Long-Term Memory in Infancy

Electrophysiological and Behavioral Measures of Declarative Memory in 14-Month-Olds

Kristina Magnusson & Emilia Thorup


Supervisor: Magnus Lindgren
Examiner: Sven Ingmar Andersson
Acknowledgments

First of all we would like to thank our supervisor, Magnus Lindgren, for his enthusiasm, inspiration and guidance throughout the work that led up to this paper. Secondly, we want to express our gratitude to Devin Terhune for his inexhaustible support in all technical matters. Thanks also to Karin Strid and Mikael Heimann, for sharing with us their experience on deferred imitation. Last but not least, we wish to extend our thanks to all children and parents who have been willing to participate in this study.
Abstract

This paper presents a pilot study investigating the development of declarative memory in infancy, by creating a stimulus material appropriate for recording Event-Related Potentials (ERPs) during memory encoding. The material was designed to establish associations between pictures, and a negative component with a fronto-central distribution (Nc) was studied as a reflection of learning. In addition an established test of declarative memory, deferred imitation (DI), was performed in order to serve as a behavioral correlate. Information concerning general development was also collected, including an assessment of productive vocabulary as measured by the Swedish version of the MacArthur Communicative Developmental Inventories. The participants were 23 typically developing 14-month-old infants. DI data from all participants was analyzed and correlated with other developmental measures. A strong positive correlation was found between productive vocabulary and DI performance. Five participants generated data for ERP analysis, which revealed the expected Nc component but failed to show the predicted patterns. This may be explained by the current presentation of the stimulus material not making the associations clear to the infants. Suggestions for improvements of the ERP design are discussed.

Key words: ERP, deferred imitation, declarative memory, infant, Nc, MacArthur CDI, ASQ
Introduction

The ability to form long-term memories, allowing us to recall our past and build up our knowledge of the world, is fundamental to the human experience and something that most adults probably do not reflect upon. Memories that are consciously accessible are termed declarative and have traditionally been studied through verbal report. Investigating the early development of this memory system has therefore been challenging considering that infants are not able to express themselves verbally. To overcome these challenges, non-verbal methods of investigating declarative memory have been developed, opening a whole new field of research. The method of deferred imitation (Meltzoff, 1985) allows infants to convey their recollections through action, namely the imitation of a precise target action earlier demonstrated by the researcher. This method is therefore a valuable tool in investigating the development of long-term declarative memory.

For a complete understanding of memory development, however, behavioral measures such as deferred imitation need to be combined with measures of actual brain activity during encoding and recall. Although many such methods, for example functional magnetic resonance imaging (fMRI), are inappropriate for young children, electroencephalogram (EEG) measurements are non-invasive and tolerable even for infants. Event-related potentials (ERPs), derived from the EEG, can then be studied in order to draw conclusions about the brain’s response to specific stimuli. This paper presents a pilot study investigating declarative memory in 14-month-old infants by combining ERP measurements with the method of deferred imitation. By developing an ERP paradigm for studying declarative memory in infants, the ambition is to create a substitute for behavioral methods that is able to obtain information of a different nature and answer a wider spectrum of questions.

Before giving a detailed description of the present study, we present a theoretical background that provides the foundation for the research project. The first part of this paper attempts to introduce the concept of declarative memory, before providing an account of the research on its development. Moreover, relevant findings on memory-related ERPs will be introduced, and finally this section will present the fundamentals of the deferred imitation paradigm.
Memory is usually understood as being composed of several systems, with qualitatively different functions and underlying brain anatomy. Though not all researchers agree to the specifics of which these systems are, a commonly used distinction is the one between non-declarative and declarative memory. The construct of non-declarative memory contains functions such as procedural learning, conditioning and perceptual representation. It generally refers to a type of memory that the individual is not consciously aware of. Parts of the striatum and cerebellum as well as cortical areas are responsible for this type of memory. These areas are early developing, and may be functional at birth. Therefore, procedural memory and learning is evident in very young infants. Declarative memory, on the other hand, is a later developing system. The brain areas responsible for declarative memory are structures in the medial temporal lobe, including the hippocampus and parahippocampal cortex (Richmond & Nelson, 2007).

Declarative memory is understood as memory that can be declared, such as memory for facts and experienced events. Declarative memory is further divided into semantic and episodic memory. The term semantic memory refers to memories for facts and general knowledge, whereas episodic memory refers to memories for personally experienced events (Richmond & Nelson, 2007). Episodic memory is commonly viewed as the capacity for time travel, that is the ability to think about one’s past. It has been suggested that episodic memory is unique to humans, and that it is more vulnerable to neuronal dysfunction than other memory systems. In terms of evolution, episodic memory is thought to have developed much later than other memory systems, and to have evolved out of semantic memory. Episodic memory shares many features with semantic memory, and it is dependent on, but also extends it. Ontogenetically, episodic memory is thought to be later developing than semantic memory (Tulving, 2002). Tulving (1993) suggests that episodic and semantic memory also differ in the kind of awareness that accompanies retrieval of memories. He refers to the type of awareness associated with semantic remembering as noetic awareness. In contrast, autonoetic awareness is associated with episodic memory retrieval, and refers to the special kind of awareness that allows humans to be conscious of the subjective time in which an event took place. According to Tulving (1993), autonoetic awareness is what makes mental time travelling possible. To further explain the distinction between episodic and semantic memory
systems, Tulving (1993) differentiates remembering from knowing. To know certain facts about the world reflects semantic memory, and the process of knowing is noetic. To remember, on the other hand, is to actively recollect facts such as when and where a personal experience occurred. Remembering reflects episodic memory, and the associated state of awareness is autonoetic.

Tulving’s (1993) understanding of episodic memory is strict in the sense that he describes the process of it as one where the individual mentally and consciously travels back in time. Reporting of it is dependent on introspection. With this strict understanding comes the notion that episodic memory cannot exist in animals or young children. Though accepting of Tulving’s view at large, other researchers have taken a more generous understanding of the concept of episodic memory, where the process of episodic retrieval is viewed as less dependent on the ability to put it into words (e.g. Aggleton & Brown, 2006). Such an understanding of episodic memory allows for the possibility of it being present also in animals and young children. Researchers advocating this view have run animal studies that have shed new light on the understanding of episodic memory and its development. Methods have also been developed that are thought to measure episodic memory in young children who have not yet started to use language (e.g. Meltzoff, 1985).

Much research has taken place since the concept of episodic memory was first proposed, and as new discoveries are made, new controversies and debates have also emerged on the topic. One is the debate about the nature of recognition memory and its relationship to episodic memory. The term recognition memory refers to a memory function that signals whether an event has been previously experienced. Some authors (e.g. Dunn, 2004) consider recognition memory and episodic memory as part of the same continuum, being supported by the same brain structures. Others (e.g. Aggleton and Brown, 2006) argue that recognition memory is composed of at least two independent processes, which serve qualitatively different functions and are supported by different brain areas. Episodic memory in the sense that Tulving (1993) originally described it, cannot be tested in animals since they lack the capacity of introspection and verbal report. By testing recognition memory, researchers have been able to establish that animals such as birds, monkeys and mice have the capacity to under certain circumstances learn contextual features such as where and when an event occurred (e.g. Dere,
Huston & De Souza Silva, 2005). In studies like these, the focus is not on episodic memory as mental time travelling, but rather on the capacity to recognize items on the basis of contextual information. Such a focus is valuable since it allows for studies on animals’ capacity to bind item and context together. Studies like these indicate that a capacity that might be referred to as a primitive form of episodic memory seems to be present in animals, and provide insights on the neural basis of episodic memory in general (Aggleton & Brown, 2006).

In order to interpret findings such as those above described in terms of what they can tell us about human episodic memory development, it is important to know more about the nature of recognition memory and its relationship to episodic memory. Advocates of various two-process models (e.g. Aggleton & Brown, 2006; Yonelinas, 2002) propose that recognition memory consists of at least two separate processes, one of which is recollection and the other familiarity detection. As Tulving (1993) used the terms knowing and remembering to distinguish between semantic and episodic processes, Aggleton and Brown (2006) use the same terms to distinguish between familiarity and recollection processes. Thus, recognition may reflect only a feeling of familiarity (knowing), or the feeling of familiarity may be verified by recollection of contextual details (remembering). Aggleton and Brown (2006) regard recollection as reflective of episodic memory. It is not clear, however, if they view familiarity detection as reflecting semantic memory, or simply regard it a different kind of recognition memory.

