Small Sets of Novel Words Are Fully Retained After 1-Week in Typically Developing Children and Down Syndrome: A Fast Mapping Study

Stella Sakhon, Kelly Edwards, Alison Luongo, Melanie Murphy, AND Jamie Edgin
Department of Psychology, University of Arizona, Tucson, Arizona
(RECEIVED November 26, 2017; FINAL REVISION May 15, 2018; ACCEPTED May 21, 2018)

Abstract

Objectives: Down syndrome (DS) is a population with known hippocampal impairment, with studies showing that individuals with DS display difficulties in spatial navigation and remembering arbitrary bindings. Recent research has also demonstrated the importance of the hippocampus for novel word-learning. Based on these data, we aimed to determine whether individuals with DS show deficits in learning new labels and if they may benefit from encoding conditions thought to be less reliant on hippocampal function (i.e., through fast mapping). Methods: In the current study, we examined immediate, 5-min, and 1-week delayed word-learning across two learning conditions (e.g., explicit encoding vs. fast mapping). These conditions were examined across groups (twenty-six 3- to 5-year-old typically developing children and twenty-six 11- to 28-year-old individuals with DS with comparable verbal and nonverbal scores on the Kaufman Brief Intelligence Test – second edition) and in reference to sleep quality. Results: Both individuals with and without DS showed retention after a 1-week delay, and the current study found no benefit of the fast mapping condition in either group contrary to our expectations. Eye tracking data showed that preferential eye movements to target words were not present immediately but emerged after 1-week in both groups. Furthermore, sleep measures collected via actigraphy did not relate to retention in either group. Conclusions: This study presents novel data on long-term knowledge retention in reference to sleep patterns in DS and adds to a body of knowledge helping us to understand the processes of word-learning in typical and atypically developing populations. (JINS, 2018, 24, 955–965)

Keywords: Down syndrome, Development, Memory, Sleep, Fast mapping, Word-learning

INTRODUCTION

Memory is a key cognitive process, influencing our ability to acquire new information, including new words. The hippocampus is a brain structure critical to explicit memory; its function is to bind together unrelated pieces of information into an integrated memory (Davachi & Wagner, 2002; Eichenbaum & Bunsey, 1995; Giovanello, Schnyer, & Verfaellie, 2009; Yonelinas, Hopfinger, Buonocore, Kroll, & Baynes, 2001). Memory for such arbitrary associations is important when learning new words. For instance, to use "Dax" as the referent for a novel object, the learner must bind that label to the object. Despite substantial evidence for the link between the hippocampus and word-learning (Duff, Hengst, Tranel, & Cohen, 2006; Mayes et al., 2004), few studies have examined the word-learning process in children with memory disorders.

In the current study, we examine children’s ability to learn new words across a 1-week delay in two encoding conditions, one of which is considered to be less reliant on hippocampal function. Using behavioral and eye tracking methods, we compare word-learning in children with and without Down syndrome (DS), a population with known hippocampal impairment and deficits in memory (Edgin, 2013; Jarrold, Baddeley, & Phillips, 2007; Nadel, 2003).

DS is a chromosomal condition that occurs when an individual has an extra copy of chromosome 21 (Lejune, Gautier, & Turpin, 1959). Individuals with DS display memory and learning difficulties, and the syndrome’s profile of hippocampal deficits is well-established in humans and animal models (Clark, Fernandez, Sakhon, Spanò, & Edgin, 2017; Menghini, Costanzo, & Vicari, 2011; Nadel, 2003). Individuals with DS also have delays in myelination, decreased density of myelinated axons of the dentate gyrus, and reduced volume of the
hippocampus (Ábrahám et al., 2012; Menghini et al., 2011; Wisniewski, 1990). Genetic mouse models of DS and human testing using neuropsychological tasks have revealed deficits on hippocampal-dependent tasks, such as pattern separation and spatial navigation (Clark et al., 2017; Courbois et al., 2013; Lavenex et al., 2015; Pennington, Moon, Edgin, Stedron, & Nadel, 2003; Smith, Kesner, & Korenberg, 2014).

In addition to memory deficits, individuals with DS have prominent language delays. Specifically, studies have shown poor vocabulary growth, grammar deficits, and deficits in speech-sound production (Mervis & Robinson, 2000; Singer Harris, Bellugi, Bates, Jones, & Rossen, 1997; Yoder, Woynaroski, Fey, & Warren, 2014). Despite these data, not all researchers have found verbal long-term memory deficits. Jarrold et al. (2007) found that verbal information was retained better than spatial information across an estimated 8-min delay, findings that are consistent with word-learning studies showing that children with DS can effectively comprehend and produce one single novel word across an hour delay (Chapman, Bird, & Schwartz, 1990). Other studies have suggested that nonverbal memory retention was also unimpaired at 24 hr (Roberts & Richmond, 2015), but follow-up studies suggested that young children with DS have deficits in temporal order memory for deferred imitation sequences across a delay of 1-month (Milojevich & Lukowski, 2016).

