Thinking in Pictures: A Fresh Look at Cognition in Autism

Maithilee Kunda (mkunda@cc.gatech.edu)
Ashok K. Goel (goel@cc.gatech.edu)
School of Interactive Computing, Georgia Institute of Technology
85 Fifth Street NW, Atlanta, Georgia 30308, USA

Abstract
In this paper, we develop a cognitive account of autism centered around a reliance on pictorial representations. This Thinking in Pictures hypothesis, inspired by the book of the same name by Temple Grandin, shows significant potential for explaining many autistic behaviors. We support this hypothesis with empirical evidence from several independent cognitive and neuroimaging studies of individuals with autism, each of which shows strong bias towards visual representations and activity. We also examine three other cognitive theories of autism—Mindblindness, Weak Central Coherence, and Executive Dysfunction—and show how Thinking in Pictures provides a deeper explanation for several results typically cited in support of these theories.

Keywords: Autism; cognitive theory; mental imagery; pictorial representation; visual reasoning.

Introduction
Ever since the discoveries of autism and Asperger’s syndrome in the 1940s by physicians Leo Kanner and Hans Asperger, autism spectrum disorders (ASDs) have been defined by the atypical behaviors that they produce. In particular, ASDs (referred to as “autism” for simplicity) are developmental conditions characterized by atypical social interactions, communication skills, and patterns of behavior and interests, as described in the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR, 2000).

While the specific causes of autism are not known, an etiological framework, shown in Figure 1, has been traced out that leads from genetic and possibly environmental factors, through neurobiological development and cognitive functioning, and finally to behavioral manifestations (adapted from Minshew & Goldstein, 1998).

Many theories have attempted to give a cogent account of the changes in cognitive functioning that lead to the behavioral characteristics of autism. Some prominent theories include: Mindblindness, which hypothesizes that individuals with autism lack a “theory of mind,” i.e. they cannot ascribe mental beliefs to other people (Baron-Cohen, 1995); Weak Central Coherence, which posits a bias towards local instead of global information processing (Happé & Frith, 2006); and Executive Dysfunction, which suggests that individuals with autism have deficits in executive functions such as planning, mental flexibility, and inhibition (Russell, 1998).

However, many individuals on the autism spectrum have given quite different introspective descriptions of their cognitive processes. One of the most famous is the account by Temple Grandin in her book Thinking in Pictures (2006). Grandin, a high-functioning adult with autism, states that her mental representations are predominantly visual, i.e. that she thinks in pictures, and that this representational bias affects how she performs a range of cognitive operations, from conceptual categorization to the interpretation of complex social cues. Numerous other individuals with autism have also informally reported becoming aware of similar biases in mental representation, suggesting that Grandin is not an isolated case.

While Grandin’s account of visual thinking has been primarily an introspective study, we aim to show that the Thinking in Pictures hypothesis does, in fact, represent a very powerful way to look at cognition in autism.

We begin by considering what it might mean to think in pictures and how this would differ from typical cognition. Second, we explore how Thinking in Pictures could account for many characteristic behaviors of autism. Third, we present relevant empirical data from a range of literature, including cognitive and neurobiological studies of individuals with autism. Fourth, we examine how Thinking in Pictures relates to other cognitive theories of autism.

What Does It Mean to Think in Pictures?
Developing a Thinking in Pictures account of autism raises many hard issues. Do individuals with autism actually think in pictures, or do they just think they do? If they actually do, does it mean they are incapable of representing verbal
information or just biased against it? To what degree do they abstract information from their visual representations, and what kinds of visual representations do they use?

We do not intend to answer all of these questions here. Instead, we propose one simple, minimal characterization of Thinking in Pictures, leaving open many details, and examine what conclusions can be drawn. This formulation does not make for a complete cognitive model, nor is it the only way to define a valid Thinking in Pictures account.

We begin with two basic assumptions. First, we assume that typical cognition utilizes both pictorial and amodal representations and reasoning. Second, we assume that the Thinking in Pictures cognitive account is characterized as using mainly pictorial and other modal representations.

