Age Differences in Unsupervised Statistical Learning

From Temporal Sequences of Visual Shapes

Daryaneh Badaly

Georgetown University

**Key Concepts:** age differences; aging; statistical learning; extraction of probabilities; higher-order temporal structure; visual shape sequences
Abstract

Research has shown that infants and young adults can learn to distinguish familiar from unfamiliar stimulus sequences differing in transitional probability (e.g., Fiser & Aslin, 2002). The current study examined whether older adults are also capable of statistical learning from a continuous sequence of shapes. In Experiment 1, after watching an animated stream of shapes, young and older adults took two different recognition tests. On both tests, subjects chose familiar shape triplets having joint probabilities of .083, from unfamiliar alternatives having null joint probabilities. Both age groups showed significant learning, in that they were able to distinguish familiar sequences that occurred in the training movie from unfamiliar, novel sequences. While older adults were in the direction of being poorer, the age difference was not significant. As well, a card sorting task was given in which participants were asked to order twelve shape cards into four sequences of three elements, based on the training movie. Young subjects generated significantly more shape sequences than the older adults. Experiment 2 minimized fatigue and attention distracting factors by reducing the training movie from 12 to 6 minutes and by placing last the battery of neuropsychological tests originally given before the familiarization session. An age difference was detected on the recognition tests, and young adults showed significant learning while older adults’ learning was only marginal. These findings show that older adults are capable of statistical learning from visual shape sequences, but to a lesser degree than young individuals.
Age Differences in Statistical Learning

From Temporal Sequences of Visual Shapes

Why examine statistical learning?

While the world may appear to be a conglomerate of disjoint, random events, people find order within the chaos. Not only do individuals find patterns and regularities to past and present happenings, but they also predict future circumstances, allowing them to make educated decisions. “Predictions serve to anticipate the next event in a temporal sequence thereby reducing uncertainty and enabling the effective preparation of motor responses” (Hunt & Aslin, 2001). Predictions depend on extracting statistical information from the environment. As predictions are central to living in a temporal milieu, statistical learning is of special importance.

Statistical learning notably plays an important role in language acquisition. Research shows that infants exhibit learning from auditory stimulus sequences based on statistical relationships. 8 month old infants can segment speech into words using only the statistical relationships which exist between neighboring speech sounds (Saffran, Aslin, & Newport, 1996). In Saffran, Aslin and Newport’s study (1996), infants were exposed to two minutes of a continuous speech stream of nonsense words, containing no acoustic or prosodic cues to word boundaries. They showed significant discrimination of words (syllable sequences heard during training) from non-words (sequences never appearing during familiarization) and part-words (sequences combining the last syllable of one word with the first two syllables of another), showing preference for the novel patterns. Words were distinguished from non-words and part-words by the transitional probabilities between their sounds. The transitional probabilities of syllable pairs were
higher for syllables within a word than between words. Words also differed from non-words and part-words as their sequence of syllables had higher frequencies. In a follow-up study, Aslin, Saffran and Newport (1998) showed that infants discriminate between words and part-words on the basis of differing transitional probabilities when frequencies are held constant. Moreover, infants demonstrate statistical learning not only in their capacity to segment speech streams, but also in their ability to subsequently discover the permissible or grammatical orderings of the words based on their distribution in sentences (Saffran & Wilson, 2003). Infants can thus track multiple levels of regularities.

Segmenting a speech stream based on statistical information is not unique to infants. Research indicates that both children and adults can discriminate words in an artificial language based on transitional probabilities (Saffran, Newport, Aslin, 1996; Saffran et. al., 1997). In Saffran et. al.’s study (1997), while first grade children and adults performed a diversion task, a continuous speech stream of randomly concatenated nonsense words played in the background; participants were not encouraged to pay attention to the auditory stimulus. The only cues to word boundaries were transitional probabilities between syllable pairs. When later asked whether words or non-words were more familiar, both children and adults chose above chance the sequence from the training session.

Regularities in non-speech sounds can also be extracted using statistical information (Saffran, Johnson, Aslin, & Newport, 1999). Adults and 8 month old infants were presented a continuous tone stream of randomly concatenated sequences. Sequences consisted of three pure tones of the same octave and the same length. In the training stimuli, sequence boundaries could only be distinguished by statistical differences. When
asked to distinguish between tone orderings that were familiar, not familiar, and partially
familiar, adult subjects showed significant learning. As well, infants reliably recognized
known sequences.

Research indicates that statistical learning from auditory sequences is not specific
to humans (Hauser, Newport, & Aslin, 2001; Toro & Trobalón, 2005). Hauser, Newport,
and Aslin (2001) assessed statistical learning in cotton-top tamarins (*Saguinus oedipus*)
with similar stimuli and testing materials used with human infants (Saffran, Aslin, &
Newport, 1996). The primates first listened to 20 minutes of a continuous stream of
synthetic speech syllables, in which four different trisyllabic nonsense words were
concatenated. The monkeys were then presented with words, non-words, and part-words.
The different groupings of syllables differed only in the frequency or probability of their
appearance in the auditory training stream. The cotton-top tamarins showed significant
learning, in that they were able to distinguish words from non-words and part-words.

Even though cotton-top tamarins are capable of statistical learning, it should be
noted that the capacities of human and non-humans differ. As Conway and Christiansen
(2001) indicate, while human infants are trained with 2 minutes of stimulus, the non-
human primates were given 20-minutes of training. As well, tamarins and adult humans
differ in their statistical learning of non-adjacent dependencies with auditory sequences
(Newport & Aslin, 2004; Newport, Hauser, Spaepen, & Aslin, 2004). Furthermore, Toro
and Trobalón (2005) showed that rats can segment words from a speech stream using
frequencies but not transitional probabilities which are used by humans. Also, rats did not
exhibit significant learning with complex, statistical patterns involving non-adjacent
elements and abstract grammar-like rules. Still, non-human species are capable of computing some statistical regularities.

Statistical learning extends beyond the auditory realm to the visuomotor domain. Using a variation of the serial reaction time task, Hunt and Aslin (2001) found that adults show learning of both joint and conditional probabilities in visuomotor sequences. Subjects responded on a button box to pairs or triplets of lights which were illuminated along a semicircle of options. The pairs or triplets of elements were randomly concatenated into a continuous sequence. Results showed that reaction times were faster for more predictive sequences, those with high conditional and joint probabilities. As well, if either the conditional or joint probabilities were held constant, participants were able to make use of differences in the second statistic. However, results suggest that subjects rely on conditional probabilities to estimate predictability within the sequences.

The visual domain also makes use of statistical learning, even in the absence of motor sequencing. Infants, sensitive to higher-order statistical features in stationary scenes of distinct shapes, can distinguish between familiar and unfamiliar patterns (Fiser & Aslin, 2002b). Fiser and Aslin (2002b) exposed 9 month old infants to multi-element shape scenes made up of base pairs and noise elements in varying combinations. Infants demonstrated a preference for element pairs that had a higher cooccurrence frequency than other pairs, as well as for pairs that had a higher predictability or conditional probability between elements of a pair. When cooccurrence frequency was held constant, infants still distinguished between pairs of varying predictability. Furthermore, Kirkham, Slemmer, & Johnson’ study (2002) demonstrated that infants as young as two months of age,
familiarized with temporal sequences of discrete shapes, can distinguish novel visual sequences using transitional probability.