Several studies indicate that the two functions of recognition memory are supported by separate regions of the brain, including the medial temporal lobe. Studies on monkeys performing visual recognition tasks have shown that neurons in the perirhinal cortex (located in the anterior parahippocampal gyrus) respond strongly to new items, but only weakly to items that have been previously seen (Brown & Aggleton, 2001; Desimone, 1996). This has been interpreted as reflecting a familiarity detection mechanism (Brown & Aggleton, 2001). The finding is supported by fMRI studies that have found decreased neuronal activity in the perirhinal cortex as a response to previously seen items in humans as well. By contrast, fMRI studies have linked activity in the hippocampus and the parahippocampal cortex to recollection processes (e.g. Ranganath et al., 2004). The notion of the hippocampal area being important for recollection is further supported by studies on amnesia patients indicating that
those with damage to the hippocampal system show impaired recollection, but partially spared familiarity detection (Aggleton & Brown, 2006). The hypothesis of separate processes is further supported by ERP studies that have identified two separate neuronal populations that are functionally and temporally dissociable and are thought to be correlates of the recollection and familiarity processes respectively (Rugg & Curran, 2007). The familiarity detecting perirhinal process has been found to be faster than the hippocampus dependent process of recollection (McElree, Dolan, & Jacoby, 1999). A mid-frontal old/new effect has been identified and proposed as an ERP correlate of familiarity detection. This is a frontally maximal deflection peaking at 300-500 ms post-stimulus, present when the individual correctly recognizes an item as familiar, even when failing to recognize its source. (Rugg & Curran, 2007). A second effect, termed the parietal old/new effect, is thought of as the ERP signature of recollection (Rugg & Curran, 2007). It is a positive-going, parietally maximal ERP effect peaking at 500-800 ms post-stimulus. The effect is present only when identification of familiar test items are accompanied by successful source judgment, and it is topographically and functionally dissociated from other posterior ERP deflections in the same time frame (Curran, 2004). Brown and Aggleton (2001) argue that such findings indicate that the separate processes of recognition memory can be doubly dissociated.

Eichenbaum, Yonelinas and Ranganath (2007) propose a three-component model that further specifies the separate functions of the brain areas involved in recognition memory. Their model, referred to as the binding of item and context (BIC) model, is further developed by Diana, Yonelinas and Ranganath (2007). The model is consistent with the idea that the perirhinal cortex is responsible for making familiarity judgments. It further proposes that the parahippocampal cortex supports recollection by encoding and retrieving contextual information, and that the hippocampus associates item and context information. These findings indicate that the pattern of activation in the medial temporal lobe during retrieval will depend on what type of cue is present. If a context is presented, successful recall should follow a pattern of activation starting in the parahippocampal cortex, followed by the association of item and context in the hippocampus and completion of the pattern in the perirhinal cortex. Presentation of an item should conversely start an involvement of the perirhinal cortex that would be followed by activation of the hippocampus and lastly the parahippocampal cortex (Diana et al., 2007). By contrast, if recollection does not occur but an
item is solely judged as being familiar, such a familiarity judgment should be supported mainly by the perirhinal cortex. The specificity of the BIC model allows for more complex predictions to be made. It should, for example, be possible to predict what sort of activation tests of associative recognition with pairs of items would elicit. Diana et al. (2007) propose that recollection of inter-item associations should elicit hippocampal activity, since the hippocampus would be responsible for linking one item to the other. If the study context would also be retrieved, it seems to follow from the model that activation of the parahippocampal cortex should occur as well. Additionally, if one of the items elicits recollection of the other, activation of the perirhinal cortex should be observed. However, Diana et al. also suggest the possibility of encoding item-pairs as single representations, which would implicate that the perirhinal cortex would be able to support associative recognition on the basis of familiarity. Studies on amnesia patients with hippocampal damage confirm this idea by showing that these individuals perform better at encoding word pairs as single units than as separate words (Quamme, Yonelinas & Norman, 2007).

Development of Declarative Memory

When studying the development of declarative memory, researchers face an obvious problem since preverbal children are not able to put their memories into words. Therefore, studies on declarative memory development have come to rely on other methods. ERP studies are an important source of information on the anatomical development of different brain structures, and behavioral tests that are considered to assess declarative memory in infants have been developed (e.g. Meltzoff, 1985). fMRI studies on healthy and amnesic adults as well as animal studies can also clarify how declarative memory in young children may function. It is important to note however, that comparing human infants to other species is not unproblematic, and neither is comparing adult brain function to that of infants.

Although the brain structures responsible for procedural memory generally mature earlier than those supporting declarative memory, some of the medial temporal lobe structures responsible for declarative memory mature relatively early as well. The hippocampus is an early developing structure with most parts formed early in gestation. Parts such as the entorhinal, perirhinal and parahippocampal cortices therefore seem to be adult-like and functional at birth (Alvarado & Bachevalier, 2000). However, the dentate gyrus of the hippocampus has a
protracted course of development, and does not reach mature status until around one year of age (Seress, 2001). The delayed development of the dentate gyrus has been suggested to have important implications for the function of declarative memory in infants. It has been hypothesized that younger infants rely mainly on the early developing structures of the hippocampus and parahippocampal area when forming declarative memories (Richmond & DeBoer, 2006). While doing so, aspects of the cue, actions and context seem to be fused together into one unitary memory representation. Young infants’ declarative memory performance therefore lacks flexibility, and for a young infant to be able to recall for example a sequence, all aspects such as the props being used and the context the sequence occurs in must remain the same as during encoding (Jones & Herbert, 2006). Representational flexibility, the ability to retrieve memories with cues and in contexts not identical to those originally encoded, is seen as an important feature of declarative memory (Eichenbaum, 1997), and the dentate gyrus is suggested to play a crucial part in the development of this ability. As the dentate gyrus matures, the infant grows increasingly efficient at relationally organizing the details of a sequence and weighing them hierarchically so that successful retrieval becomes less context dependent (Richmond & DeBoer, 2006).

Richmond and DeBoer argue that the development of the dentate gyrus represents a qualitative shift in the nature of infants’ memory representations. Nelson (1995) even proposes that the earliest form of hippocampus dependent memory does not share enough features with mature declarative memory to be termed declarative. Instead he proposes the term pre-explicit memory to describe this function. Pre-explicit memory, he argues, does not only rely on the early developing parts of the hippocampus, but also on other limbic structures such as the amygdala. Somewhere between 6 and 12 months, Nelson proposes that the pre-explicit form of declarative memory is replaced by a more adult-like form of declarative memory. Other researchers (e.g. Hartson et al., 1998) argue that there is no evidence of a sudden improvement in declarative memory performance during infancy, and therefore suggest that the development of declarative memory is gradual rather than involving a shift. Gradual age-related improvement on the deferred imitation task (Meltzoff, 1985) seems to reflect anatomical development of the medial temporal lobe, which has been interpreted as supporting the hypothesis that the development of declarative memory is gradual rather than stepwise (Jones & Herbert, 2006).
Regardless of whether they advocate a continuous development of declarative memory, or one involving a qualitative shift, most researchers agree that while rudimentary declarative memory is present early in development, continued development of encoding, retention and retrieval processes contribute to changes in declarative memory performance throughout infant development (Richmond & Nelson, 2007). It has been suggested that the rapid myelination during the first year of age accounts for the great changes in processing speed that have been found to affect encoding time, allowing older infants to encode information faster than younger (Richmond & Nelson, 2007). Jones and Herbert (2006) propose that young infants’ lack of representational flexibility also may be a contributing factor, since equally attending to all details of an event may require longer encoding time. Whereas the development of encoding speed mainly occurs during the first year of infancy, age-related changes in retention have been found to continue into early childhood (Hayne, 2004). It has been suggested that the dentate gyrus plays a central role in storage and consolidation, and that the protracted development of this structure thus may account for these changes (Bauer, 2005). When it comes to retrieval, the development of the dentate gyrus may again play an important role, since it has been suggested to account for changes in the ability of flexible retrieval (Richmond & DeBoer, 2006). When discussing the contribution of various brain structures, it is important to keep in mind that these do not operate in isolation. The brain is an integrated system, and though the hippocampus and related structures are crucial for declarative memory, it is important to note that other brain areas, for example the prefrontal cortex, are involved as well.