Modal representations are close to perceptual modalities, such as vision, while amodal representations are removed from perception. At one end of this representational spectrum, illustrated in Figure 2, are analogical, imagistic or pictorial representations that preserve topology and have a one-to-one structural correspondence with the world, but no abstractions of the world (Glasgow & Papadias, 1992). At the other end are amodal representations that contain only abstractions of the world, e.g. in the form of propositions. Along this modal-amodal spectrum are various descriptive representations in the forms of symbols. Descriptive representations close to the modal end of the spectrum are perceptually grounded, as in perceptual symbol systems (Barsalou, 1999).

Figure 2: Representational spectrum of a cat, illustrating progression from modal to amodal representations.

Pictorial representations, which lie towards the modal end of the spectrum, capture quantitative, or metric, information; amodal representations are especially good at expressing non-metric relations (Chandrasekaran, 2006). For example, a pictorial representation of a cat consists of considerable metric information about the cat, such as the relative sizing of her body parts, their spatial layout, etc. (Kosslyn, 1994). In contrast, amodal representations such as the word “cat” or “behind (tail, body)” contain no metric information about the cat (even though spatial relationships like “behind” can be represented).

As a result, in pictorial representations, abstract, non-metric concepts such as the function of an object, the causal relations between events, or the intentionality of an agent are, at best, implicit.

A bias toward pictorial representations also affects knowledge organization, in particular the formation of conceptual categories. Consider some basic types of categories: prototype, exemplar, and qualia (Thagard, 2005). Qualia-based categories can be supported using pictorial representations, since they are based on perceptual similarity. Exemplar categories, which consist of examples of a particular concept, are also supported since no further abstraction is required. However, prototype categories require abstractions from examples to form the uninstantiated prototype that defines the basic category.

Finally, some basic inferences in typical cognition inherently use pictorial representations, such as the mental rotation of images (Shepard & Cooper, 1982). While other types of inference rely on amodal representations, such as propositional logic, alternate methods for some of these inferences, or variations of them, could exist using pictorial representations. For example, many theories of analogical reasoning focus on amodal matching, where syntactic relationships among the various components of a concept (often representing causal relations) are used to establish the similarity between two cases (Gentner, 1983). However, it has been shown that visual knowledge alone is sufficient for analogical transfer (Davies & Goel, 2008).

**Thinking in Pictures and Autism**

The definition of autism is centered on three areas: social interaction, communication, and stereotyped patterns of behavior and interests. Each of the following paragraphs addresses a subset of the behaviors in each area, which are listed in the *DSM-IV-TR* (2000).

Atypical social behaviors of autism include a lack of seeking to share enjoyment with others and a lack of social or emotional reciprocity. Both of these types of behavior rely on an ability to infer the mental states of others, which is a highly abstract concept that cannot easily be represented pictorially. Without this concept, individuals with autism would have difficulty in desiring to induce certain mental states in others, like enjoyment, and in reciprocating or even perceiving emotional or social intentions.

Communication issues in autism include the delayed development or inappropriate use of language along with deficits in imaginary play. Thinking in Pictures explicitly allows for problems in verbal language development from a bias against amodal representations. Regarding the latter, it has been shown that symbolic play in typically developing children evolves from using objects that share perceptual similarity with the target representation to objects that are perceptually dissimilar (Ungerer et al., 1981), suggesting a progression from play that is perceptually grounded to play that is free from perceptual constraints. Accordingly, imaginary play deficits in autism could be explained by a bias towards perceptually grounded representations.

Finally, autism is also characterized by stereotyped patterns of behavior and interests, such as a preoccupation with parts of objects or an adherence to nonfunctional routines. Function, as discussed previously, is a qualitative
concept not well-suited to pictorial representations. Without functional interpretations, object use by children with autism could remain centered on visual or other perceptual features, within the sensorimotor-based frameworks of early developmental play (Fenson et al., 1976). Several studies have indeed shown less (Stone et al., 1990) or less complex (Williams, Reddy, & Costall, 2001) functional play in children with autism compared to their typically developing peers. Also, as noted previously, pictorial representations do not provide for the explicit representation of causality, which could lead to nonfunctional routines, for instance if routines were structured temporally instead of causally.