Given the importance of examining long-term word retention in individuals with DS and the literature suggesting that word-learning may be linked to the hippocampus under certain encoding conditions, we examined the ability of individuals with and without DS to retain words across a 1-week delay. We used measures of explicit memory recognition as well as implicit eye tracking measures to measure preferential looking to target words. The relational memory exhibited via eye tracking measures has been proposed to engage the hippocampal system; therefore, eye tracking can provide additional information about learning mechanisms (Hannula & Ranganath, 2009; Richmond & Nelson, 2009).

One theoretical perspective that has been important for understanding the processes involved when acquiring arbitrary associations (e.g., word-learning) is the complementary learning systems model of memory (McClelland, McNaughton, & O’Reilly, 1995). This model proposes that sparse representations are readily indexed by the hippocampus and medial temporal lobe systems. Under this model, distributed representations are more slowly formed by the neocortex (Davis & Gaskell, 2009; McClelland et al., 1995, but also see McClelland, 2013). The complementary learning systems model proposes that information initially supported by the hippocampus and neocortex is consolidated over time and becomes less dependent on the hippocampus. Much work has been conducted regarding the role of sleep in this process (Diekelmann & Born, 2010). However, a few studies investigating a learning mechanism called fast mapping (FM), have provided some evidence suggesting that the rapid acquisition of novel arbitrary associations can be achieved independently of the hippocampus when words are learned under certain conditions (Sharon, Moscovitch, & Gilboa, 2011, but also see Greve, Cooper, & Henson, 2014; Smith, Urgolites, Hopkins, & Squire, 2014; Warren & Duff, 2014; Warren, Tranel, & Duff, 2016).

FM is an incidental, exclusion-based learning procedure commonly used to explain how young children can rapidly acquire language after being exposed to a word once or a few times. In a typical FM task, participants are presented with a familiar item, a novel item, and the name of the novel item for each trial. Participants can infer that the label is referencing the novel item and not the familiar item because they already know the name of the familiar item. The FM procedure has three key characteristics, stipulating that the novel word: (1) is encoded incidentally, (2) is introduced in context with an already-known item, and (3) has a meaning that is apparent through inference (Coutanche & Thompson-Schill, 2014). In a matched condition using explicit encoding (EE), which is known to be hippocampal-dependent and to benefit typical adults compared to FM (Atir-Sharon, Gilboa, Hazan, Koilis, & Manevitz, 2015; Coutanche & Thompson-Schill, 2014; Merhav, Karni, & Gilboa, 2015; Sharon et al., 2011), participants are presented with just one novel item and its name and are instructed to remember the novel association.

In Sharon et al. (2011), four of six patients with hippocampal damage were able to form novel arbitrary associations better through FM compared to EE, suggesting an alternate cortical route for consolidation. Initial findings from this study suggested that amnesic patients are able to rapidly acquire arbitrary novel associations using FM, independent of the hippocampus. Merhav et al. (2015) further investigated whether semantic associations acquired through FM can be integrated directly into cortical regions. Participants’ functional brain responses were measured during the four-alternative forced-choice recognition test of the associations they acquired. Compared to the EE condition, the FM condition had significantly increased activity in the anterior temporal lobe (ATL). This finding, in concert with a similar study (Atir-Sharon et al., 2015), suggests that an ATL-related network is responsible for associations learned through FM.

In total, these studies provide a potential alternative learning strategy for novel arbitrary associations mediated by the ATL. As such, the current study seeks to examine word-learning based on hippocampal and hippocampal-independent ways of encoding and retaining words in individuals with DS and mental-age matched typically developing (TD) children to determine if long-term retention might be enhanced in DS with FM.

Poor sleep also negatively influences word-learning in children with and without DS (Gómez & Edgin, 2015). Individuals with DS have sleep disruptions, including obstructive sleep apnea, and previous work has shown that these sleep disruptions correlate with poorer language outcomes (Breslin et al., 2014; Edgin et al., 2015). However, little work has examined sleep quality in reference to the word-learning process over delays of days to weeks. Given the profile of sleep disturbance in this sample, we also examined whether or not sleep may impact long-term
retention in either group (Breslin et al., 2014; Churchhill, Kickehefer, Landis, & Ward, 2012; Edgin et al., 2015).