Deferred Imitation

As mentioned earlier, behavioral tests have been developed that are considered to tap declarative memory in infancy. One such method is deferred imitation, which assesses infants’ ability to encode, retain and reproduce a sequence of actions. The deferred imitation task developed by Meltzoff (1985) includes a number of sequences performed with specially designed toys. The experimenter demonstrates a novel sequence of actions performed with a toy that the infant is not allowed to touch. The demonstration is followed by a delay, after which the infant is handed the toy. The idea is that the infant’s ability to reproduce the action earlier demonstrated reflects declarative memory. When the task was developed, control conditions were included to establish that infants normally do not spontaneously perform the
target action if handed the toy without prior demonstration. Therefore, learning can be inferred when the target action is being produced (Meltzoff, 1985).

Imitation has long been known as an important means of learning in infancy. When infants do not have the opportunity to immediately reproduce the action observed, they may form a memory representation of the sequence to later guide their behavior (Jones & Herbert, 2006). Since deferred imitation is based on observation without motor involvement, it is thought to function as a nonverbal assessment of declarative memory (Meltzoff, 1988). This notion is further supported by the finding that adult amnesic patients have difficulties when tested with an adult-appropriate version of the deferred imitation task (McDonough, Mandler, McKee & Squire, 1995). Furthermore, performance on the deferred imitation task is affected by the same variables that influence adults’ performance on other declarative memory tasks. For example, retention is influenced by demonstration time, length of delay and change in context (Hayne, Boniface & Barr, 2000). These findings all indicate that performance on the deferred imitation task relies at least in part on the medial temporal lobe.

Though several researchers agree that deferred imitation seems to assess declarative memory, the question still arises of what kind of declarative memory it reflects. Meltzoff (1985) argues that in order to perform the target behavior, simply recognizing the toy would not be enough. In order to reproduce the action, the infant needs to recall the experimenter performing it. This line of reasoning suggests that episodic memory processes are active. The findings earlier discussed that suggest hippocampal development as crucial for deferred imitation performance further supports the idea, since the hippocampus and related structures have been found to support episodic memory. It is important to note however, that it yet remains unclear how much the form of episodic memory that may be present in infants has in common with the adult form. It seems somewhat unlikely that infants would possess the ability to travel mentally in time, which Tulving considered to be an essential aspect of episodic memory. If one on the other hand chooses to focus on the more primitive form of episodic-like memory that may be present in animals and young infants (e.g. Aggleton & Brown, 2006), deferred imitation seems to be a useful and promising tool for assessing it.
The Event-Related Potential Technique as a Method of Investigating Infant Memory

Studying cognitive processes in young children presents special challenges with regards to many of the methods used with adults, such as fMRI. Infants who cannot follow instructions or remain still are very difficult to study with fMRI, and the procedure can also be quite unpleasant and frightening to young children. In contrast, electrophysiological methods that measure EEG and ERPs can be used successfully even with very young children (DeBoer, Scott & Nelson, 2007). These methods are non-invasive and do not require the child to respond overtly. By placing electrodes across the scalp, the sum of postsynaptic potentials in the brain can be recorded. EEG measures the brain’s continuous electrical activity while ERPs, derived from the overall EEG, reflect the brain’s response to a specific event or stimulus. ERPs are therefore very useful in studying specific cognitive processes such as memory and language.

The ERP components found in children often differ significantly from those found in adults. This can be understood in terms of the major changes that occur in the brain throughout childhood, such as myelination resulting in increased processing speed and thus reduced latency of the ERPs. Some ERP components found in children not only change latency or amplitude throughout development, but seem to disappear completely, presumably developing into other components. One such component is the negative central component (Nc) which is one of the most extensively studied event-related potentials in the visual modality in infants (de Haan, 2007). The Nc is a fronto-central negative wave with a latency ranging from approximately 400 to 800 ms after stimulus onset (Csibra, Kushnerenko & Grossmann, 2008). While the latency of the Nc has been found to decrease during the first year of life, the amplitude increases, thus becoming more negative (Webb, Long & Nelson, 2005).

Courchesne (1978) investigated the presence of Nc in different age groups and found that Nc was still commonly present in 10- to 14-year-olds but then gradually disappeared during adolescence. The functional significance of the Nc is not entirely clear, but one of the earliest interpretations, namely that it reflects an aspect of attention that regards the perception of salient and interesting events (Courchesne, 1978), still holds a strong position.

The Nc has commonly been studied within the oddball paradigm, in which one stimulus is presented frequently and another (the oddball) is presented infrequently. In infants 3 months
of age and older, the Nc has a greater amplitude in response to the oddball than to the frequently presented stimulus, presumably because the oddball is novel and unexpected, and thus more salient in its context (e.g. Courchesne, Ganz & Norcia, 1981). Still, the interpretation of the Nc is not straightforward, as the amplitude is not always greater in response to a more novel stimulus. For instance, when presented with pictures of the mother’s face and a stranger’s face (with equal probability) children up to 2 years of age showed a larger Nc amplitude to the mother’s face, whereas the opposite pattern was found in children older than 45 months (Carver et al., 2003). This indicates that the significance of the mother’s face varies with age.

As the Nc is influenced by novelty and item familiarity, it is of great interest in understanding memory processes in young children. Courchesne et al. (1981) present the hypothesis that as well as reflecting orientation toward interesting events, the Nc may also reflect the strength of memory traces, as the amplitude in his study was greater for items with which the subject had less experience. This particular hypothesis is in conflict with the results that young infants show a larger Nc amplitude to the mother’s face, and is still controversial. Nevertheless, de Haan (2007) and Csibra et al. (2008) both conclude their discussions about the Nc by stating that there must be some element of recognition memory related to the Nc, considering the results showing the component’s sensitivity to familiarity and novelty even when the stimuli are presented with equal probability.

It is clear that interest and motivation plays an important role in the function of the Nc. Intuitively, things become more interesting as the child starts to understand and master the material presented, while interest declines when the material becomes too repetitive. In a recent study by Torkildsen et al. (2008a, in press), a group of 20-month-olds were given a fast mapping task. They were taught associations between nonsense words and their referents, which were pictures of imaginary creatures or objects. The authors observed the modulation of the Nc component in two groups of children, those determined to have high versus low productive vocabularies. In the high productive vocabulary group, a nonlinear pattern was found where the amplitude increased during the first few presentations of the associations, but then started to decrease. This indicates increased interest as the child first detects something familiar, but decreased interest as the material becomes too repetitive. In contrast, the low
productive vocabulary group showed a continuously more negative amplitude with increased repetitions, which indicates that their learning process is slower (Torkildsen et al., in press). The learning phase was later followed by a phase where the associations were violated, by presenting a picture with a different but equally familiar word, or with a completely novel word. At this stage, the authors studied the N400, a component shown to be sensitive to semantic violations in children as well as adults (Friedrich & Friederici, 2004). Children with high productive vocabularies showed an increase in the N400 amplitude to these violations, which led to the conclusion that they recognized the violations and thus had learned the earlier associations (Torkildsen et al., 2008). This study demonstrates how motivation and interest modulates the amplitude of the Nc, but also illustrates the component’s significance when studying memory processes.

Certain studies have managed to combine behavioral measures of declarative memory with ERP measures focusing on differences in Nc wave pattern. Carver, Bauer and Nelson (2000) used deferred imitation with 9-month-old infants to measure recall memory after a 1 month delay. In addition, they examined ERP responses to pictures of old and new sequences one week after demonstration, as a measure of the infants’ recognition memory. Interestingly, it was found that infants who showed evidence of recall memory by reproducing the correct sequences also differentiated between pictures of old and new sequences at an electrophysiological level. Nc amplitudes in this group of infants were larger to new sequences than to old sequences, while infants who did not manage to show recall lacked this differential activation.

