Advancements in magnetic imaging techniques have revolutionized our ability to study the developing human brain in vivo. The ability to noninvasively image both anatomy and function in healthy volunteers, including young children, has already enhanced our understanding of brain and behavior relations. The application of these techniques to developmental research offers the opportunity to further explore these relationships and allows us to ask questions about where, when and how cognitive abilities develop in relation to changes in underlying brain systems. It is also possible to explore the contributions of maturation versus learning in the development of these abilities through cross-sectional and longitudinal research involving training and intervention procedures. Current imaging methodologies, in conjunction with new and rapidly evolving techniques, hold the promise of even greater insights into developmental issues in the near future. These methodologies and their application to development and learning are discussed in the current paper.

**Key Words:** MRI; fMRI; imaging; development; learning; maturation

## INTRODUCTION

Research into the development of cognitive and behavioral abilities has a long history and has produced a large knowledge base on which to build. The recent combination of cognitive research and neuroimaging techniques allows questions about how these abilities develop in relation to the development of underlying brain systems. For example, while brain imaging studies with adults have explored questions relating brain structure and function [i.e., what areas are active for particular cognitive abilities], developmental imaging has begun to question where and when maturational changes occur in the brain and how these changes relate to development of cognitive abilities [e.g., Casey, 2002]. It is also possible to explore the contributions of maturation versus learning in the development of these abilities through cross-sectional and longitudinal research involving training and intervention procedures [e.g., McCandliss and Noble, 2003, in this issue]. As a relatively young field there are many questions yet to be explored, but developmental neuroimaging provides an opportunity to formulate and test hypotheses that were inaccessible just a decade ago.

Although other imaging techniques have been available for longer periods of time (e.g., positron emission tomography, event-related potentials), the evolution of noninvasive techniques (e.g., functional MRI, magneto-encephalography) with good spatial resolution has provided an enormous boost to neuroimaging research. Functional magnetic resonance imaging (fMRI) capitalizes on magnetic differences between oxygenated and deoxygenated blood to generate blood oxygen level dependent (BOLD) signals. In brief, hemoglobin in the blood becomes paramagnetic in the deoxygenated state and can be compared with oxygenated blood as a naturally occurring contrast agent. During brain activity localized changes in blood flow increase the level of oxygenated hemoglobin (and reduce deoxygenated hemoglobin). This blood flow change can be detected as an increase in the MR signal in that brain location [Kwong et al., 1992; Ogawa et al., 1990]. This signal change provides the base measure of fMRI and allows for spatial resolution on the order of a few millimeters [Hu et al., 1997; Menon et al., 1995], with the temporal resolution as short as a few seconds [Boyton et al., 1996; Dale and Buckner, 1997].

In this paper we will discuss several important issues relevant to developmental and clinical neuroimaging research. These issues include anatomical, physiological, and psychological differences between children and adults, as well as general issues that must be considered when making between-group comparisons. As developmental research often makes comparisons across a wide range of ages and/or performance levels, strategies for equating performance or for dealing with performance differences are usually required. In addition, we emphasize the development of age-appropriate and scanner-appropriate tasks that children can complete to a reasonable level of performance. Several of these issues and strategies will be illustrated in the context of an empirical study of development and learning in healthy children and adults.

## DEVELOPMENTAL DIFFERENCES IN BIOLOGICAL PROCESSES

There are many anatomical and physiological differences between children and adults that are relevant to imaging research. These differences include; height and weight, size, and head, respiration rate, heart rate and blood pressure, as well as differences in synaptic density and the development of myelination [e.g., Casey et al., 2002; Gaillard et al., 2001]. Given these differences one might expect significant variability in the BOLD response between children and adults. However, previous research suggests that the BOLD response is quite similar for

---

Grant sponsor: NIH; Grant number: R01 MH63255.
*Correspondence to: Matthew Davidson, 1300 York Ave, Box 140, Weill Medical College of Cornell University, New York NY 10021. E-mail: mcd2002@med.cornell.edu
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/mrdd.10076
children and adolescents, relative to adults. Although there is individual variability, to date most of the developmental research has shown fairly similar activation patterns in children and adults [Casey et al., 1995; Gaillard et al., 2000; Nelson et al., 2000; Thomas et al., 1999]. In addition, recent research has measured both the time course and peak amplitude of the BOLD response and found similar hemodynamic responses in children and adults with similar behavioral performance [Schlaggar et al., 2002]. These results suggest that basic anatomic and physiological differences are not a limiting factor when comparing school-aged children and adults.