Evidence for Thinking in Pictures

Knowledge Representation

We discuss the results of four experiments that examined the uses of pictorial versus verbal (amodal) representations in individuals with autism as compared to typically developing individuals. First, in studies of picture and word recall tasks in the typical population, it has been consistently found that more pictures are recalled than words. One prominent explanation of this effect is dual-coding theory, which posits that pictures are encoded both pictorially and verbally whereas words are encoded only verbally, and that the “dual encoding” of pictures aids in their recall (Paivio, 1991). One study of this task in children found that the control group exhibited a significant picture-superiority effect of this kind, as expected, but the same effect was not statistically significant in the autism group (Whitehouse, Maybery, & Durkin, 2006), which the authors of that study suggest results from the children with autism not benefiting from a second, verbal encoding of the pictures.

Second, in the typical population, the recall of short words is generally better than the recall of longer words, but this effect can be eliminated by articulatory suppression, suggesting that verbal encoding is used to some extent (Cowan et al., 2003). Furthermore, this effect is still seen if pictures are used instead of words, but only in subjects of a certain age, starting around 7 to 8 years old (Hitch et al., 1989). In one study of this type of picture recall task in children (Whitehouse, Maybery, & Durkin, 2006), the pictures had either long or short labels, and subjects were asked to either remain silent or verbalize each label. As expected, the control group performed significantly better with the short labels than with the long labels, and whether they verbalized the labels had no effect. In contrast, the autism group exhibited a much smaller word-length effect overall, and the effect was smaller in the silent condition than in the verbalizing condition. These results suggest that the children with autism verbally encoded the pictures to a lesser extent than did the control group, and also that their use of verbal encodings increased when prompted to verbalize the labels.

The third experiment looked at the effects of articulatory suppression on a task-switching test (Whitehouse, Maybery, & Durkin, 2006). Children were given a sequence of pairs of numbers to add or subtract alternately, and they either had to remain silent or to repeat “Monday” as a form of articulatory suppression (AS). The control group performed far better when they were silent than under AS. However, the autism group showed no difference between the silent and AS conditions, suggesting that they did not use verbal representations to guide their task-switching.

The fourth experiment looked at a word-completion task in which semantic priming was provided using either picture or word cues (Kamio & Toichi, 2000). The control group performed similarly under both conditions, but the autism group performed much better with picture cues than word cues. This suggests that the individuals with autism were better able to retrieve verbal information through pictorial representations than through other verbal representations.

Together, results from these four experiments suggest that individuals with autism are biased towards using pictorial representations over verbal ones in several tasks. They seem to encode pictures pictorially until required to produce a verbal representation and also to rely more on pictorial representations than verbal ones for recall, task-switching, and semantic retrieval.

Knowledge Organization

As discussed earlier, prototype-based categories cannot be instantiated easily using pictorial representations. Several studies show evidence for limitations in the use of prototype categories in autism. In one study, children with autism who successfully applied rule-based categorization on one task were unable to apply prototype-based categorization on another (Klinger & Dawson, 2001). Individuals with autism were also shown under a variety of tasks to exhibit fairly typical abilities in concept identification but significantly lower abilities in concept formation, which relies on the use of prototypes (Minshew, Meyer, & Goldstein, 2002). Another study showed a preference in individuals with autism for categorizing objects based on physical attributes instead of on more abstract qualities, namely for sorting books by color instead of by genre as typically developing individuals did (Ropar et al., 2007).

Visual Attention and Reasoning

Much empirical evidence has shown that individuals with autism are adept at certain tasks relying on pictorial modes of processing. One such task is the Embedded Figures Task (EFT), in which a small, simple shape must be found within a larger figure. Numerous studies have shown that individuals with autism are often more accurate (Shah & Frith, 1983) or more efficient (Joliffe & Baron-Cohen, 1997) on the EFT than typically developing individuals.

Recent studies have looked at another visual search task in which a target must be found amid a group of distracters that share either shape alone (feature search) or shape or color (conjunctive search) (Plaisted, O’Riordan, & Baron-Cohen, 1998; O’Riordan et al., 2001). Results showed that individuals with autism had significantly faster search times than typically developing individuals and, unlike the control
group, had the same search times under both feature and conjunctive search conditions. Even more unusual were findings that, while typically developing individuals showed a characteristic linear increase in search time as the number of distractors increased in conjunctive search, the increase in search times for the autism group remained fairly flat.

These results suggest that individuals with autism might be using fundamentally different search strategies than the typical population. While the view of Thinking in Pictures presented in this paper does not explicitly deal with visual attention, these results could be explained by attentional strategies that are not mediated by verbal representations, which might prove more efficient for visual tasks.