Adults, as well, show statistical learning in the visual domain. Young adults are sensitive to statistical information in stationary spatial structures (Fiser & Aslin, 2001). In Fiser and Aslin’s study (2001), subjects were shown 144 shape scenes during 7 minutes of training. Twelve basic shapes were used in the training and organized into 6 base pairs of different orientations (horizontal, vertical, and oblique). Each shape scene consisted of 3 base pairs presented on a 3x3 grid. After the training session, subjects were asked to choose from a base pair and a non-base pair which two-element set was more familiar. The joint probability for base pairs was .50, while the probability of non-base pairs was less than .02. Subjects reliably distinguished base pairs from non-base pairs when learning was position-dependent and position-independent. Furthermore, subjects were able to discriminate base pairs from non-base pairs using conditional probabilities when joint probabilities were held constant.

Young adults also show statistical learning from temporal shape sequences (Fiser & Aslin, 2002a). In Fiser and Aslin’s study (2002a), twelve shapes were organized into four base triplets. During a familiarization session, participants watched an animated movie made up of randomly concatenated base triplets. The joint probability for any base triplet was .083. After the training session, subjects were asked to choose the more familiar sequence, comparing a base triplet to either an impossible triplet (null probability) or a part triplet (joint probability of .027). Participants reliably distinguished base triplets from both impossible and part triplets. Furthermore, subject learned to distinguish
familiar from unfamiliar shape pairs using conditional probabilities, when joint probabilities were held constant.

Another sensory domain from which individuals are capable of extracting statistical information is the tactile sphere (Conway & Christiansen, 2005). Conway and Christiansen (2005) presented participants with sequences of vibrations applied to the finger tips of their preferred hand. The sequences followed an artificial grammar. After being familiarized with a set of tactile sequences, subjects were tested on their ability to classify novel sequences as legal (generated with the same rules as the training sequences) or illegal (not following the same finite-state grammar). Legal and illegal sequences differed from each other in terms of the statistical relationship between adjacent elements. Subjects reliably classified sequences using cooccurrence statistics.

While different modalities mediate statistical learning from sequential input, quantitative and qualitative modality constraints have been found (Conway & Christiansen, 2005). Auditory learning exhibits a quantitative advantage compared to tactile and visual learning even when training performance and stimulus element perceptibility are controlled for. Furthermore, tactile learning places a heavier importance on initial elements while auditory learning relies more on final input.

Why evaluate age differences?

Both decline and stability of cognitive abilities are evident in old age. Three types of age-related changes exist: life-long declines, late-life declines, and life-long stability (Hedden & Gabrieli, 2004). Life-long declines affect information processing, working memory, and episodic memory. Late-life declines, on the other hand, affect short-term memory, vocabulary, and semantic knowledge. Meanwhile, autobiographical memory,
emotional processing, and automatic processes remain relatively unchanged (Hedden & Gabrieli, 2004).

Representation and control are linked to whether cognitive abilities are maintained or worsened in old age (Craik & Bialystok, in press). Representations, or schemas of knowledge, remain relatively stable in old age, though the formation of new representations becomes more difficult and access of old schemas depends on use and practice. Meanwhile, cognitive control – “the set of fluid operations that enable intentional processing and adaptive cognitive performance” – wanes after young adulthood. Control is linked to the “ability to inhibit attention to irrelevant stimuli, to maintain task set and select choices in line with the current goals, to hold information in working memory, and to reflect on integrated higher-order rules” (Craik & Bialystok, in press). As processing higher-order rules, such as statistical regularities in temporal sequences, depends on control, deficits in control during old age may predict difficulties in learning from statistical information.

Prull, Gabrieli and Bunge (2000) further outline age-related deficits. While impairments in non-declarative memory vary, declarative memory deficits are more pronounced for strategic than non-strategic tasks. Prull, Gabrieli and Bunge (2000) review different strategic tests on which older adults perform less well than young adults. Specifically, impairments on frequency and temporal ordering tasks suggest statistical learning may be impaired in older adults.

In addition, studies involving explicit memory tests such as the weather prediction task and association tasks indicate that older adults may have difficulty with statistical learning. Healthy older adults are impaired on the weather prediction task, a measure
which relies on awareness of underlying structure, hypothesis testing ability, and working memory capacity (Price, 2005). The task has a probabilistic structure, where learning depends on information accrued over many trials. Similarly, tasks dealing with statistical learning in temporal sequences present a probabilistic structure. Moreover, research with verbal associations indicates that older adults have a “decreased ability to encode and retrieve associations among units of information or attributes within events” (Naveh-Benjamin, 2000). If older individuals have trouble with associating units, they may show an inferior capacity to young adults in remembering elements connected to one another by statistical relationships.

More specifically, research indicates that statistical learning based on auditory sequences is impaired with age (Love, 2000). In Love’s experiment, subjects engaged in a serial reaction time task while, as in Saffran et al. (1997), a stream of randomly concatenated nonsense words played in the background. Subjects were then tested on discrimination between familiar and unfamiliar words, which only differed in transitional probabilities during the training. Older subjects did not perform significantly above chance. Young participants, on the other hand, did show learning indicating that there was an age difference. However, the study’s small n should be noted, as the lack of power may have affected significance levels. Furthermore, older subjects’ impaired performance may be due to difficulties with speech perception; the study does not indicate whether or not older individuals could in fact hear correctly the artificial speech sounds.

The current study, which comprised of two experiments, examined aging and statistical learning of visual sequences. First, the goal was to replicate Fiser and Aslin’s
results (Experiment 1; 2002a) with young participants. Second, the study aimed to
determine whether there exist age differences in adults’ sensitivity to temporal structure
of visual shape sequences differing in statistical features. Young and older adults were
tested using the stimuli and testing material from Fiser and Aslin (Experiment 1; 2002a).
In Experiment 1, a change was made to the stimuli: its length was doubled to 12 minutes.
As it was hypothesized that older adults would show significantly inferior learning to
young adults, familiarization was lengthened so that any deficits noted would not be due
to inadequate training. In Fiser and Aslin’s set of experiments (2002a), on those with
increased training, subjects tended towards improved performance. The present
Experiment 2, on the other hand, conserved Fiser and Aslin’s original length of
presentation, 6 minutes (2002a).

A second change was made to Fiser and Aslin’s procedure (2002a). In addition to
the animated recognition test used by Fiser and Aslin (2002a), participants completed a
stationary recognition task. As older adults show evidence of working memory deficits,
this second measure was added in which the sequences to compare did not need to be
held in memory as they were readily available on screen.
Experiment 1

Method

Participants

Twenty-four participants were tested, with two removed from the study. The resulting age groups were eleven younger adults (5 men, 6 women, mean age = 20.6, range 20-24) and eleven older adults (7 men, 4 women, mean age = 73.5, range 65-84) (Table 1). One of the removed participants was an older subject with a score of 23 on the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975). Subjects with MMSE scores below 25 were rejected because such scores may indicate dementia (Mungas, 1991; Grut, Fratiglioni, Viitanen, & Winblad, 1993); the study’s aim was to look at healthy aging. The other removed participant was a young adult who did not pay attention during the training session of the study. The subject’s lack of attention was indicated on a post-test questionnaire. As well, the subject answered a cell phone call in between the training and the testing sessions.

