-units. Similarly, incorporating conceptual structure into the models could radically alter people’s construction paths—heads and eyes on structures that look like animals, or wheels on structures that look like vehicles could serve as the chunks or sub-goals to be built. The role of conceptual knowledge in the reproduction of complex figures has a long history in the domain of chess, where experts are known to reproduce board configurations using sub-structures that reflect high-level concepts such as attack and defend (Chase & Simon, 1973).

In our simple construction task, errors were relatively rare. Errors were characterized as either spatial relationship or block identity errors. We hypothesize that spatial errors indicate problems with spatial working memory, in translating information observed in the model into the working copy. Block identity errors, on the other hand, may involve object working memory, or a prioritization of spatial configuration over color information. These error types are likely linked to the relative simplicity of the models we used; analysis of error patterns for more complex models may well reveal more variation.

We see our method as a powerful way to examine the nature of sub-goals and errors, applicable to a variety of visual-spatial construction tasks involving conceptual or perceptual chunks. The extent to which observed construction paths and construction errors change over variations in the target model would provide insight into how building principles change across target types. In addition, our method permits evaluation of variation in construction paths across different participant populations including construction experts compared to novices, and developmental populations of children at different ages.

We believe that our analytic method has great potential for revealing the fine-grained nature of many tasks that require step-by-step actions, which in turn require rich cognitive capabilities, including representation of the goal as well as strategies for moving from a start state to an end state. Such general task requirements are ubiquitous throughout life—from the toddler who learns to operate an iPad to the adult who learns to cook a gourmet meal. Our insight is that understanding complex skills requires a fine-grained and precise approach, exemplified by the method we have introduced. Block construction serves as a first demonstration of the utility of our approach, but it is by no means the end.

Acknowledgments

This work was supported by NSF grant #1561278.

References

Abrahamsen, M., & Eilers, S. (2011). On the asymptotic enumeration of Lego structures. *Experimental Mathematics, 20*, 145-152. https://doi.org/10.1080/10586458.2011.564539

Akshoomoff, N. A., & Stiles, J. (1996). The influence of pattern type on children's block design performance. *Journal of the International Neuropsychological Society*, *2*, 392–402.

Bailey, M. W. (1933). A scale of block constructions for young children. *Child Development*, *4*(2), 121–139. https://doi.org/10.2307/1125591

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723–767.

Brosnan, M. J. (1998). Spatial ability in children's play with Lego blocks. *Perceptual and Motor Skills*, *87*, 19–28.

Caldera, M., Culp, A., O Brien, M., Tuglio, R., Alvarez, M., & Huston, A. (1999). Children's play preferences, construction play with blocks, and visual-spatial skills: Are they related? *International Journal of Behavioral Development*, *23*(4), 855–872. https://doi.org/10.1080/016502599383577

Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cognition and Instruction*, *26*(3), 269–309. https://doi.org/10.1080/07370000802177177

Casey, B. M., & Bobb, B. (2003). The power of block building. *Teaching Children Mathematics*, *10*(2), 98–102.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Cheng, Y. L., & Mix, K. S. (2012). Spatial training improves children's mathematics ability. *Journal of Cognition and Development*, *15*(1), 2–11. https://doi.org/10.1080/15248372.2012.725186

Frick, A., Hansen, M. A., & Newcombe, N. S. (2013). Development of mental rotation in 3- to 5-year-old children. *Cognitive Development*, *28*(4), 386–399. https://doi.org/10.1016/j.cogdev.2013.06.002

Hoffman, J. E., Landau, B., & Pagani, B. (2003). Spatial breakdown in spatial construction: Evidence from eye fixations in children with Williams syndrome. *Cognitive Psychology*, *46*, 260–301. https://doi.org/10.1016/S0010-0285(02)00518-2

Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education*, *86*(2), 151–158. https://doi.org/10.1002/j.2168-9830.1997.tb00278.x

Kell, H. J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2013). Creativity and technical innovation: Spatial ability's unique role. *Psychological Science*, *24*(9), 1831–1836. https://doi.org/10.1177/0956797613478615

Nath, S., & Szücs, D. (2014). Construction play and cognitive skills associated with the development of mathematical abilities in 7-year-old children. *Learning and Instruction*, *32*, 73–80. https://doi.org/10.1016/j.learninstruc.2014.01.006

Stannard, L., Wolfgang, C. H., Jones, I., & Phelps, P. (2001). A longitudinal study of the predictive relations among construction play and mathematical achievement. *Early Child Development and Care*, *167*(1), 115–125. https://doi.org/10.1080/0300443011670110

Stiles-Davis, J. (1988). Developmental change in young children's spatial grouping activity. *Developmental Psychology*, *24*(4), 522–531. https://doi.org/10.1037//0012-1649.24.4.522

Stiles, J., & Stern, C. (2001). Developmental change in spatial cognitive processing: Complexity effects and block construction performance in preschool children. *Journal of Cognition and Development*, *2*(2), 157–187.

Stiles, J., Stern, C., Trauner, D., & Nass, R. (1996). Developmental change in spatial grouping activity among children with early focal brain injury: Evidence from a modeling task. *Brain and Cognition*, *31*(1), 46–62. https://doi.org/10.1006/brcg.1996.0024

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2014). Finding the missing piece: Blocks, puzzles, and shapes fuel school readiness. *Trends in Neuroscience and Education*, *3*(1), 7–13. https://doi.org/10.1016/j.tine.2014.02.005

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2016). Links between spatial and mathematical skills across the preschool years. *Monographs of the Society for Research in Child Development*.

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematics skills. *Child Development*, *85*(3), 1062–1076. https://doi.org/10.1111/cdev.12165.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817–835. https://doi.org/10.1037/a0016127

Wolfgang, C. H., Stannard, L., & Jones, I. (2003). Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child Development and Care*, *173*(3), 467–475. https://doi.org/10.1080/0256854010959495

Zhang, M., Igarashi, Y., Kanamori, Y., & Mitani, J. (2016). Component-based building instructions for block assembly. In *Computer-Aided Design and Applications* (pp. 55–59). Vancouver, Canada: Taylor & Francis. https://doi.org/10.1080/16864360.2016.1240450