There are also psychological differences that must be considered when making comparisons between age groups. For example, children may experience more anxiety than adults in the testing situation, which may influence their willingness to participate or their ability to understand and comply with the requirements of the study. These differences are generally greater for younger children, relative to adolescents and adults, and can vary as a function of gender and puberty [Spear, 2000]. Children may also experience greater difficulty with the task and/or adopt an alternative strategy for performing the task. Strategies for minimizing these differences and equating both the testing experience and subjects' performance as closely as possible will be discussed. In all cases we should identify and quantify individual and group differences, and use this information in our analyses and interpretations to reduce spurious effects and more appropriately characterize the abilities under investigation.

ACCLIMATION TO THE SCANNER ENVIRONMENT

Most people experience some level of discomfort or anxiety in unfamiliar situations, including university and hospital settings. This anxiety is often greater when being tested or when performance is being measured. An important step in alleviating this anxiety is to create an environment that is relaxed and comfortable for participants. For children, a waiting room (or play room) with toys and activities can provide a pleasant distraction while parents complete the necessary paperwork. This setting should allow the children to relax and spend time with a researcher prior to starting the testing session. As in behavioral studies, neuroimaging researchers need to be alert to signs of anxiety throughout the testing session and proceed to the next level only when the child appears comfortable with the current activity.

In neuroimaging research, these steps are particularly important, as the participants will be exposed to a new setting, with loud and potentially frightening equipment. Many imaging researchers have found simulation devices to be extremely valuable tools for working with children (see Fig. 1). These devices can provide an early indicator of anxiety or claustrophobia that may be experienced in the relatively small bore of the magnet. A simulation session can introduce a child to the sights, sounds and experiences of the scanner setting, as well as give the child (and parents) an opportunity to ask questions about the scanning procedure. The simulator session can also be useful for training participants on the task (if desired or necessary) and can provide an opportunity to give feedback about movement and inappropriate behavior while in the scanner.

Recently, several groups have implemented training techniques to help children recognize when they are moving and improve their ability to lie still. Briefly, a motion detection system within the head-coil is connected to the video presentation system such that excessive motion disrupts the video display for several seconds, producing a break in the movie that the child is watching. Through a gradual reduction in the acceptable range of motion, children learn to relax and stay still to avoid disruptions in the movie. In our experience, this increased awareness and understanding of the limits for acceptable movement translates into fewer motion artifacts during the actual scanning session.

MOTION CORRECTION

Both MRI and fMRI are quite sensitive to artifacts produced by subject motion. Even small movements of the head and neck can significantly reduce the quality of the images. These artifacts are more likely when working with special populations and can be a major concern when working with young children. In addition to the simulation and training procedures discussed above, several of the methods used to minimize motion in adults are also effective with children. These methods may involve head restraints (e.g., pillows, foam padding, head and chin straps or bite bars) and periodic reminders to lie still and minimize head motion.

Even with all of these precautions it is not possible to eliminate movement completely. As a result, it is necessary to correct for motion using off-line motion correction algorithms [e.g., Woods et al., 1992]. These algorithms can realign images to correct for a certain degree of motion and then provide measures of the amount of correction that was required.

In developmental research it is important to consider the amount of correction required for individuals and groups of subjects. In our lab we generally use a threshold of less than half a voxel of mean motion as the criterion for including each subject in the final analyses. We then make comparisons of motion variance between groups to ensure similar distributions. Imaging results can be distorted greatly by differences in motion artifacts and motion correction distributions.

PARADIGM DEVELOPMENT

Some paradigm development issues relate to all forms of developmental research while others are specific to the imaging environment. For example, it is always important to use age-appropriate tasks that allow children to achieve reasonable levels of performance. These tasks must be sufficiently interesting to keep the children engaged long enough to collect adequate data. Young subjects become frustrated with tasks that are too difficult and/or too boring, affecting both the results and the attrition rate of the study. The overall length and structure of the scanning session are also important factors, as children tend to lose focus faster than most adults. Each of these issues will play a role in the experimental design, along with factors such as the number of conditions to be tested and the number of data points required for sufficient statistical power.