Finally, while many of these types of studies cast their findings as evidence for isolated skills or “islets of ability” in individuals with autism, recent research has suggested that, given the opportunity to reason pictorially, individuals with autism can exhibit significantly higher measures of general intelligence than shown on standard tests.

In particular, a study conducted using groups of both children and adults with autism demonstrated that their performance on Raven’s Progressive Matrices fell into dramatically higher percentile ranges than their performance on Wechsler scales, a discrepancy not seen in the typically developing control groups (Dawson et al., 2007). In fact, whereas a third of the children with autism fell into the mentally retarded range on the Wechsler scales, only 5 percent did so on the Raven’s test, with a third of them scoring at the 90th percentile or higher (as compared to none in this range on the Wechsler scales). Raven’s Progressive Matrices, while a pictorial test, is cited in this study as requiring “the ability to infer rules, to manage a hierarchy of goals, and to form high-level abstractions.” This result is strongly in accord with the Thinking in Pictures theory of autism, which provides for the pictorial execution of all of these high-level reasoning processes.

**Neurobiological Evidence**

Here, we discuss two neuroimaging studies using functional MRI that relate to representations and attentional strategies, respectively. The first study looked at differences in brain activation between individuals with autism and typically developing individuals when they had to answer true/false questions about high or low imagery sentences (Kana et al., 2006). High imagery sentences included statements like, “The number eight when rotated 90 degrees looks like a pair of eyeglasses,” while low imagery sentences included statements like, “Addition, subtraction, and multiplication are all math skills.” The control group showed a significant difference between the high and low imagery conditions, with the high imagery condition eliciting more activity from temporal and parietal regions associated with mental imagery as well as from inferior frontal regions associated with verbal rehearsal. In contrast, the autism group showed similar activation in both conditions, with less activity in inferior frontal language regions than the control group in the high imagery condition, and greater activity in occipital and parietal visual regions in the low imagery condition. These results suggest that individuals with autism rely on visuospatial brain regions to process both high and low imagery sentences, unlike typically developing individuals who use these areas more for high imagery sentences and use verbal areas for low imagery sentences.

The second fMRI study looked at the differences in brain activation between individuals with autism and typically developing individuals while they performed the Embedded Figures Task (EFT) which, as described earlier, is a visual search task (Ring et al., 1999). While many brain regions showed similar activation between the two groups, the control group showed greater activation than the autism group in prefrontal cortical regions that are associated with working memory and in particular serial search. In contrast, the autism group showed greater activation in occipito-temporal regions that represent low level visual processing and have been linked to mental imagery (and possibly motion). These results suggest a difference in high-level attentional strategy between individuals with autism and typically developing individuals while performing the EFT, with typically developing individuals recruiting a working-memory-based serial search strategy and individuals with autism using an imagery-based strategy.

**Other Cognitive Theories of Autism**

**Mindblindness**

Mindblindness hypothesizes that individuals with autism lack a “theory of mind,” or ability to ascribe beliefs to other beings (Baron-Cohen, 1995). This limitation could lead to atypical social and communicative behaviors. The classic study used in support of this theory is the false-belief task, in which two characters (typically dolls named Sally and Ann) are shown alongside two baskets. Sally places a marble in her basket and exits the room, after which time Ann switches the marble from one basket to another. When Sally returns, the subject is asked in which basket Sally will search for her marble. Responding that Sally will look in the first basket, where she still supposes the marble to be, requires ascribing a false belief to Sally, i.e. a belief that does not match the state of the real world.

The original study (Baron-Cohen, Leslie, & Frith, 1985) looked at three groups for this task: children with autism (mean age 12 years), children with Down’s syndrome (mean age 11 years), and a control group (mean age 4 years). Both the Down’s syndrome group (who had a lower mean verbal mental age than the autism group) and the control group (who had a lower mean physical age than the autism group) averaged percent-correct scores in the mid-80s, while the autism subjects scored only in the mid-20s.

While Mindblindness holds that theory of mind is a distinct mental mechanism (Baron-Cohen, 1995), one can also approach theory of mind from the standpoint of representations in general. As discussed in the previous section, the formation of concepts like intentionality or mental states is difficult using pictorial representations. If
these concepts were made accessible through pictorial representations, for instance through diagrams or metaphors, then we might expect to see improvements in theory of mind capabilities. A recent study (Wellman et al., 2002) used cartoon-drawn thought-bubbles to teach children with autism about mental states. After this training, most of the children passed standard false-belief tasks that they had previously failed as well as other theory of mind tasks.