Overall, the two age groups were well matched. As seen in Table 1, the two groups did not differ on measures of vocabulary or digit span. However, older adults did have significantly lower MMSE scores, and, following a typical pattern, performed significantly worse on the Digit Symbol.

All participants were compensated $15, except for one younger volunteer who received extra credit for a Georgetown University psychology course. The older adults were recruited through newspaper advertisement; the younger adults were Georgetown University undergraduates who responded to e-mail requests for volunteers.
Table 1. Participant Characteristics for Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Young Adult</th>
<th>Older Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>5 M, 6 F</td>
<td>7 M, 4 F</td>
</tr>
<tr>
<td>Age**</td>
<td>20.64 (1.21)</td>
<td>73.55 (4.82)</td>
</tr>
<tr>
<td>Education*</td>
<td>14.18 (0.60)</td>
<td>16.09 (2.02)</td>
</tr>
<tr>
<td>Mini Mental Status Examination*</td>
<td>29.91 (0.30)</td>
<td>28.82 (1.17)</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>51.00 (10.83)</td>
<td>49.91 (9.89)</td>
</tr>
<tr>
<td>Digit Symbol: Coding**</td>
<td>97.91 (10.13)</td>
<td>60.55 (14.67)</td>
</tr>
<tr>
<td>Digit Symbol: Pairing@</td>
<td>15.55 (3.01)</td>
<td>9.30 (3.62)</td>
</tr>
<tr>
<td>Digit Symbol: Recall**</td>
<td>8.36 (0.81)</td>
<td>6.60 (1.35)</td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>12.36 (1.50)</td>
<td>11.82 (2.52)</td>
</tr>
<tr>
<td>Digit Span Backward</td>
<td>7.91 (2.12)</td>
<td>7.73 (1.95)</td>
</tr>
</tbody>
</table>

p < .01 *
p < .001 **

@ Score for older adults is based on 10 subjects; one participants’ data were removed because of an intervening fire alarm during the test.

Stimuli

Participants were presented a training movie on an iMac with a 1,024-pixel x 768-pixel, 15 inch monitor. The training movie, kindly provided by Fiser & Aslin, replicated the procedure of their Experiment 1 (2002a). In the continuous movie, a black bar remained at the center of the screen. From behind the bar emerged, one at a time, one of twelve distinct shapes (Figure 1). Each shape moved back and forth horizontally, with a constant speed. Once a shape emerged from behind the black bar, it would proceed to the edge of the screen, then return in the opposite direction behind the black bar. The following shape would then emerge from the opposite side of the rectangle (Figure 1).

The order of the shape sequence was semi-structured. The twelve shapes were organized into four temporal base triplets. When one shape appeared, it did so in the context of its base triplet. However, the order of the triplets was semi-random “so that no repetitions of base triplets or triplet pairs (e.g. Triplet₁, Triplet₃, Triplet₁, Triplet₃) occurred and the number of base triplets was identical in each third of the movie” (Fiser
The joint probability of any base triplet’s appearance was .083 (Fiser & Aslin, 2002a).

The training session in Fiser & Aslin (2002a) and the one in this experiment do differ in one respect. The movie used in this study was 12 minutes instead of 6 minutes. The doubled duration, used in Fiser & Aslin Experiment 2A (2002a), was used to facilitate learning. Having hypothesized a deficit in the older adults, exposure was increased to assure that any deficit was not due to insufficient training.

**Procedure**

Participants first completed an informed consent approved by the Georgetown Institutional Review Board. They then completed a biographical and a Health Screening Questionnaire (Christensen, Moye, Armson & Kern, 1992). As well, several neuropsychological tests (see Table 1) were administered prior to the main experimental task.

In the main task, participants viewed the 12 minute movie created by semi-randomly concatenating base triplets (Fiser & Aslin, 2002a). The training directions, based on Fiser and Aslin’s instructions (2002a), urged participants to pay attention to the movie. However, the instructions did not indicate what would later be tested or what the purpose of the experiment was.

Two tests, referring to the training session, were administered. The first test was a replication of Fiser & Aslin’s (2002a) Experiment 1. Each trial presented two animated sequences of three elements, separated by a one second pause. On all thirty-two trials, participants were asked to choose the more familiar sequence based on the training movie. One of the sequences was a base triplet (p = 0.83) and the other an impossible triplet (p =
Impossible sequences were constructed in such a manner that they could never have appeared in the movie. Participants were given unlimited time to answer and urged to guess if unsure.

Participants then completed a second test based on the training session. The test, created in Microsoft PowerPoint, consisted of 32 trials – each presented on a different slide. Each slide consisted of two stationary sequences of three shapes. The two sequences were presented vertically, with one to the left (labeled A) and one to the right (labeled B). Figure 2 shows a sample slide. Participants were asked to choose which sequence occurred more often in the movie, Sequence A or Sequence B. They were urged to guess if they were not sure. Answers were recorded on a separate sheet. Since the test was designed in Microsoft PowerPoint, which allows slides to be viewed out of order, the experimenter was present during the task. The proctor controlled the mouse and keyboard to make sure that participants viewed questions sequentially, without going back to previous questions.

In each trial, one of the two sequences was a base triplet. The other was either an impossible triplet from the animated test or a new impossible triplet. Repeated and novel impossible triplets had a null joint probability. 16 trials presented repeated impossible triplets and 16 trials presented novel impossible triplets. Each impossible-base combination was tested twice in two different orders so that a sequence occurred equally often on the left versus the right of the screen.

Participants then responded to a questionnaire relating to the movie and subsequent tests. Questions probed whether or not individuals had paid attention, what
kind of expectations people had while watching the movie, and whether or not sequences had been noticed.

The experiment concluded with a card sorting task. Participants were given a deck of 12 cards, each depicting one of the twelve shapes presented in the training movie. The bottom of each shape was indicated. Participants were asked to order the twelve cards into 4 sequences of three elements based on the training session.

Results

Recognition Tests:

On the animated recognition test (from Fiser & Aslin, 2002a), both age groups’ mean accuracy is significantly above chance (50 %), as is illustrated by Figure 3. In accordance with Fiser and Aslin’s results (2002a), the young group showed statistically significant learning, in that they were able to distinguish familiar sequences that occurred in the training movie from unfamiliar, novel sequences. In addition, the older group showed statistically significant learning. On the stationary recognition test (an addition to Fiser & Aslin, 2002a), both age groups significantly discriminated familiar from unfamiliar sequences, as seen in Figure 4. T-values are reported in Table 2.