In an effort to discover the sources of these individual differences, Bauer et al. (2003) developed a similar design with the ambition to investigate a possible association between ERP data during encoding and delayed recognition, as well as later recall. The goal was to determine whether the infants who failed in the recognition and recall tests did so because of a storage failure or because of an encoding failure. The method used, however, did not in fact record ERP data during encoding, that is during the demonstration of the sequences, but rather during an immediate recognition test following the final demonstration session. The results indicated that the two groups of infants (those who demonstrated recall and those who did not) did not differ in immediate ERP data, but did differ at the 1 week delayed recognition
The authors draw the conclusion that the infants did not differ in how well they encoded the events but rather in the effectiveness of memory storage. The results of this study are certainly interesting in themselves, but considering that no data was recorded at the time of actual encoding, this particular conclusion may be premature.

**Aim of the Present Study**

Up to this point in time no studies of declarative memory in infants have presented ERP data from the actual time of encoding. Using only the method of deferred imitation this type of data is understandably difficult to obtain, since the ERP technique requires a large number of trials while the deferred imitation paradigm allows only a few demonstrations. The present study sought to overcome this problem by using both the method of deferred imitation and a stimulus material more in line with traditional ERP paradigms. One of the types of stimuli most appropriate for ERP recording is slides presented on a computer screen, allowing for strict control of presentation time and number of trials. As has already been mentioned, earlier studies of the Ne component in infants have used this type of stimulus to investigate familiarity detection, for instance using the oddball paradigm (e.g. Courchesne et al., 1981). The material used in these studies has primarily consisted of single items presented at different probability, with or without a familiarization phase. The aim of the present study was to create a stimulus material appropriate for ERP recording that would tap the same underlying memory processes as the deferred imitation task. Theoretically, an associative memory task would to the same extent as deferred imitation require recollection memory as opposed to simple familiarity. As discussed earlier, the model used by Diana et al. (2007) predicts that tests of inter-item associations should engage the hippocampus, as the hippocampus would be needed in linking one item to the other. A slideshow of pairs of objects would therefore prompt a similar type of learning as deferred imitation, which also is assumed to be dependent on the hippocampus.

The study presented in this paper had the ambition of adding to the results found by Bauer et al. (2003) by developing a new method of investigating declarative memory in infants, more suitable for ERP recording, while allowing for interesting parallels between this new method and the more established paradigm of deferred imitation. The idea of the new part of the experimental design was to build associations between two pictures, each pair consisting of an
animal and an object, and observe the modulation of the Nc component with increased repetitions of the picture-pairs. At the end of each test phase the associations were violated by presenting the same animals together with a different but equally familiar object, as well as a completely novel object, in order to test the infant’s learning of the associations. This design allowed for the distinction between the learning of the actual inter-item associations and the simple detection of familiarity versus novelty. With this method it was possible to collect unique ERP data during actual encoding, and in order to tie the results together with earlier studies, the deferred imitation paradigm was used in a separate part of the design.

Since the study involved developing an entirely new type of stimulus material, one of the aims was to evaluate the effectiveness of this material. In addition, four hypotheses were formulated concerning the expected ERP results. In accordance with the results obtained by Torkildsen et al. (in press), Nc amplitude was expected to gradually increase with the first few presentations of the picture-pairs, and then decrease during the last repetitions. The Nc in response to the violations of the associations should once again be larger (compared to the final presentation of the associated pairs, presentation 8), similar to the N400 effect in response to the semantic violation in the Torkildsen et al. (2008) study. An increased Nc following the presentation of familiar pictures in new combinations would indicate that the infant had learned the original associations. An unchanged Nc amplitude at this point, however, would indicate that the infant never learned the associations between the pictures and thus did not react when they were presented in new combinations. An increased Nc amplitude to the completely novel picture would be expected in both cases. The use of deferred imitation (DI) in the procedure provides the potential for correlations between DI test results and the electrophysiological data. Torkildsen et al. (2008) were able to divide their participants into two groups, one high productive vocabulary group and one low productive vocabulary group, and found different patterns of Nc modulation in the two groups. Similarly, this study would be able to divide the participants into high and low achievers on deferred imitation. These groups may then show different patterns of Nc modulation, allowing the DI results to predict the pattern of electrophysiological activation.
Hypotheses

Hypothesis 1. An Nc component (a negative deflection at a latency of approximately 350-600 ms) will be present in response to the second picture in the presented pairs.

Hypothesis 2. The average amplitude in the Nc time window will increase across the first presentations and then decrease during the last presentations.

Hypothesis 3. Nc amplitude will be greater in response to the two violation conditions (mixed & novel) compared to presentation 8.

Hypothesis 4. There will be a difference between high and low achievers on the deferred imitation task, concerning Nc modulation as well as Nc responses to the violation conditions compared to presentation 8.
Method

Participants

The participants were 23 infants (10 boys/13 girls) within the age-range of 14 months +/- 42 days, with a mean age of exactly 14.0 months at the time of testing. After running experiments on 5 participants the ERP stimulus material was altered, thus preventing data from these infants from being included in analysis. Due to a technical error during EEG recording, data was lost for the following 13 participants, resulting in meaningful ERP data from only 5 participants. The deferred imitation procedure was carried out with all participants, and all parents were given questionnaires to fill out concerning their child’s general development. However, some parents did not return the questionnaires, leaving data on general development from 18 children available for analysis. The infants were recruited through e-mails sent out to employees at Lund University, advertisements posted around the university campus and information campaigns at day care centers, as well as advertisements at a local toy store, and an online magazine for parents. Parents were given written information about the procedure (see appendix A), and signed a consent form (see appendix B) before the children participated.

Materials

Deferred imitation. Five toys developed by Meltzoff (1985, 1988) were used. Three of these were developed for 14-month-olds, including a small pull toy, a string of beads that was to be lowered into a cup, and a collapsible cup. In addition, two toys originally developed for 9-month-olds were used, in order to potentially increase the variance of responses. These toys were a wooden construction with a hinge, and a plastic egg that produced a rattling sound when being shaken.

Developmental measures. In order to receive information about the infants’ general level of development, the parents were asked to fill in a few questionnaires. The Swedish Early Communicative Development Inventory (SECDI) (Berglund & Eriksson, 2000; Eriksson & Berglund, 1999) is a Swedish translation of MacArthur-Bates Communicative Development Inventories (CDI Advisory Board, 2006). The SECDI questionnaire used was Words & Gestures, designed for children younger than 16 months. This questionnaire
assesses language comprehension, language production and communicative gestures. Additionally, the parents were asked to fill out an age-appropriate version (14 months) of a Swedish version of Ages and Stages Questionnaires (ASQ) (Paul H. Brookes Publishing Co., 2006), an instrument originally developed for screening for developmental delays, which assesses the infant’s level of development in various areas including language and motor development.

**ERP experiment.** A total of 144 pictures were chosen from the website Clipart.com (Jupiterimages Corporation, 2008). Half of the pictures were of animals, and the other half showed inanimate objects. Most of the pictures were drawn and some were photographs. Vivid pictures in bright colors were chosen to attract infants’ attention, and objects that were thought to be generally familiar to infants (e.g. a pacifier) were chosen. The pictures were presented in a slideshow on a computer screen, grouped in pairs, with each pair containing one picture of an animal and one of an object.

EEG data was recorded with HydroCel Geodesics Sensor Nets, which are appropriate for use with infants due to their properties of allowing high impedance and not requiring abrasion. One infant net, size 44-47 cm, was used as well as a pediatric net, size 47-51 cm, and in one case an adult net, size 51-54 cm.