The scanning session is usually comprised of both structural and functional data collection. In general researchers collect at least one set of structural images to allow localization of the functional activations. In our lab we collect a high resolution 3D volume (for anatomical studies similar to those described in Kennedy et al., 2003, in this issue) as well as a set of in-plane images that cover the whole brain and that are used for localizing the functional activity. For the functional data collection, imaging tasks typically are organized into a number of discrete ‘runs’ with rest breaks in between. The appropriate number and length of these runs is influenced by the groups involved and by the number of conditions being tested but should be kept as short as possible when working with children. Our developmental tasks are usually divided into several runs lasting 4 to 5 minutes each. When combined
with the set-up time and short breaks between runs this combination of structural and functional scans lasts for approximately one hour.

In addition to considering the number and length of these runs, it is important to use an appropriate experimental design. Initially, fMRI task designs were fairly limited and most studies involved some variant of a block design. In these designs many trials of the same type are presented in short sequences or
blocks within a run. The number of blocks depends on the number of conditions being tested and the length of the run. The analysis involves averaging activations across blocks of similar trial types and comparing to blocks of other trial types. Single trial or event related designs have become quite popular as they can be used to isolate activations for individual trials or subcomponents of trials. These designs typically involve long inter-trial intervals to accommodate the hemodynamic response. This response has been shown to take approximately 6 seconds to peak and 12 or more seconds to return to baseline [e.g., Rosen et al., 1998]. The introduction of rapid mixed trial designs has dramatically improved the temporal resolution of fMRI and provides much greater flexibility in task designs [Dale and Buckner, 1997]. These designs can distinguish events occurring 2 to 4 seconds apart and are more analogous to standard cognitive/behavioral tasks, with randomly intermixed trials and more reasonable inter-trial intervals.

Although these new designs can significantly reduce the amount of time needed to collect adequate imaging data it is still important to make the paradigm age-appropriate. The task must be easy to explain, relatively easy to complete, and must engage the child long enough to collect MRI data. We generally present each task as a game, complete with ‘levels’ for the child to master, and often in a cartoon or superhero framework. For example, Durston et al. [2003] used the Pokemon cartoon characters in a variant of the go/nogo task with both healthy and ADHD children. All of the children were familiar with these characters and found it easy to comprehend the instructions, ‘push the button for each of the characters, except for Meowth.’ These characters are quite colorful and presumably more engaging for the children than simple symbols or shapes used in previous versions of this task [e.g., Casey et al., 1997].

In order to make between group comparisons it is important to design tasks that allow comparable levels of performance for subjects in both groups. This step is relevant to comparisons of clinical groups [e.g., Chapman and Chapman, 1978] as well as developmental groups. Several approaches have been used to overcome the confound of performance differences between age groups. Based on the results of the pilot testing it may be possible to more closely equate performance by modifying the task or by including parametric manipulations. These manipulations should leave the subjects performing essentially the same task but with varying degrees of difficulty, such that children are able to perform at levels comparable to adults. Thomas et al. [1999] used an ‘n-back’ task in an effort to equate performance between groups, with children performing a 1-back version (detection of a repeat of a stimulus 1 trial back) while adults did a 2-back version of the task (detection of a repeat of a stimulus 2 trials back). Importantly, even though all subjects started out with equal performance, the children’s performance decreased as a function of time on task in the scanner while the adults improved. This finding highlights the importance of a thorough behavioral analysis in conjunction with the imaging analysis.

A second approach to making group comparisons is to correlate the behavioral and imaging dependent measures with each other and with the age of the subjects. With this approach it is possible to measure age related changes and performance related changes independently. These analyses can help to dissociate the contribution of developmental and performance differences with brain activation patterns. For example, activation in one region may correlate with the subjects’ ages while activation in a different region may correlate with accuracy or RT performance. It is important to note that performance and age are often correlated so it may be necessary to run additional tests (e.g., partial correlations or multiple regression models) to effectively evaluate these contributions [Casey et al., 1995; Klingberg et al., 2002]. This approach will be discussed in more detail in the sections that follow.

A final consideration in paradigm development is the amount of interaction and/or feedback allowed between the researchers and the participants. In fMRI research it is quite difficult to interact with the participant during data collection. It is possible to interact between runs but this is generally limited to video and audio communication, rather than personal contact. This limitation may be a greater concern with younger children, who may require extra contact and encouragement to maintain performance throughout the task. The appropriateness of feedback is often dependent on the particular task being used, or question being asked, but can be given during the task (between trials or blocks of trials) and/or during the break between runs. Again this feedback is generally limited to video or audio presentation and may be a greater problem for the children.