| Table 2. Single Sample T-Tests on the Recognition Test Scores of Experiment 1 |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|                             | Older Adults               |                             | Young Adults                |                             |
|                             | Animated Test              | Stationary Test – Novel Probes | Stationary Test – Repeated Probes | Stationary Test – Cumulative |
|                             | t(10) =                    |                             |                             |                             |
| Animated Test               |                             |                             |                             |                             |
| Stationary Test – Novel Probes | 2.25*                      |                             |                             |                             |
| Stationary Test – Repeated Probes | 2.92*                      |                             |                             |                             |
| Stationary Test – Cumulative | 1.39                       |                             |                             |                             |
|                             |                             |                             |                             |                             |
|                             | Young Adults               |                             |                             |                             |
|                             | Animated Test              | Stationary Test – Novel Probes | Stationary Test – Repeated Probes | Stationary Test – Cumulative |
|                             | t(10) =                    |                             |                             |                             |
| Stationary Test – Novel Probes | 3.01*                      |                             |                             |                             |
| Stationary Test – Repeated Probes | 3.22**                     |                             |                             |                             |
| Stationary Test – Cumulative | 1.79                       |                             |                             |                             |
|                             |                             |                             |                             |                             |
| p < .05                      | *                           |                             |                             |                             |
| p < .01                      | **                          |                             |                             |                             |
For both age groups, while mean accuracy on the stationary recognition test is significantly different from chance, a difference may be found between the scores on questions with novel impossible triplets versus repeated ones. For both age groups, performance is significantly above chance for the novel impossible triplet questions but not for those with repeated imposibles (Table 2). The scatterplot in Figure 5 shows the distribution of scores for each question type and age group. Repeated impossible triplets are probes which appeared in the animated recognition test, while novel imposibles are sequences which are first encountered in the second test. The fact that the animated test familiarizes subjects with repeated probes may explain why participants do poorer on questions with repeated impossible triplets during the stationary recognition. A breakdown by age and question type for mean accuracy on the stationary recognition test can be seen in Figure 6.

The age difference in learning was not significant on either kind of recognition test (Animated Test: t(20) = 1.35, p > .10; Stationary Test: t(20) = .95, p > .35), but it was in the direction of being poorer for the old group as can be seen in Figures 3 and 4. In addition, the correlation between age and accuracy, though negative, was not statistically significant for either test (Animated Test: r(20) = -0.29, p > .10; Stationary Test: r(20) = -0.23, p > .30). Figure 7 gives a scatterplot of the association between age and mean accuracy.

Scores on the two recognition tests did not differ significantly for either age group (Young: t(20) = 0.05, p > .95; Older: t(20) = 0.22, p > .80). Interestingly, the correlation between the two tests, depicted via the scatterplot in Figure 8, was significant for the young adults (r(9) = 0.91, p < .0001), but it was not for the older adults (r(9) = 0.47, p
> .10), indicating that the use of a single measure might not adequately assess how much
the older people have learned.

*Card Sorting Task:*

When assessing participants’ ability to generate triplets on the card sorting task, 
triplet formation was analyzed twice, once taking into account the order of elements
within the sequence and once not taking it into account. When internal order was
considered, if shape1-shape2-shape3 was a correct triplet, it had to be generated as
shape1-shape2-shape3 to be considered correct. However, when internal order was not
considered, the triplet shape3-shape1-shape2 would also have been correct.

While the young subjects were able to generate sequences of shapes above chance
on the card sorting task (Internal order considered: \( t(10) = 3.61, p < .01 \); Internal order
not considered: \( t(10) = 3.49, p < .01 \)), the older adults were not (Internal order considered:
\( t(10) = 0.87, p = .41 \); Internal order not considered: \( t(10) = 1.04, p = .32 \)), an age
difference that was statistically significant. Younger adults produced significantly more
triplets than older adults, whether the order of elements within the triplet was taken into
account (\( t(20) = 3.39, p < .01 \)) or not (\( t(20) = 2.39, p < .05 \)). Figure 9 illustrates for both
age groups the mean number of correct triplets formed – either taking into account
correct internal sequence order or not. In addition, a negative correlation was found
between age and triplet making when considering order (\( r(20) = -.60, p < .01 \)) and when
not (\( r(20) = -.48, p < .05 \)); this correlation is shown in the scatterplot of Figure 10.
Furthermore, only 2 older versus 7 younger adults were able to form triplets when
element order was not taken into account; 1 older versus 7 younger adults put together
triplets when element order was considered.
All subjects who scored above 80% correct on the two recognition tests were able to form at least 1 triplet (order not taken into account). Moreover, there was a significant positive correlation between subjects’ maximum recognition test score and triplet making, both when internal sequence order was looked at ($r(20) = 0.79, p < .0001$) and when it was not ($r(20) = 0.86, p < .0001$), as seen in Figures 11 and 12.

In addition to triplet making, pair making was considered for the card sorting task. Pairs could consist of either the first two, last two, or first and last elements of a base triplet. When analyzing pair making, element order within pairs was not considered. As well, the placement of pair elements within a single triplet was not taken into account.

Younger adults produced significantly more pairs than older adults ($t(20) = 2.21, p < .05$). This age difference was not consistent across types of pairs, though. While a significant age difference existed for making pairs of first and last elements of a base triplet ($t(20) = 2.40, p < .05$), only a marginal age difference was found for making first element pairs ($t(20) = 2.07, p = .052$) and last element pairs ($t(20) = 1.57, p = .131$). Figure 13 illustrates for each pair type the mean number of correct pairs formed by young and older adults. Moreover, a significant negative correlation was found between pair making and age ($r(20) = -.45, p < .05$). The correlation varied in significance for the different types of pairs, though. While age was significantly correlated to making pairs of first elements ($r(20) = -.44, p < .05$) and pairs of first and last elements ($r(20) = -.47, p < .05$), the correlation was not reliable for generating pairs of last elements ($r(20) = -.35, p = .11$). Figure 14 illustrates the relationship between pair making and age.

Furthermore, the ability to make pairs was significantly correlated with subject’s maximum recognition test score ($r(20) = .90, p < .0001$). The correlation was significant
beyond the .0001 level for all types of pairs. The relationship between pair making and maximum recognition test score is shown in Figure 15.

**Questionnaire:**

The post-test questionnaire first evaluated participants’ expectations during the training session. Seven of the subjects (2 older and 5 young) reported that they watched out for patterns or sequences of shapes. Of those seven adults, five (1 older and 4 young) performed significantly above chance on both recognition tests, scoring above 90% correct. Of the eight participants who scored above 75% on both tests, three did not predict the later measures would relate to order or sequence. Other testing expectations related to shape recall or recognition, the number of different shapes, frequency of individual shapes, and “how far [shapes] moved left to right.” A representative response of what a subject anticipated is: “Either identifying the order/sequence of appearing shapes, identifying how many patterns or if there was one at all, or being asked to recreate/draw the shapes from memory.”

The questionnaire also asked participants if they had noticed reoccurring patterns in the training movie. Eight older adults and seven young adults indicated having observed reoccurring sequences. However, only five older adults could recall repeating pairs of shapes and none could identify triplets. Meanwhile, six younger subjects recalled at least one reoccurring triplet and four identified at least two triplets. All subjects who scored above 75% on both discrimination tests either recalled a repeating pair or triplet from the movie.

On the post-test questionnaire, participants were asked: “Which kind of test did you find easier?” Subjects’ answers are summarized in Table 3. Most older adults (8 of
11) answered that both recognition tests were of equivalent difficulty. Three answered that the test with stationary sequences was easier. The older subject with the best performance stated: “I noticed 2 sequences which made the stationary test easier.”