**Procedure**

**Deferred imitation.** For the DI procedure, the infant sat on the parent’s lap, facing the experimenter, with a small table in between. The five toys were demonstrated in a randomized sequence. For each toy the action was demonstrated three times for a total of approximately 20 seconds. The infant was not allowed to touch the toys during presentation. For the wooden construction, the demonstrated action was to push the vertical extension down so it would lie flat on top of the larger piece of wood. For the plastic egg the demonstrated action was to shake it to produce a rattling sound. For the pull toy the separate parts of the toy were pulled apart. The string of beads was lowered into the plastic cup, and for the collapsible cup the demonstrated action was to have it fully collapsed by pushing it with the palm. Following the demonstration of all five toys was a break during which the EEG recording took place. Depending on the duration of this procedure the break lasted for between 30 and 45 minutes.
Thereafter the infant was again seated at the table. The instructor now handed the infant the toys one by one in the same order as they were earlier demonstrated, and the infant was invited to handle the toys by the Swedish phrase “varsågod” (“there you go”). A second experimenter was present in the room, in order to score the results. All sessions were video recorded, in case any ambiguities would arise during the immediate scoring.

Developmental measures. At the end of the visit the parent was handed an envelope containing the SECDI as well as the ASQ. The parent was asked to complete the questionnaires as soon as possible and return them by mail to the experimenters.

ERP experiment. Two experimenters were always present during preparations for the ERP experiment, allowing one person to concentrate solely on keeping the infant comfortable while the other person measured the child’s head, applied the Geodesic Sensor Net and measured impedance. During application, the infant was seated on the parent’s lap and the assisting experimenter used soap bubbles and toys to distract the infant. The infant remained on the parent’s lap during the experiment, and a white cardboard screen was placed around them in order to screen out distractions. The light was dimmed so that the stimuli would become more salient.

ERPs were recorded from 128 or 124 scalp locations, depending on the child’s head circumference. For infants with heads larger than 47 cm, the 128-channel pediatric or adult sensor nets were used which included orbital sensors, something the smaller infant net lacked. Impedance was kept under 50 kΩ. The EEG was recorded with a Net Amps 300-amplifier, without an online filter and with a sampling rate of 250 samples/second. A vertex reference was used during recording, but an average reference was applied to the data during analysis.

The stimulus material consisted of a slideshow of pictures presented on a 17 inch computer screen (34x27 cm) positioned approximately 35 cm from the infant. The slideshow was divided into blocks, allowing for a short break in between blocks if necessary. The pictures were presented in pairs, with three picture-pairs (consisting of one animal and one object) per block, where each pair was presented eight times, in a pseudo randomized order (the same pair was never presented twice in a row). Between the two pictures forming a pair appeared a
red screen accompanied by a whistling sound, with the volume adjusted to a comfortable listening level, and between each pair appeared a white screen. At the end of each block each animal was paired with one of the other, equally familiar, objects, as well as one novel object. The pictures and red screen were presented for 1000 ms, while the white screen was shown for 1500 ms.

Figure 1. Illustration of the different conditions.
The stimulus material contained a total of 16 blocks, although none of the infants completed all of them. The average number of blocks completed was 4.6. The number of blocks completed depended on each individual infant’s attention span, and the experiment was terminated when the infant no longer attended to the stimulus or became fussy. The entire procedure was video recorded to allow for the elimination of trials where the infant was not attending to the stimulus. The infants on average saw 86% of the material presented.

**Data Analysis and Scoring**

**Deferred imitation.** During the imitation session, the observing experimenter timed the infant and took notes of whether he or she succeeded at producing the target action within 20 or 30 seconds, or not at all. The exact time was recorded from when the infant first touched the toy until the target action was produced. The original time limit used by Meltzoff (1985, 1988) was 20 seconds. The reason for using an additional, more generous, time limit was that a previous study carried out on Swedish children suggested that these may take slightly longer time to produce the target action than American children. This was interpreted as reflecting cultural differences in interaction between children and adults, where Swedish children are socialized into being relatively more cautious (Heimann & Meltzoff, 1996). In the experimental setting, American experimenters also tend to carry the interaction at a higher pace than the Swedish experimenters, which may influence the children’s response pattern. The actions produced by the infant were scored following Meltzoff’s (1985, 1988) guidelines. From that follows that the wooden extension would have to be folded through an arch of at least 45 degrees toward the vertical piece in order for the action to be scored as successful. The plastic egg would have to be shaken back and forth. The separate parts of the pull toy would have to be completely separated. For the string of beads no more than a third of it
would be hanging outside the cup for the action to be considered successful. The cup would have to be fully collapsed. The infant was awarded one point for each successful action, which allowed for a maximum of 5 points.

**Developmental measures.** The Ages & Stages Questionnaire for 14-month-olds consists of five scales assessing different areas of the child’s development. The areas included are communication, gross motor, fine motor, problem solving and personal-social. Each scale gives a maximum of 60 points. For the statistic analysis the subscales were combined to create a total score of maximum 300 points. The SECDI Words & Gestures questionnaire yields three main scores: vocabulary comprehension (number of words the child comprehends, max. 370 points), vocabulary production (number of words the child produces, max. 370 points) and Total Gesture Score (actions the child produces, max. 62 points). These scales were treated separately in the statistical analysis.

**ERP experiment.** Before doing the ERP analysis, the video and data were synchronized and used to determine when the infant looked away, and these sections were manually marked in the data file to be excluded from the analysis. The data was filtered using a bandpass of 0.5 to 20 Hz, and then segmented into epochs focusing on the second picture in the pair, in order to capture the establishment of an association between the first and the second picture. The epochs lasted from 150 ms before stimulus onset to 1500 ms after onset, thus including the entire presentation of the second picture, as well as 500 ms of the following white screen. Epochs were averaged and grouped according to the different conditions, which were the following: each of the eight presentations of the associated pictures, the familiar picture in a new combination (mixed condition), and the completely novel picture (novel condition).

For the automatic artifact detection procedure, the max-min amplitude accepted was set to 300 μV, which took into account the fact that young children’s EEG has larger amplitudes than adults, and a moving average of 40 ms was used. It was decided not to separately reject segments including eye blink and eye movement artifacts, considering that the infant electrode nets lacked orbital sensors and that eye artifacts are a relatively smaller problem in infant EEG data. Segments including more than 20 bad channels were rejected in the automatic procedure. A visual inspection was later conducted where in certain cases segments
with more than 20 bad channels were included, as long as the bad channels were evenly and peripherally distributed, thus not affecting the fronto-central electrodes of interest. During this stage, the earlier markings from the video inspection were taken into account, thus excluding segments where the infant was not attending to the material, unless the infant started to look away at the very end of a segment.

Channels marked as bad were replaced by the average of the ERP values of the surrounding electrodes. The data was re-referenced to an average reference. Average referencing was considered the best option for the present study, since using a mastoid reference would have resulted in an extensive amount of artifacts, as children tend to rub and grab the sides of the net. Vertex referencing makes the ERP waveforms of the electrodes closest to the vertex appear flat, and is therefore not a suitable method for Nc detection, since the Nc component has a fronto-central distribution. The data was also PARE-corrected, which compensates for the fact that it is impossible to evenly cover the entire scalp area with electrodes. Thereafter, baseline correction was performed, using the 100 ms before stimulus onset as a baseline. The data was averaged so that data from each subject generated a separate average ERP waveform for each condition. Grand averaging resulted in separate ERP waveforms for each condition across subjects.

**Statistical analysis.** A multiple regression was carried out on the data from the deferred imitation task and the questionnaires. The deferred imitation score was used as dependent variable, and the various questionnaire scales as independent variables. In addition, a Spearman Rank Order Correlation was performed on parameters of interest.

For the ERP analysis, a cluster of fronto-central electrodes (4, 5, 6, 11, 12, 13, 19, 20, 112, 118) where the Nc component was most prominent was used. The time window of 350-600 ms was used to define the Nc, and average amplitudes for each subject and each condition during this time window was then extracted for the statistical analysis. The effect of condition was computed using a one-way repeated measures ANOVA with condition (mixed, novel, presentation 8) as within-subject factor.
Figure 3. Layout of 128-channel HydroCel Geodesic Sensor Net. The cluster of electrodes chosen for analysis is highlighted.
Results

Deferred Imitation and Developmental Measures

The average number of successful target actions in the deferred imitation task was 2.78 within the 20 seconds time limit (SD = 1.45, N = 23). The results were spread across the entire scale, from 0 to 5. When using the more generous time limit of 30 seconds the mean score was 3.09 (SD = 1.47, N = 23). In the following steps of the analysis, the score for the 20 seconds time limit was used, as this is the time limit originally suggested by Meltzoff (1985, 1988). The mean ASQ total score was 229.72 (SD = 42.3, N = 18). For the SECDI Words and Gestures, the mean vocabulary comprehension score was 103 (SD = 78, N = 17), the mean vocabulary production score was 21.94 (SD = 27.49, N = 17), and for the total gesture score the mean was 29.59 (SD = 9.27, N = 17). Note that for one child only ASQ data was submitted, therefore N = 17 for the SECDI questionnaire, but N = 18 for ASQ.