Also, a dissociation has been found between children with autism and typically developing children in their abilities to represent different kinds of “false” states. In one study (Leslie & Thaiss, 1992), two groups of children were tested on standard false belief tasks, in which subjects had to reason about beliefs that did not match the state of the real world, as well as on false photograph and false map tasks, in which subjects had to reason about external visual representations that did not match the state of the real world. The control group performed well on the false belief tasks but poorly on the false photograph and false map tasks. As expected, the autism group performed poorly on the false belief tasks, but they actually outperformed controls on the visual tasks. This suggests that, while impaired in their understanding of abstract false beliefs, the children with autism had access to richer pictorial representations or stronger pictorial reasoning skills than the control group.

Weak Central Coherence
Weak Central Coherence hypothesizes that individuals with autism have a limited ability to integrate detail-level information into higher-level meanings, or are at least biased towards local instead of global processing (Happé & Frith, 2006). This trait is presumed to account for some of the stereotyped patterns of behaviors and interests in individuals with autism. Also, superior performance on certain tasks like the EFT is explained with the rationale that in individuals with autism, visual search is unhindered by potentially distracting gestalt perceptions.

However, as described above, these results can also be explained under the Thinking in Pictures hypothesis by enhanced visual attentional strategies that could arise from a bias towards pictorial representations. Other evidence for Weak Central Coherence often includes verbal tests, such as deficits in homograph pronunciation in sentence contexts (as cited in Happé & Frith, 2006). These tests, while putatively measuring local, word-level versus higher-order, sentence-level processing, can also be interpreted as tests of verbal reasoning skills, which would be impaired under the Thinking in Pictures account.

Executive Dysfunction
The final cognitive theory of autism that we discuss is Executive Dysfunction, which hypothesizes that individuals with autism have limitations in their executive functions such as planning, mental flexibility, and inhibition, among others (Russell, 1998). Many studies cited in support of the Executive Dysfunction theory include verbal tests of memory and inhibition and sorting-based tests of mental flexibility. However, in the Thinking in Pictures account, we would expect to see atypical performance in these verbal and category-dependent areas.

Furthermore, as cited earlier, a recent study (Dawson et al., 2007) showed that both children and adults with autism performed considerably better on Raven’s Progressive Matrices than on Wechsler scales of intelligence. Raven’s Progressive Matrices are deemed to test fluid intelligence, which includes “coordinated executive function, attentional control, and working memory” (as described by that study). Therefore, these results do not seem to indicate a general executive dysfunction in individuals with autism.

One possible explanation, using the Thinking in Pictures hypothesis, is that individuals with autism have deficits in executive functions that are verbally mediated but not in executive functions that are (or can be) pictorially mediated. This view is consistent with current models of working memory that propose two distinct storage components—the phonological loop and the visuospatial sketchpad—that operate under a central executive (Baddeley, 2007).

Conclusions
Our aim in this paper has been to show that Thinking in Pictures has significant strengths as a cognitive account of autism that are not to be found in existing theories, both in terms of its explanatory breadth regarding the behaviors of autism as well as the depth to which it can account for many different pieces of supporting empirical evidence.

Of course, the range of autistic behaviors that any of these theories, including Thinking in Pictures, can explain of and by itself remains an open issue, as does the question of whether any of these theories might apply more to one particular subset of the autistic population than another. It is possible that a full cognitive account of autism requires a combination of theories, or that developing a coherent theory requires the identification of specific subgroups of individuals on the autism spectrum beyond what has already been established.

Acknowledgments
Kunda was supported by the Office of Naval Research through the NDSEG fellowship program, and Goel’s work was partially supported by an NSF IIS Grant (#0534622) on Multimodal Case-Based Reasoning in Modeling and Design. The authors would like to thank Gregory Abowd, Rosa Arriaga, Nancy Nersessian, Eric Schumacher, melody Moore, and Irfan Essa for their helpful comments, as well as members of the Design Intelligence Lab and the Autism Research Group. Also thanks to Andrea Ray Chandler for the use of her cat sketch.

References