Meanwhile, most younger subjects (6 of 11) indicated that the stationary recognition test was easier (Table 3). One repeated explanation was that the stationary test provided subjects with more time to compare sequences and make a decision. Interestingly, one young subject who scored significantly below chance noted that the second test was easier because one could “compare the 1st [test] and 2nd [test] to each other.” Three young adults indicated that the animated recognition test was easier, while two others stated the two assessments were equivalent in difficulty. The animated test was judged easier because “it was more similar to the training video.” One subject stated, “I was familiar w/ the movement, not w/ three stationary pictures.” Another participant summed up well why some thought the two tests of equivalent difficulty: “I had to get used to seeing them stationary and visualize the movement to make sure, but for the first test I had to pay more attention because I couldn’t compare the two choices at once.”

| Table 3. Breakdown of Participants’ Recognition Test Preference in Experiment 1 |
|-----------------------------------------------|--------|--------|
|                                               | Young  | Older  |
| Animated Test                                 | 27.27% | 0%     |
| Stationary Test                               | 54.55% | 27.27% |
| Neither                                       | 18.18% | 72.73% |

The questionnaire also asked: “When answering questions, did you feel that the sequence you did not choose as more familiar had been presented fewer times or just not at all in the training movie?” For both age groups, the modal answer was “fewer times.” “There were some patterns that kept reoccurring, but there seemed to be a lot of randomness in between, so I’m thinking the other patterns did appear at some time.” Only
one young participant replied “never,” when in fact the impossible sequences – the wrong answers – had never occurred in the video. However, two young learners who responded “other” grasped that the incorrect choices had appeared very few times. “Most of the time I felt it had not been presented at all, but a few times I thought it had just been presented once or twice.”

In summary, the questionnaire revealed that: 1) subjects expected to be tested on a variety of aspects of the training movie, including sequences of shapes; 2) younger subjects and participants with high scores on the recognition tests were better able to identify reoccurring pairs and triplets; 3) while most older subjects found the two recognition tests to be of equivalent difficulty, most young subjects preferred the stationary test; and, 4) in both age groups, most people believed that the incorrect choice on the recognition tests was a sequence that had appeared in the training fewer times than the correct choice (instead of never).

Discussion

On the two recognition tests, both young and older adults discriminated familiar from unfamiliar visual sequences based on statistical information. While older participants were in the direction of being poorer, the age difference was not significant. Of note, the correlation between scores on the two recognition tests was significant for young adults, but not older subjects, indicating that the use of a single measure may not adequately gauge older subjects’ learning. When subjects judged which test they found easier, most older adults responded that the two tests were of equivalent difficulty, while most young participants assessed the stationary test as less difficult. Furthermore, the age difference for generating familiar triplets on the card sorting task was significant. Young
adults were able to generate shape sequences but older adults were not. Triplet making was negatively correlated with age and positively correlated with maximum recognition test score. Moreover, even though young participants were significantly better at generating familiar pairs of elements overall, the age difference was not consistent across the different types of pairs. In addition, young subjects and participants with high scores on the recognition tests identified more correct reoccurring pairs and triplets on the post-test questionnaire.

While no age difference in statistical learning was found, the study’s small n should be taken into account. As well, the inability to detect an age difference may be due to the younger participants’ variable and low scores. A t-test between the animated test scores of the current study and those of Fiser and Aslin’s experiment (2002a) reveals that the present study’s younger participants performed significantly worse ($t(17) = 2.20, p < .05$). Two differences in procedure may explain why the current study’s young subjects performed more poorly. First, the training movie was 12 minutes long, double the length used by Fiser and Aslin (2002a). Second, a battery of neuropsychological tests was added to the study. This series of tests, which took approximately 20 minutes to complete, was performed by subjects before they viewed the training movie. The lengthened familiarization and added tasks may have bored young subjects, who were then less attentive to the training. Subjects were questioned about paying attention to the movie on the post-test questionnaire and the one subject who reported not paying attention was excluded from results. However, other participants who reported paying attention were observed whistling during the recognition tests and fidgeting excessively in their chairs.
Research suggests that statistical learning in the visual domain is affected by attention. Baker, Olson, & Behrmann (2004) found that attention to individual shapes is required for statistical learning. Subjects were presented with shapes that were either connected by a bar or not. When participants were required to pay attention to the location of both shapes, statistical learning was not affected by whether or not the two stimuli were connected. Connectedness of shapes did influence learning, however, when subjects were only asked to attend to one of the shapes’ location. Both perceptual grouping and explicit attention allowed individual shapes to be attended to and thus enabled statistical learning.

The role of attention is not limited to the visual domain. Research indicates that word segmentation based on statistical regularities in auditory sequences is impaired when attention is diverted (Toro, Sinnett, & Soto-Faraco, 2005). Word segmentation drops to chance levels when individuals are required to perform a difficult task or an auditory activity while passively listening to the training speech stream. Performance is affected as attention resources are depleted.

Furthermore, “normal aging is associated with episodic memory impairments, and when young adults are made to encode information under [divided attention] conditions, their memory performance is reduced and resembles that of old adults working under full attention conditions” (Anderson et. al., 2000). Notably, Anderson et. al. (2000) have found that left inferior prefrontal activity is reduced similarly by aging and by divided attention during encoding of episodic memory. Thus, if the current study’s young subjects were inattentive during the training movie when statistical information would be encoded, then their reduced performance on the animated recognition test, as compared to
Fiser and Aslin’s subjects (2002a), is to be expected. Moreover, one would expect not to find an age difference on the recognition task if older subjects were working under full attention conditions. Since older adults are more likely to worry about performing poorly on a test they know looks at learning and memory, they are more likely to have paid close attention to the training and not have become distracted.

Since a lack of attention may have affected the results on the recognition tests in Experiment 1, a second study was created to minimize boredom and fatigue. In Experiment 2, two key changes were made to take into account the role of attention. First, the training session was reduced back to the 6 minutes originally used by Fiser and Aslin (2002a). Second, the battery of neuropsychological tests, which takes approximately 20 minutes to complete, was given to subjects last instead of before the training session.

The aim of the second experiment was to determine whether age deficits in learning would appear under these modified conditions. We expected that young adults would improve compared to Experiment 1, as they would be more likely to remain attentive during training and test. However, it seemed possible that older adults would show the reverse pattern, with lower learning scores in Experiment 2 than in Experiment 1, because the amount of exposure to the regularity was cut in half. Overall, we predicted that an age deficit would emerge.

**Experiment 2**

**Method**

*Participants*

Twenty-two new participants were tested: eleven young adults (1 man, 10 women, mean age = 19.6, range 19-21) and eleven older adults (5 men, 6 women, mean age =
72.9, range 65-89) (Table 4). As seen in Table 4, the two groups did not differ in MMSE scores. However, young adults performed significantly better on Vocabulary, Digit Symbol, and Digit Span.

| Table 4.  
Participant Characteristics for Experiment 2 |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Young Adult</td>
</tr>
<tr>
<td>Gender</td>
<td>1 M, 10 F</td>
</tr>
<tr>
<td>Age***</td>
<td>19.64 (0.67)</td>
</tr>
<tr>
<td>Education***</td>
<td>13.64 (0.67)</td>
</tr>
<tr>
<td>Mini Mental Status Examination</td>
<td>29.91 (0.30)</td>
</tr>
<tr>
<td>Vocabulary*</td>
<td>54.73 (6.28)</td>
</tr>
<tr>
<td>Digit Symbol: Coding***</td>
<td>91.82 (8.06)</td>
</tr>
<tr>
<td>Digit Symbol: Pairing**</td>
<td>15.00 (3.55)</td>
</tr>
<tr>
<td>Digit Symbol: Recall*</td>
<td>7.91 (1.14)</td>
</tr>
<tr>
<td>Digit Span Forward*</td>
<td>12.82 (1.40)</td>
</tr>
<tr>
<td>Digit Span Backward*</td>
<td>9.18 (2.64)</td>
</tr>
</tbody>
</table>

One younger and all older participants were compensated $15. All others received extra credit for a Georgetown University psychology course. The older adults were recruited through newspaper advertisement; the younger adults were Georgetown University undergraduates recruited from a Research Methods & Statistics course and through the Georgetown Research Volunteer Program.