Table 1

Descriptive statistics for DI and developmental measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>DI 20 sec</td>
<td>2.78</td>
<td>1.45</td>
<td>0-5</td>
<td>23</td>
</tr>
<tr>
<td>DI 30 sec</td>
<td>3.09</td>
<td>1.47</td>
<td>0-5</td>
<td>23</td>
</tr>
<tr>
<td>Ages &amp; Stages Questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASQ Total</td>
<td>229.72</td>
<td>42.3</td>
<td>145-300</td>
<td>18</td>
</tr>
<tr>
<td>ASQ Communication</td>
<td>43.89</td>
<td>13.56</td>
<td>20-60</td>
<td>18</td>
</tr>
<tr>
<td>ASQ Gross Motor</td>
<td>52.5</td>
<td>11.54</td>
<td>20-60</td>
<td>18</td>
</tr>
<tr>
<td>ASQ Fine Motor</td>
<td>45.83</td>
<td>9.89</td>
<td>30-60</td>
<td>18</td>
</tr>
<tr>
<td>ASQ Problem Solving</td>
<td>46.11</td>
<td>11.83</td>
<td>20-60</td>
<td>18</td>
</tr>
<tr>
<td>ASQ Personal-Social</td>
<td>41.39</td>
<td>12.81</td>
<td>20-60</td>
<td>18</td>
</tr>
<tr>
<td>SECDI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary comprehension</td>
<td>103</td>
<td>78</td>
<td>12-227</td>
<td>17</td>
</tr>
<tr>
<td>Vocabulary production</td>
<td>21.94</td>
<td>27.49</td>
<td>0-111</td>
<td>17</td>
</tr>
<tr>
<td>Total gesture</td>
<td>29.59</td>
<td>9.27</td>
<td>15-49</td>
<td>17</td>
</tr>
</tbody>
</table>

The multiple regression analysis generated a significant correlation between DI performance and SECDI vocabulary production score (r = 0.59, p < 0.01). This measure alone explained
31% of the variance in DI performance, when R squared was adjusted for sample size. None of the other questionnaire measures were significantly correlated with the outcome variable.

Table 2

Pearson correlation coefficients between behavioral measures and DI 20 sec

<table>
<thead>
<tr>
<th></th>
<th>Pearson r</th>
<th>Sig.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASQ Total</td>
<td>0.15</td>
<td>0.27</td>
<td>18</td>
</tr>
<tr>
<td>Vocabulary comp.</td>
<td>0.29</td>
<td>0.13</td>
<td>17</td>
</tr>
<tr>
<td>Vocabulary prod.</td>
<td>0.59</td>
<td>0.01</td>
<td>17</td>
</tr>
<tr>
<td>Total gesture</td>
<td>0.13</td>
<td>0.31</td>
<td>17</td>
</tr>
</tbody>
</table>

Considering the large standard deviation for the vocabulary production score, a Spearman Rank Order Correlation was performed in order to test the strength of the relationship between this measure and DI score. The Spearman correlation coefficient was 0.62 and was found to be significant at the 0.01 level.

**ERP Results**

*Nc findings.* Grand-average waveforms showed a negative deflection across fronto-central electrodes (4, 5, 6, 11, 12, 13, 19, 20, 112, 118), with a latency of approximately 350-600 ms, throughout all presentations of the initially associated pictures. The average amplitudes within the analyzed time window differed across conditions, and for presentation 6 the average amplitude was in fact positive. Visual inspection revealed that this was due to the entire waveform being more positive, and a relatively negative deflection was still present within the specified time window.
Figure 4. Grand-average waveforms at fronto-central electrodes for each presentation condition and for an average of all 8 presentations. The Nc time window is highlighted. Note that negative is plotted upwards in this and all other figures.
The two conditions in which the associations were violated differed greatly from one another. The novel condition elicited an ERP pattern similar to that found in most of the initial presentations. However, for the mixed condition no Nc-like pattern was found. Instead, the waveform was consistently positive.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures}
\caption{Grand-average waveforms at fronto-central electrodes for the mixed and novel conditions.}
\end{figure}

It is reasonable to assume that the pattern found in all conditions but the mixed condition, corresponds to an Nc component.

\textbf{Nc modulation.} Average amplitudes during the time window of interest (350-600 ms) were calculated for each condition. According to the second hypothesis, Nc amplitude would increase during the first few presentations, and decrease during the last ones. However, the Nc modulation found did not follow the predicted pattern. Rather, the amplitude alternately increased and decreased across conditions. The highest amplitude was found in the middle of the series of presentations. The findings that contradicted the predicted pattern were the fact that the average amplitude was lower in presentation 3 than in presentations 2 and 4, and that it increased during the final two presentations.
Comparisons between main conditions. The third hypothesis stated that the average Nc amplitude would be greater in response to the violating conditions than to the final presentation of the associated pairs. The obtained results did not follow this prediction. As reported above, the ERP waveform for the mixed condition did not show an Nc component at all. Although the novel condition did show an Nc, the amplitude was in fact lower than that for presentation 8.
window. The descriptive statistics showed that the standard deviations for the different conditions were fairly large.

Table 3

<table>
<thead>
<tr>
<th></th>
<th>Mean Amplitude (μV)</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed</td>
<td>4.52</td>
<td>3.64</td>
<td>5</td>
</tr>
<tr>
<td>Novel</td>
<td>-3.29</td>
<td>10.13</td>
<td>5</td>
</tr>
<tr>
<td>Presentation 8</td>
<td>-7.14</td>
<td>16.25</td>
<td>5</td>
</tr>
</tbody>
</table>

The difference between the three conditions was non-significant \( F(2,8) = 1.93, p = 0.22 \). The p-value is reported using the Greenhouse-Geisser correction for violations of sphericity, while uncorrected degrees of freedom are presented. The effect size, calculated using partial eta squared, was 0.33.
Discussion

The aim of this study was to create and evaluate a stimulus material appropriate for ERP recording that would tap the same underlying memory processes as the deferred imitation task. As well as evaluating the design, the idea was to establish a connection between the DI results and the pattern of electrophysiological activation in response to the ERP stimuli. Since, unfortunately, a large proportion of the ERP data was lost, it was not possible to investigate such a relationship statistically. Therefore the ERP results and the behavioral measures are primarily discussed separately.

Implications of Results

Deferred imitation and developmental measures. The results from the deferred imitation task successfully differentiated between the children. The mean score fell close to the middle of the scale (2.73/5), with a range from 0-5. These results are similar to those obtained by Strid, Tjus, Smith, Meltzoff and Heimann (2006) from a group of 26 Swedish children, (M = 1.5, SD = 0.81), when accounting for the fact that only three toys were used.

The results indicate that, as established by earlier studies, the deferred imitation task is a useful and stable test. While administrating the test, however, the experimenters were given the impression that there may be factors other than memory ability that influence a child’s performance on deferred imitation. Some children were very cautious and refrained from interacting with the toys at all, naturally leading to a low score. In these cases it is difficult to determine whether the low score was in fact due to the child not remembering the target actions, or if the child’s behavior could be related to a certain temperament. Another factor that may affect a child’s performance is interaction with the parent present in the room. Parents were not meant to talk or in other ways interact with the child during the task, but in certain cases the child was very focused on the parent, showing the toys to him or her, and perhaps due to this shift of focus failed to perform the target actions. It also occurred, in two cases, that the parent actively interacted with the child, commenting on the toys presented. In these particular cases, however, the children did not appear to be affected by the interaction. It would, of course, have been problematic if many parents interacted with, and actively encouraged, their children during the task.
It also became clear that DI is a test that requires great accuracy from the experimenter in order to be reliable. Seemingly small mistakes on part of the experimenter may run the risk of affecting the child’s performance. When watching the video recordings of the DI sessions, it was discovered that in a few cases the toys were placed slightly farther from the child than intended. In each case, the toys were still within the child’s reach, but the additional effort required on behalf of the child to reach them may have affected performance. In these particular cases the children still performed well, but if a poor score was obtained it would have been difficult to determine the true reason behind it.