DATA ANALYSIS

Multiple dependent measures can be used to detect differences in performance and activation patterns. It is important to begin the analysis by looking at the behavioral data generated by the individual subjects. A thorough analysis of reaction time (RT) and accuracy performance will give an indication of how well individual subjects followed instructions and whether they remained on task throughout the scan. Dramatically different performance (among individuals or groups) will likely bias the imaging results. Some degree of variability is inevitable, but subjects who are not performing the task at roughly the same level as the others most likely will have different activation patterns as well. Therefore, important decisions need to be made about inclusion or exclusion of individuals.

In the imaging analyses one can measure differences in activation for experimental conditions and subject groups in a number of ways. For example, one can compare two conditions to determine the location of activations (regions of interest, ROIs) that differ between conditions or groups. In addition to the location, it is possible to detect differences in the volume of activation (the number of voxels activated) and/or the intensity of activation (the change in signal strength for these voxels). In event-related designs it is also possible to assess differences in the time course of activation (e.g., time to peak signal) for particular regions [Cohen et al., 1997]. Within group analyses often use t-test or f-map statistics to compare two conditions on one or more of these measures. These statistics are essentially subtraction techniques that evaluate mean or variance differences in activation between the two conditions in each voxel. Most importantly, the ANOVA allows one to directly test group X condition interactions.

Multifactorial ANOVAs or general linear models (GLMs) can be used and have several advantages over simple t-test comparisons. In particular, multivariate analyses allow each task condition to be compared simultaneously against all other conditions. These analyses are less prone to family-wise error and can be used to test for nonspecific trends and subject variability in the MR signal by including run and/or subject as factors in the model. As discussed above it is also possible to run correlations between age and behavior and the different measures of activation. These correlations generally involve calculating individual percent difference scores between two con-
ditions (e.g., experimental – control/control) for each of the measures (excluding age). This form of normalization removes overall subject differences in activation and performance but provides measures of proportional differences among conditions that can then be correlated.

APPLICATION TO DEVELOPMENT AND LEARNING

In a recent study [Casey et al., 2002], we asked children and adults to perform a stimulus-response task that contained both compatible and incompatible mappings. We were particularly interested in exploring how activation patterns differ for mature and immature brains when learning new spatial stimulus-response associations. Eight adults (mean age 24.5 years) and eight children (mean age 8.8 years) were scanned while performing three different stimulus-response conditions. In all conditions subjects were shown a centrally presented digit on each trial and were required to press one of three buttons that were associated with the digits. In the compatible condition participants pressed the first button in response to the number 1, the second button in response to the number 2 and the third button in response to the number 3 (i.e., 1-2-3 mapping). In the incompatible conditions participants had to learn new associations for the presented numbers. For example, the response mapping of 3-1-2 required a press of the first button when they saw a “3,” the second when they saw a “1” and the third when they saw a “2.” A second incompatible mapping (2-3-1) was also used, with all three conditions presented in an ABCCBA design that was repeated in four identical runs.

The behavioral results from this study showed that children had more difficulty than adults overriding a well-learned stimulus-response mapping in favor of a new one with the incompatible mappings, producing significantly more errors than for the compatible mappings. Both adults and children were slower to respond in the incompatible conditions, however, their RT difference scores (latency to respond to an incompatible mapping versus a compatible one) were not significantly different. The imaging results showed activation in the basal ganglia, inferior frontal/insular cortex, thalamus and hippocampal regions. A two-way ANOVA testing the interaction between group and condition showed children to have larger volumes of activity in the basal ganglia and hippocampal regions for the incompatible mapping conditions, relative to the compatible condition (see Fig. 2).

Percent difference scores were then calculated for each of the measures (incompatible–compatible/compatible) for use in correlation analyses. Correlations between activations and performance showed a negative relationship between RTs and intensity of activation in the hippocampal area, with a greater increase in this region for individuals with the smaller increases in latency for incompatible mappings. Similar comparisons between accuracy performance and intensity of activation revealed positive correlations in the caudate nucleus and thalamus.

Fig. 2. Brain regions showing a robust MR signal change for the interaction of group (children, adults) by condition (incompatible, compatible). (© 2002 by the Journal of Neuroscience. Reprinted with permission.)