Stimuli and Procedure

The stimuli and procedure were the same as in Experiment 1, except for two modifications. First, the familiarization movie was 6 minutes instead of 12 minutes long. As well, the series of neuropsychological tests given to subjects (Table 4) was administered after the card sorting task instead of before the training session. These two changes were made to minimize fatigue and maximize attention during the training movie and subsequent tests.
Results and Discussion

Recognition Tests:

Parallel to the results in Experiment 1 and in Fiser and Aslin’s study (2002a), on the animated recognition test, the young group showed statistically significant learning, in that they were able to distinguish familiar sequences that occurred in the training movie from unfamiliar, novel sequences. However, in contrast with Experiment 1, the older group’s mean accuracy was not significantly above chance (50%), as is illustrated by Figure 16. On the stationary recognition test, young adults significantly discriminated familiar from unfamiliar sequences while older participants did not, as seen in Figure 17. However, older adult’s learning was marginal when referring to their maximum test score. T-values are reported in Table 5.

<table>
<thead>
<tr>
<th></th>
<th>Older Adults</th>
<th>Animated Test</th>
<th>Stationary Test – Novel Probes</th>
<th>Stationary Test – Repeated Probes</th>
<th>Stationary Test – Cumulative</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t(10) =</td>
<td>1.46</td>
<td>2.58*</td>
<td>0.77</td>
<td>1.71</td>
<td>2.15***</td>
</tr>
<tr>
<td>Young Adults</td>
<td>t(10) =</td>
<td>13.20**</td>
<td>31.30**</td>
<td>11.66**</td>
<td>18.56**</td>
<td>18.56**</td>
</tr>
</tbody>
</table>

While the older group’s mean accuracy on the stationary recognition test is not significantly different from chance, a difference may be found between the scores on questions with novel impossible triplets versus repeated ones. Performance is significantly above chance for the novel impossible triplet questions but not for those with repeated impossibles (Table 5). Young subjects’ learning, on the other hand, is consistent across the question types of the stationary test. A breakdown by age and
question type for mean accuracy on the stationary recognition test can be seen in Figure 18.

The age difference in learning is significant on both recognition tests (Animated Test: \( t(20) = 6.41, p < .0001 \); Stationary Test: \( t(20) = 6.62, p < .0001 \)). As can be seen in Figures 16 and 17, the older group performed more poorly than young subjects on the recognition tests. In addition, the correlation between age and accuracy was significant for both tests (Animated Test: \( r(20) = -0.81, p < .0001 \); Stationary Test: \( r(20) = -0.81, p < .0001 \)). Figure 19 gives a scatterplot of the association between age and mean accuracy.

Scores on the two recognition tests did not differ significantly for either age group (Young: \( t(20) = 1.88, p = .07 \); Older: \( t(20) = 0.40, p = .70 \)). Furthermore, the correlation between the two tests was significant for both young adults (\( r(9) = 0.71, p < .05 \)) and older adults (\( r(9) = 0.90, p < .001 \)).

Card Sorting Task:

Mirroring the findings of Experiment 1, while young subjects were able to generate triplets above chance on the card sorting task (Internal order considered: \( t(10) = 4.35, p < .01 \); Internal order not considered: \( t(10) = 7.59, p < .0001 \)), the older adults were not (Internal order considered: \( t(10) = 1.39, p = .19 \); Internal order not considered: \( t(10) = 0.89, p = .39 \)), an age difference that was statistically significant. Young adults produced significantly more triplets than older adults, whether the order of elements within the triplet was taken into account (\( t(20) = 3.82, p < .01 \)) or not (\( t(20) = 6.84, p < .0001 \)). Figure 20 illustrates for both age groups the mean number of correct triplets formed – either taking into account correct internal sequence order or not. In addition, a negative correlation was found between age and triplet making when considering order (\( r(20) = - \)
.64, p < .01) and when not (r(20) = -.82, p < .0001); this correlation is shown in the scatterplot of Figure 21. Furthermore, only 2 older adults versus all 11 younger adults were able to form triplets when element order was not taken into account; 2 older versus 10 younger adults put together triplets when element order was considered.

All subjects who scored above 70% correct on the two recognition tests were able to form at least 1 triplet (order not taken into account). Moreover, there was a significant positive correlation between subjects’ maximum recognition test score and triplet making, both when internal sequence order was looked at (r(20) = 0.67, p < .001) and when it was not (r(20) = 0.83, p < .0001), as seen in Figures 22 and 23.

Furthermore, young adults produced significantly more pairs than older adults (t(20) = 7.33, p < .0001). This age difference was consistent across all types of pairs. Figure 24 illustrates for each pair type the mean number of correct pairs formed by young and older adults. Moreover, a significant negative correlation was found between pair making and age (r(20) = -.84, p < .0001). The correlation with age was significant beyond the .001 level for all types of pairs. The relationship between pair making and age is shown in Figure 25.

The ability to make pairs was also significantly correlated with subject’s maximum recognition test score (r(20) = .88, p < .0001). The correlation was significant beyond the .001 level for all types of pairs. The relationship between pair making and maximum recognition test score is shown in Figure 26.

**Questionnaire:**

The post-test questionnaire first evaluated participants’ expectations during the training session. As in Experiment 1, participants expected to be tested on a variety of
aspects of the movie, including sequences or patterns of shapes, the number of different shapes, and the frequency of individual shapes.

The questionnaire also asked participants if they had noticed reoccurring patterns in the training movie. Eight older adults and all 11 young adults indicated having observed reoccurring sequences. However, only two older adults could recall repeating pairs of shapes and only one could identify a triplet. Meanwhile, six younger subjects recalled at least one reoccurring triplet. All subjects who scored above 75% on both discrimination tests either recalled a repeating pair or triplet from the movie.