The scores on the Ages & Stages Questionnaire for 14-month-olds were compared to Norwegian normative data (Janson & Smith, 2003). On all 5 subscales, all children obtained scores that are equal to or above the Norwegian cutoff scores. These cutoff scores were obtained by subtracting 2 standard deviations from the mean, and clinicians are recommended to pay extra attention to children achieving below these scores. The results from this study indicate that the participants were all typically developing infants.

For the SECDI, normative data from Swedish children has been provided by Eriksson and Berglund (2002). When comparing the results from the present study to the normative data, the vocabulary production score of this sample (M = 21.94) stands out as much higher than the estimated population mean (M = 8). Even after excluding the outlier, the mean score is unusually high (16.38), falling in the 70\textsuperscript{th} to 80\textsuperscript{th} percentile according to the normative data. This may be due to the fact that many parents of the children included in the sample had an academic background. Many parents were in some way connected to the university, as students or researchers, and most had a high level of education. Therefore, these children are most likely not as representative of the general population as those in Eriksson and Berglund’s study, where 228 children were randomly selected from a county in Sweden.

From the reasoning above, it could be expected that the other two SECDI scores would also be higher in this sample than in the general population, which is not the case. The vocabulary comprehension score (103) closely corresponds to the estimated population mean (107.5). However, the standard deviation is very large (78) and the individual scores range from 22 to 227. It seems natural that a measure such as vocabulary comprehension shows a large spread
in this age group, as children develop at very different paces with regards to language. The sample included is also relatively small (N = 17). However, it is also reasonable to question the stability of a measure that is based on parental report of how many words their child understands. Compared to vocabulary production this is a rather subjective measure, that probably reflects the parents’ attitudes as well as the child’s actual word comprehension.

The average total gesture score of the present sample (29.59) falls below the mean of the normative data (M = 36), corresponding to the 30th percentile. This may be a random effect due to the small sample included. At first glance, this result seems to conflict with the theory that the parental academic background has influenced the developmental scores. However, it is possible, and reasonable, that this academic background has a larger influence on language development than other, non-verbal aspects of development.

In the statistical analysis, a relationship was found between vocabulary production score and deferred imitation performance. Both parametric and non-parametric tests revealed large significant positive correlations between these two measures (Pearson’s r = 0.59, Spearman Rank-Order Correlation = 0.62). Whether high productive vocabulary leads to high deferred imitation skills, or the other way around, is impossible to say based on these results alone, and it is also possible that a third factor influences both vocabulary production and deferred imitation performance.

In the case of a direct causal relationship between the two variables it seems reasonable that declarative memory may have an effect on word learning. When learning words, children have to build associations between words and their referent, a process naturally relying on memory. From this reasoning follows that a correlation between DI and vocabulary comprehension should be expected as well. That such a relationship was not found is in one way peculiar, but one should consider the argument already made concerning the apparent subjectivity of this particular measure. It is, perhaps, not unreasonable that the more robust and objective measure of language development, productive vocabulary, shows a stronger relationship with another variable. Vocabulary comprehension, in fact, showed the second largest correlation with deferred imitation, although it did not reach statistical significance. It
is imaginable that a stronger relationship could have existed, if a more objective way of measuring this entity was found.

A third factor, influencing both variables, could also account for the discovered correlation. One possible such factor is temperament. Both productive vocabulary and deferred imitation are productive measures. The DI task measures if the child produces a certain action, just like productive vocabulary measures how many words the child actually says. There may exist general differences between children’s tendency to actively display their abilities. It is conceivable that some children have the capacity to produce words without doing so, for example waiting until they are more confident at pronouncing them. Similarly, some children may remember the DI target actions without producing them. Likewise, other children may be more inclined to demonstrate their abilities right away, without hesitation. Personality traits such as introversion versus extroversion could possibly be responsible for this difference.

Aside from temperamental influences, another factor that could possibly affect both variables is parental stimulation. Extensive verbal stimulation from parents may have the effect of increasing the child’s productive vocabulary, as well as offer opportunities to practice and develop memory skills.

Productive vocabulary has been identified as a strong predictor in another study similar to the present one. As previously discussed, Torkildsen et al. (2008, in press) found that productive vocabulary predicts fast mapping abilities, that is the ability to quickly establish connections between words and their referents. In that study the Nc component was used to measure establishment of word-referent associations, in the same way that it was used in this study to reflect the learning of picture associations. It is obvious that fast mapping involves memory processes, which is why the Nc was a relevant component to study. The design in Torkildsen et al.’s study follows the same fundamental principles as the ERP design presented in this study, and could essentially be considered a memory task. The idea underlying the present study is that the ERP design tested would reflect the same memory processes as the deferred imitation task. A correlation between deferred imitation performance and ERP responses to the computerized stimuli was expected. As this relationship was not possible to explore directly, it is interesting that a measure that was found to predict responses to Torkildsen et al.’s similar stimulus material also was found to predict the task fundamentally related to the
ERP design tested in this study, namely deferred imitation. This could be taken as support for the assumption that the present ERP design does in fact tap the same underlying memory processes as deferred imitation. To further speculate, one could imagine that memory capacity, as reflected by deferred imitation performance and response to the fast mapping task in Torkildsen et al.’s study, is essential for development of productive vocabulary.

Electrophysiological responses. As presented previously, four hypotheses were formulated pertaining to the ERP experiment. The first one stated that an Nc component would be present in response to the second picture of the pairs. A negative deflection at 350-600 ms was indeed identified and interpreted as an Nc. The presence of this component indicates that the children were attentive and aware of the pictures, which suggests that a continuation of the experiment with some modulations would be meaningful. It seems reasonable to assume that the experiment, with a revised design, has the potential to show the expected patterns.

The second hypothesis regarded Nc modulation. It was predicted that the average amplitude in the Nc time window would increase during the first few presentations and then decrease during the latter ones. Such a pattern was not found. As expected, the largest amplitude was found in the middle of the presentations, whereas the amplitude was low in the beginning and after the fourth presentation. A few anomalies, however, made the modulation divert from the expected pattern. The electrophysiological response to the third presentation showed an unexpected reduction in amplitude, and for the last two presentations the amplitude started to rise again. These findings are difficult to explain, and the unstable pattern is likely due to the small number of participants and too little data per participating child.

The third hypothesis suggested that Nc amplitude would be greater in response to the two violation conditions (mixed and novel), than to presentation 8. Again, the results were not according to the hypothesis. Perhaps the most unexpected result was a positivity that was found in response to the mixed condition. A reasonable interpretation seems to be that the infants failed to understand the associations between pictures. If interpreting the material as single pictures rather than as pairs, the mixed condition would simply be understood as presentations of familiar pictures already seen eight times. The absence of an Nc would thus
be understandable. However, the increase in amplitude during the final two presentations does not fit this explanation, and may again be due to the small amount of data. Statistically, the differences in amplitude were non-significant. Considering the small amount of participants and the large standard deviations within conditions, this is not surprising. The large standard deviations in particular, contribute to the unstable results and unclear patterns.

The fourth hypothesis stated that a difference concerning Nc modulation as well as Nc responses to the violation conditions would be expected between high and low achievers on the deferred imitation task. Due to the small amount of usable ERP data, it was not possible to divide participants into separate groups, which would have been necessary in order to test this hypothesis.

**Suggestions for Improving the Design**

Because of the limited statistical power, it is hard to draw strong conclusions about the meaning of the ERP results. An important aim of the study was however to evaluate the design, and the remaining part of this discussion will focus on possible improvements that may increase the efficiency of the experiment. Attempts already made in this direction will also be discussed.