Fig. 3. The regions of significant activation for the incompatible versus compatible conditions, for all subjects, are presented in the top half. The graphs in the bottom half show the percent change in MR signal significantly correlated with performance and age for each of the regions presented in the top half. Closed circles are data from children and open circles are data from adults. (© 2002 by the Journal of Neuroscience. Reprinted with permission.)
inferior frontal cortex, with more activation in these regions for subjects making a greater number of errors (see Fig. 3).

Taken together, these results suggest it is possible to dissociate the contributions of frontostriatal and hippocampal related circuitry in the learning of new spatial stimulus-response associations with developmental populations. In particular, frontostriatal circuitry appears to be sensitive to interference from the well-learned association and greater for children while hippocampal circuitry appears to be involved in the learning of new associations between spatial stimulus-response mappings which did not show a developmental effect across this age range. The positive correlation between number of errors and frontostriatal activation suggests that subjects who were experiencing greater interference were activating these regions to a greater extent. The negative correlation between hippocampal activation and RT difference scores suggests that subjects who learned the new associations relatively well and therefore could produce the responses quickly activated the hippocampal regions to a lesser degree. Both of these regions showed larger volumes of activation, extending more ventrally, in the children and suggest that regional activity may be less focal in the immature brain relative to the mature brain.

One interpretation of this less focal pattern of activity in children is that it reflects a delay in maturity of network circuity, with more diffuse patterns of activity found in a relatively immature network. Interestingly, this pattern of refinement was found by Luna et al. [2001] when comparing children and adolescents to adults in a task that required suppression of eye movements. In that study younger children showed a greater divergence in activations than adolescents, relative to adults. They suggest that relatively immature networks may be a key factor influencing performance and activation differences between these age groups. Although it is difficult to assess the maturation of projections between brain regions with fMRI, new imaging techniques are currently being used in an effort to trace fiber tracts within the brain as described by Watts et al. [2003], in this issue. The evolution of diffusion tensor imaging (DTI) offers a method for measuring the extent of connectivity between particular regions and potentially track the development of networks within the brain. This technique is already being used with healthy and clinical populations to assess regional differences in the degree of brain myelination [Klingberg et al., 1999]. With further refinements it should be possible to reliably quantify changes in myelination as a function of development and to better assess the effects of maturation in brain structures and brain networks on performance and activation.

CONCLUSIONS

Although fMRI provides a safe and reliable method for exploring the brain structures and networks that underlie cognitive abilities, a great deal of care is needed when acquiring and analyzing developmental data. It is important to minimize anxiety and motion artifacts for both children and adult participants. Every effort should be made to develop age and scanner appropriate tasks that allow children to perform at levels comparable to adults. This may require modifications to the testing procedure, such as using cartoon characters or computer games strategies, or to the task structure, such as context or parametric manipulations. Particular statistical approaches may still be needed to overcome group differences in performance and/or activation and allow more appropriate comparisons. These approaches can include multivariate ANOVAs and general linear models as well percent difference scores and correlation analyses. As with most between group comparisons it is important to identify and control for any additional factors that may produce spurious effects and confound the interpretations. The learning study used as an example in this paper touched on many of these issues and highlighted the correlation approach to analysis of performance and activation data across groups.

The combination of developmental research and neuroimaging methods has already produced a better understanding of the relationships between cognitive and brain development. However, as these methods and others, such as diffusion tensor imaging and pharmacological MRI, continue to evolve and are brought together there is potential for even greater understanding. The diffusion tensor technique provides an opportunity to examine fiber tracts and myelination in the developing human brain, [Li and Noseworthy, 2002; Ulug, 2002]. With this technique it should be possible to examine the mechanisms underlying network development in typically developing children as well as developmental disorders such as dyslexia [Klingberg et al., 1999]. The pharmacological MRI technique may provide a measure of transmitter influences on the BOLD signal and could be extremely helpful in addressing neurochemical questions in developmental research [e.g., Vaidya et al., 1998]. Even though many of these techniques are still evolving, and have only recently been applied to developmental questions, it is clear that they will continue to have a significant impact on evolving theories of behavioral and neural development in both typically and atypically developing populations [see this issue Durston, 2003; Eigsti and Shapiro, 2003; McCandliss and Noble, 2003; McEwen, 2003].

REFERENCES


Klingberg T, Vaidya CJ, Gabrieli JD, et al. 1999. Myelination and organization of the frontal