Participants were also asked which recognition test they found easier; answers are summarized in Table 6. As in Experiment 1, most older adults (6 of 11) answered that both recognition tests were of equivalent difficulty. Three answered that the test with stationary sequences was easier and two replied that the animated test was less difficult. Meanwhile, similarly to Experiment 1, most younger subjects (7 of 11) indicated that the stationary recognition test was easier. One individual answered that the test with animated sequences was easier and three replied that the two recognition tests were equally difficult. Interestingly, one younger subject remarked: “I found the second test easier because having the first test helped identify some reoccurring sequences. It helped me remember them better.”

| Table 6. Breakdown of Participants’ Recognition Test Preference in Experiment 2 |
|-----------------------------|-----------------------------|
|                            | Young          | Older          |
| Animated Test              | 9.09%          | 18.18%         |
| Stationary Test            | 63.64%         | 27.27%         |
| Neither                    | 27.27%         | 54.55%         |

The questionnaire also asked: “When answering questions, did you feel that the sequence you did not choose as more familiar had been presented fewer times or just not
at all in the training movie?” For both age groups, answers were quite varied, as seen in Table 7. One young subject, who scored perfectly on both recognition tests, had a very interesting explanation for choosing “other” in response to this question. “It was possible [that the incorrect sequences] occurred as part of 2 chunks mixed together. [For example,] if a sequence goes 123 123, the one that wasn’t more familiar could have been 312. I’m not sure if they appeared like this in the movie.” In the current study, incorrect sequences were impossible triplets. This subject, however, noticed that one could additionally test participants with part triplets, as was done by Fiser and Aslin (2002a).

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Answer</th>
<th>Fewer times</th>
<th>Not at all</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older Adults</td>
<td>45.45%</td>
<td>18.18%</td>
<td>36.36%</td>
<td></td>
</tr>
<tr>
<td>Young Adults</td>
<td>27.27%</td>
<td>36.36%</td>
<td>36.36%</td>
<td></td>
</tr>
</tbody>
</table>

To sum up, the questionnaire revealed that as in Experiment 1: 1) subjects expected to be tested on a variety of aspects of the training movie, including sequences of shapes; 2) younger subjects and participants with high scores on the recognition tests were better able to identify reoccurring pairs and triplets; and 3) while most older subjects found the two recognition tests to be of equivalent difficulty, most young subjects preferred the stationary test. However, unlike Experiment 1, participants varied in their assessment as to whether the incorrect choice on the recognition tests was a sequence that had appeared in the training fewer times than the correct choice or never.

Experiment 1 vs. Experiment 2:

Experiment 2 was designed to create optimal attention conditions, conducive to statistical learning. A shortened training session was used (6 vs. 12 minutes) and the main
task and training were placed at the beginning of the procedure. As in Experiment 1, young subjects significantly discriminated and generated familiar sequences. However, the variability of their scores on the recognition tests decreased from Experiment 1. Furthermore, young adults scored significantly better on the stationary recognition test in Experiment 2 than in Experiment 1 ($t(20) = 2.19$, $p < .05$). Scores did not differ between the animated recognition tests of the two experiments ($t(20) = 1.63$, $p = .12$). Thus, young adults’ performance is negatively impacted by distraction and fatigue.

Older adults, however, showed the reverse pattern in that they were in the direction of performing more poorly in Experiment 2 than in Experiment 1. Although they had shown above chance discrimination in Experiment 1, their performance was only marginally above chance in Experiment 2 when looking at the maximum recognition test scores. The difference between older subjects’ test scores on the two experiments was not significant (Animated: $t(20) = 0.83$, $p = .42$; Stationary: $t(20) = 0.60$, $p = .55$).

Furthermore, as in Experiment 1, older adults did not significantly produce familiar sequences on the card sorting task. Thus, older adults do not appear to have been affected by the same attention factors as young adults. As older individuals are more likely to worry about impairment on tests of memory and learning, they are more likely to have focused during the training session. However, older adults’ performance in Experiment 2 may have been limited by the reduced exposure to the training movie. Older subjects may require more training to show significant learning. As well, the small $n$ may explain the lack of significant learning from visual sequences.

As in Experiment 1, a significant age difference emerged for generating familiar sequences of shapes based on the familiarization session. In contrast with Experiment 1,
the two age groups significantly differed in their ability to distinguish familiar from
never-seen sequences. Thus, statistical learning from visual sequences appears to be
impaired with old age.

**General Discussion**

The present experiments, first of all, aimed to replicate the findings of Fiser and
Aslin (2002a). As in Fiser and Aslin’s study (2002a), young adults were able to
discriminate familiar from unfamiliar sequences of shapes based on statistical
information. The current study as well showed that young adults are able to generate
familiar sequences of shapes. The second aim of the present study was to assess the
differences between young and older adults’ statistical learning from visual sequences.
Older adults were impaired at generating familiar shape sequences. Furthermore, older
participants were poorer than their younger counterparts at discriminating familiar from
unfamiliar sequences of shapes. However, when subjects were is conditions of distraction
and inattention, young participants, but not older adults, showed decreased performance.
When subjects were placed in such distraction conditions, the two age groups did not
differ in distinguishing previously-seen from novel sequences. Young adults thus perform
best with minimum fatigue. Furthermore, older adults’ discrimination of familiar
sequences may be enhanced with increased exposure, as mean accuracy on the
recognition tests was above chance with 12 minutes but not 6 minutes of training.

In Experiment 1, for both age groups, questions on the stationary recognition test
with repeated versus novel impossibles differed in significance. While learning was
statistically significant for questions with novel probes, it was not for questions with
repeated impossibles. The same pattern appeared for older subjects in Experiment 2.
However, the same was not true for the young group in Experiment 2. Experiment 2’s young group outperformed all others. Thus, it appears that, unless learning is maximal, a difference exists between the questions on the stationary test with repeated impossible sequences and those with novel probes. The fact that the animated test familiarizes subjects with repeated probes may explain why participants do poorer on questions with repeated impossible triplets during the stationary recognition test.

Furthermore, it is possible that participants’ scores on the stationary recognition test are affected by learning from the animated test. However, the effect can only be slight as scores of the two recognition tests do not significantly differ from each other for either age group and in either experiment. Moreover, learning during the recognition tests may affect participants’ ability to generate sequences on the card sorting task, as well as their capacity to identify reoccurring sequences from the training movie on the post-test questionnaire. Nonetheless, even if subjects do learn not only from the training session but also from the tests, all participants were given the same opportunity to amass knowledge as the procedure’s order was held constant.

Moreover, a few general points on testing older adults should be noted. First, “recruitment methods that place high demands on older people, such as volunteering to come to a university for testing, might over-represent the highest performing older adults” (Hedden & Gabrieli, 2004). In addition, recruiting older individuals from newspaper advertisement may over-represent well-educated older adults. Therefore, the older adults tested in the present study may have performed better than the average older person. While high-performing older adults may have been over-sampled, the nature of the test may have promoted low performance in older participants. Rahhal, Hasher, and
Colcombe (2001) have shown that older subjects show decreased performance on tasks when they know they are being tested on learning and memory. In Rahhal, Hasher, and Colcombe (2001), older and younger adults were given an explicit memory test on newly learned trivia. Age differences appeared when instructions emphasized memory, but disappeared with neutral instructions. In the current study, as subjects were coming into the Cognitive Aging Lab, they knew that some aspect of learning and memory would be tested. Even if testing instructions never used the word memory, it was associated to the tasks simply by virtue of the lab’s purpose. Furthermore, as participants were told they would be tested on the training movie’s content, they understood that they had to remember something about the familiarization.

Viewing past research in light of the current study’s findings, it is evident that statistical learning of visual sequences is not limited to infants (Fiser & Aslin, 2002b; Kirkham, Slemmer, & Johnson, 2002) and young adults (Fiser & Aslin, 2002a), but extends to older individuals. Furthermore, the present experiments support Love’s work (2000). Love (2000) reported that older adults were impaired at discriminating familiar auditory sequences, while the current study found that older adults are poorer than their younger counterparts at discriminating and generating familiar visual sequences.