There appears to be two major problems to correct for as the study continues. First of all, it seems probable that the infants understood the material as single pictures rather than pairs. Since breaking associations between pictures is a crucial part of the experiment, it is essential to alter the design so that the children are in fact able to establish the associations that are to be broken. Secondly, not enough data was retained from each individual. The children were not attentive enough during the presentations, and the many conditions resulted in too few segments in each. The violations were presented at the end of each block when the infants in many cases had already lost interest. It would be desirable to find means of increasing the infants’ attention so that more blocks could be presented to each child, and a larger percentage of each block could be retained.

Considering the issue of the children not understanding the associations between pictures, the form of presenting the pictures must be altered. The pictures are now presented after each
other, and it appears to be unclear to the infants that the red screen and the sound effect are supposed to indicate association. Presenting both pictures of a pair simultaneously may be a possible solution. When it comes to the problem of the children becoming inattentive, an attempt must be made to make the stimulus material more appealing to infants. Using animations instead of still pictures may be a good idea, since moving pictures tend to attract more interest. Placing the pictures in a context could also contribute to making the material more interesting. A possible suggestion for a new stimulus material that takes all issues discussed into consideration, is to place an animated wizard at the center of the screen, and have him pull an animal out of a hat, and thereafter an object out of a second hat, so that the two are displayed together. Though probably more interesting, a design like this could possibly increase the number of eye movement artifacts, and it would therefore be crucial to display the animal and object as close to the center of the screen as possible.

During the course of the study, efforts have continuously been made to enhance the infants’ attention. A screen was used to shield off the child from various distractions in the lab. The light was dimmed in order to make the computer screen more salient. Using a larger screen may be a possible further improvement, but that might again lead to an increased number of eye artifacts. After running a few experiments, it was concluded that the infants were not paying enough attention to the stimulus material, and some alternations were made. The presentation rate was increased to reduce the risk of infants getting bored and looking away. At the same time, the number of presentations per trial block was increased. A possible solution to the current problem of having too few usable segments per condition would be to keep the fast presentation rate, but again decrease the number of repetitions, thus decreasing the number of conditions. Each block would then be shorter, allowing for an increase in the total number of blocks completed.

In conclusion, this pilot study has confirmed that deferred imitation is a stable measure that is well suited for use in future modifications of the present research project. Further, it has shown that interesting relationships can be found between performance on this task and other developmental measures. The strong correlation discovered between DI performance and productive vocabulary can be the object of further research, not only in the present research project. Considering the ERP experiment, the most important insights regard possible
improvements of the design. Most importantly, the associations between pictures need to be strengthened, and measures need to be taken to ensure that as much data as possible from each individual will be secured. With these modifications, there are indications that the proposed hypotheses may be successfully tested in the next phase of this research project.
References


44


Appendix A: Participant Information

Ett nytt forskningsprojekt startar:
Långtidsminne och nervsystemets utveckling hos små barn

Vid Linköpings och Lunds universitet genomförs under åren 2008-2010 ett forskningsprojekt där målet är att bättre förstå hur minnet utvecklas hos barn som är 9 och 14 månader gamla. Vi är särskilt intresserade av att studera långtidsminne hos barn innan språkutvecklingen tagit fart, framförallt vill vi undersöka i vilken grad mätningar av hjärnans aktivitet kan hjälpa oss att bättre förstå när och hur små barn minns. Projektet genomförs i samarbete mellan Institutionen för beteendevetenskap och lärande (IBL), Linköpings universitet, och Institutionen för psykologi (IPL), Lunds universitet.


Den studie som nu genomförs har som huvudsyyfte att undersöka hur fysiologiska processer kan hjälpa oss att bättre förstå individuella variationer i förmågan i hur 9- och 14-månader gamla barn minns olika händelser. Vi tror att skillnader i hur barn minns avspeglar mognadsmässiga olikheter i nervsystemets utveckling. Olikheter som kan fångas upp av elektrofysiologiska metoder och ge oss en bättre förståelse för hur psykologiska minnesobservationer samvarierar med fysiologiska mätningar av hjärnans elektriska aktivitet. Även om den nuvarande studien bara syftar till att kartlägga hur den "normala" eller "typiska" utvecklingen ser ut, hoppas vi att resultaten i förlängningen kommer att leda till att bättre metoder skapas för att bedöma små barns utvecklingsnivå.

Vi vill nu fråga dig om du kan låta ditt barn delta i studien. I ett första skede är det 14 månader gamla barn som vi kommer att undersöka, medan studien längre fram kommer att fokusera på 9 månader gamla barn. Ett deltagande innebär ett besök på Humanistlaboratoriet i Lund (som finns på Språk- och litteraturcentrum) som tar ca 1 ½ timme. De metoder vi använder är väl utprövade och brukar upplevas som motiverande för barnen. Vissa delar innebär att barnet får se någon använda en leksak och senare imitera beteendet, och under ett moment mäter vi även hjärnans aktivitet (EEG). Mätning av hjärnaktiviteten görs genom att man sätter på barnet ett registreringsnät (elektroder som fångar upp hjärnans elektriska

Detta ger möjlighet att samla in värdefull information om hjärnaktiviteten hos barn vid inlärning. EEG-undersökningar har länge använts för att undersöka hjärnans funktioner och medför inga obehag eller risker för den som undersöks.

Att delta i studien är naturligtvis helt frivilligt och skulle du säga ja nu, men senare ångra dig så går det också bra. Du kan när som helst dra dig ur utan att behöva motivera orsaken till detta. Den analys vi kommer att genomföra baserar sig på gruppjämförelser och inga personuppgifter kommer att läggas in i någon datafil (t ex personnummer eller namn).

Ansvariga för projektet är professor Mikael Heimann, Linköpings universitet, docent Magnus Lindgren, Lunds universitet samt fil dr Mikael Johansson, Lunds universitet. Om du och ditt barn väljer att delta i studien kommer ni att träffa psykologkandidaterna Kristina Magnusson och Emilia Thorup som utför det praktiska arbetet i forskningsprojektet.

Om du kan tänka dig att låta ditt barn medverka så skall du fylla i bifogad samtyckesförklaring, samt kontakta Kristina Magnusson eller Emilia Thorup på nedanstående telefonnummer eller e-mail för att boka en tid. Har du funderingar eller frågor så svarar vi gärna på dessa.

Linköping och Lund i januari 2008

Kristina Magnusson
psykologkandidat
Lunds universitet

Emilia Thorup
psykologkandidat
Lunds universitet

Mikael Heimann
Professor/ leg psykolog
Linköpings universitet

Magnus Lindgren
docent
Lunds universitet

Mikael Johansson
universitetslektor
Lunds universitet

Institutionen för beteendevetenskap och lärande Linköpings universitet
Hus I
581 83 Linköping

Institutionen för psykologi Lunds universitet
Box 213
221 00 Lund
Projekt: Långtidsminne i spädbarnsåldern och dess elektrofysiologiska korrelat.

Föräldrainformation

Vi vill med denna information erbjuda Er medverkan i en forskningsundersökning om minnesutveckling under barns första två levnadsår.

Vi kommer att dels låta barnen observera och utföra vissa handlingar med komplicerade leksaker (motsvarande sk. ”pedagogiska leksaker”), dels göra EEG-undersökningar medan barnen tar del av en minnesuppgift på en datorskärm. EEG-data lagras i dator och sammanställs i efterhand. EEG-undersökningar har länge använts för att undersöka hjärnans funktioner och medför inga obehag eller risker för den som undersöks.

Undersökningen går till så att barnet får sitta i vårdnadshavarens knä, och ett registreringsnät sätts på huvudet. Sedan registreras EEG-aktiviteten med hjälp av en dator medan barnet tittar på ett bildspel på en datorskärm. Detta kan ge oss värdefull information om hjärnaktiviteten hos barn vid inlärning. Undersökningen tar totalt mellan 1-2 timmar.

Deltagandet är naturligtvis frivilligt och undersökningen kan avbrytas när helst Ni önskar. Alla uppgifter behandlas konfidentiellt. Ingen vetenskaplig redovisning kommer att göras så att något enskilt barn kan känna igen. Vi kan tyvärr inte ge information om individuella barn.

Vi har tagit del av den skriftliga informationen och medger att vårt barn deltar i studien.

--------------------------------------------------
Ort, datum Underskrift