In addition, the present experiments complement research investigating the implicit and explicit aspects of statistical learning. Statistical learning may not be uniformly implicit or explicit across domains, as modality differences exist for such learning (Conway & Christiansen, 2005). Conway and Christiansen (2005) reported that statistical learning of tactile sequences appears to be implicit. Vibration sequences, produced from a finite-state grammar, were applied to subject’s fingertips. During a
testing session, participants classified novel sequences appropriately based on the statistical structure of the training. However, participants were unable to verbalize the nature of sequence legality. Thus, tactile learning, which occurs in the absence of feedback and apparently without participant’s awareness, may be implicit. Furthermore, Love (2000) suggests that statistical learning from auditory sequences may be implicit as it does not require explicit attention. Moreover, Baker, Olsen and Behrmann (2004) suggest that visual statistical learning may have an explicit component as it is not simply a passive process. In their study, Fiser and Aslin (2002a) do not classify statistical learning from visual sequences as either implicit or explicit, but instead use the term “unsupervised.” The present experiments, however, indicate that it is at least partially an explicit process. Participants who performed the best of the recognition tests were those who generated the most sequences on the card sorting task and identified the most reoccurring pairs and triplets on the questionnaire. Furthermore, one subject was able to describe the structure of the training movie and of the possible recognition tests. Thus, it is not clear whether the age deficits in this study reflect implicit processes or not.

Further research will be required to determine the sensitivity of older adult’s statistical learning. In the current study, older adults discriminated between triplets with joint probabilities of .083 and null joint probabilities. However, it remains unknown whether or not they can distinguish between two sequences with nonzero joint probabilities. Thus, future studies should test older subjects with base triplets and part triplets as in Fiser and Aslin (2002a). Moreover, further research will be required to determine which temporal statistics older adults make use of. Notably, research should assess whether or not older observers are sensitive to conditional probabilities.
In summary, both older and younger adults can extract higher-order temporal statistics from a continuous stream of shapes. However, older adults are poorer than young participants at both discriminating and generating familiar sequences. Furthermore, distraction and fatigue in the experimental situation can greatly influence younger individual’s performance on tasks involving statistical learning.
References


Acknowledgements

This research was supported by the National Institute of Health Grant R37AG15450. I am grateful to Dr. Darlene Howard and Ilana Bennett for their guidance, mentoring, and reviewal throughout this research, Dr. Richard Aslin and Dr. József Fiser for providing the training stimuli and animated recognition test used in these experiments, Dr. Hans Engler for aiding with calculations of statistics and probabilities relating to the card sorting task, and Dr. Chandan Vaidya and Dr. Fathali Moghaddam for giving feedback on the research as directors of the honor’s program.
Figure 1. Consecutive screenshots of the training movie, illustrating a base triplet.
Figure 2. Sample question from the stationary recognition test. A is a repeated impossible sequence while B is a base triplet. The correct answer is B.
Figure 3. Mean accuracy on animated recognition test (with standard error bars) for young and older adults in Experiment 1.
Figure 4. Mean accuracy on stationary recognition test (with standard error bars) for young and older adults in Experiment 1.
Figure 5. Scatterplot of the association between age and accuracy for questions with novel and repeated impossible sequences on the stationary recognition test of Experiment 1.
Figure 6. Mean accuracy on Experiment 1’s stationary recognition test (with standard error bars) for young and older adults by question type – with novel or repeated impossible triplets.
Figure 7. Scatterplot of the association between accuracy on the recognition tests and age in Experiment 1.
Figure 8. Scatterplot of the association between the stationary and the animated test scores of Experiment 1.
Figure 9. Mean number of correct triplets formed – either taking into account correct internal sequence order or not – on the card sorting task of Experiment 1 for young and older adults (with standard error bars). Chance for correct triplets with correct internal order is 0.012 triplets; chance for correct triplets with any internal order is 0.073 triplets.
Figure 10. Scatterplot of the association between age and number of triplets formed – when taking into account correct internal sequence order and when not – on the card sorting task of Experiment 1.
**Figure 11.** Scatterplot of the association in Experiment 1 between maximum score on the two recognition tests and number of triplets formed on the card sorting task when not taking into account correct internal sequence order.
Figure 12. Scatterplot of the association in Experiment 1 between maximum score on the two recognition tests and number of triplets formed on the card sorting task when taking into account correct internal sequence order.
Figure 13. Mean number of correct pairs formed – for pairs consisting of the first two, last two, or first and last elements of a base triplet – on the card sorting task of Experiment 1 for young and older adults (with standard error bars).
Figure 14. Line chart depicting the number of correct pairs formed for each pair type by subject age in Experiment 1. For each pair type (first element, last element, and first and last element), the minimum number of pairs which can be formed is 0 and the maximum is 4. The minimum number of total pairs which can be formed is 0 and the maximum is 12.
Figure 15. Line chart depicting the number of correct pairs formed for each pair type by the maximum recognition test score in Experiment 1. For each pair type (first element, last element, and first and last element), the minimum number of pairs which can be formed is 0 and the maximum is 4. The minimum number of total pairs which can be formed is 0 and the maximum is 12.
**Figure 16.** Mean accuracy on animated recognition test (with standard error bars) for young and older adults in Experiment 2.
Figure 17. Mean accuracy on stationary recognition test (with standard error bars) for young and older adults in Experiment 2.
Figure 18. Mean accuracy on Experiment 2’s stationary recognition test (with standard error bars) for young and older adults by question type – with novel or repeated impossible triplets.
Figure 19. Scatterplot of the association between accuracy on the recognition tests and age in Experiment 2.
Figure 20. Mean number of correct triplets formed – either taking into account correct internal sequence order or not – on the card sorting task of Experiment 2 for young and older adults (with standard error bars). Chance for correct triplets with correct internal order is 0.012 triplets; chance for correct triplets with any internal order is 0.073 triplets.
Figure 21. Scatterplot of the association between age and number of triplets formed – when taking into account correct internal sequence order and when not – on the card sorting task of Experiment 2.
Figure 22. Scatterplot of the association in Experiment 2 between maximum score on the two recognition tests and number of triplets formed on the card sorting task when not taking into account correct internal sequence order.
Figure 23. Scatterplot of the association in Experiment 2 between maximum score on the two recognition tests and number of triplets formed on the card sorting task when taking into account correct internal sequence order.
Figure 24. Mean number of correct pairs formed – for pairs consisting of the first two, last two, or first and last elements of a base triplet – on the card sorting task of Experiment 2 for young and older adults (with standard error bars).
Figure 25. Line chart depicting the number of correct pairs formed for each pair type by subject age in Experiment 2. For each pair type (first element, last element, and first and last element), the minimum number of pairs which can be formed is 0 and the maximum is 4. The minimum number of total pairs which can be formed is 0 and the maximum is 12.
Figure 26. Line chart depicting the number of correct pairs formed for each pair type by the maximum recognition test score in Experiment 2. For each pair type (first element, last element, and first and last element), the minimum number of pairs which can be formed is 0 and the maximum is 4. The minimum number of total pairs which can be formed is 0 and the maximum is 12.