Therefore, the current study aimed to determine whether individuals with DS show deficits in word-learning across delays in two encoding conditions (FM and EE) and in reference to sleep. The central hypothesis, based on data suggesting that the hippocampus is compromised in DS, is that individuals with DS will learn novel arbitrary associations better through FM than EE and that their EE performance will be impaired in relation to controls. We expected that individuals with DS would have a delayed or absent effect of preferential looking to target items during retrieval compared to the TD group for the EE condition. Finally, we expected that sleep quality would be poorer in individuals with DS. Based on previous work, better sleep should relate to better word retention in both groups, and specifically in the EE condition, which relies on hippocampal representations and replay of those bindings for cortical consolidation.

METHODS

Participants

Twenty-six 11- to 28-year-old individuals with DS (M = 18.70 years; SD = 4.80; 15 male) and twenty-six 3- to 5-year-old TD controls were recruited (M = 4.52 years; SD = .71, 13 male). DS was verified by karyotype report. The DS and TD group achieved similar raw verbal [t(50) = .01; p = .92] and nonverbal [t(50) = −.82; p = .42] scores on the Kaufman Brief Intelligence Test – second edition (K-BIT-II). The DS group had a mean verbal raw score of 26.73 (range = 1–54; SD = 13.13), a mean verbal standardized score of 46.27 (range = 40–71; SD = 9.45), a mean nonverbal raw score of 14.27 (range = 0–22; SD = 4.87), and a mean nonverbal standardized score of 48.19 (range = 40–67; SD = 8.42). The TD group had a mean verbal raw score of 26.42 (range = 11–47; SD = 8.94), a mean verbal standardized score of 106.19 (range = 73–130; SD = 15.28), a mean nonverbal raw score of 15.69 (range = 1–34; SD = 7.41) and a mean nonverbal standardized score of 106.96 (range = 58–145; SD = 20.93).

On another benchmark 12-item list-learning task from a NIH-funded memory battery (the Arizona Memory Assessment for Preschoolers and Special Populations, A-MAP, Clark et al., 2017), we found that the DS sample showed impaired list learning retention both at immediate (t(41) = −2.37; p = .02) and long-term delays (>10 min, t(41) = −2.10; p = .04), a finding matching the memory profile reported in the previous literature (Pennington et al., 2003).

All groups were recruited through local and parent organizations and advertisement. The exclusion criteria for this study included: past head injury or brain trauma, incident of loss of consciousness, accidental poisoning, chemotherapy or radiation, enrollment in a clinical trial, and uncorrected vision or hearing impairments. An additional exclusion criterion for the DS group was an autism diagnosis. All experimental procedures were approved by the University of Arizona Institutional Review Board.

Sample size of 26 individuals per group was estimated for a power of .80 and alpha equal to .05 to detect large effect sizes for between-samples condition differences (Cohen, 1992). Our sample size was commensurate with previous investigations of memory in Down syndrome and exceeds the samples comparing FM and EE in previous work (Sharon et al., 2011).

Materials

Actigraphy

The Actiwacht-2 (Phillips Respironics Mini-Mitter, Bend, OR) was used to obtain sleep data. According to standard collection procedures for pediatric groups, participants wore the watch for 7 days (Sadeh, 2011). Data for four participants, from each group, were not collected due to technical errors or refusal. Parents also completed a 1-week sleep diary to verify the watch’s recordings. Data were collected in 30-s epochs and analyzed using commercially available software (Respironics Actiware 5.71.0, Bend, OR). Actigraphy data were scored at the medium sensitivity threshold (activity counts = 40/min), with sleep onset and sleep end marked by a period of 3 and 5 min of immobility or more, respectively (Meltzer, Montgomery-Downs, Insana, & Walsh, 2012).

Each epoch of data was assessed as sleep or wake, based on whether the activity score exceeded the medium threshold. The actigraphy variables of interest were: average sleep efficiency (percent of time spent asleep from sleep onset to offset), average sleep time (time spent asleep minus any periods of wake), average wake after sleep onset (time spent awake), average wake percentage (percent of time spent awake from sleep onset to offset), and average sleep fragmentation (an index of restlessness based on the sum of mobile time and immobile time that lasts less than a minute during the night). Averages were taken across all nights.

Word-learning stimuli

Two lists of words (versions A and B), were designed in a computer-based paradigm, each containing four novel and two familiar words. All of the six words were in the categories of “fruit” or “animals,” and each list had two novel words from each category (Figure 1). The familiar items were at the preschool level (e.g., based on the MacArthur-Bates CDI, Fenson, Dale, & Reznick, 1993). Most stimuli were

![](https://www.cambridge.org/core/image/?captive=0&doi=10.1017/S1355617718000450&file_id=5135561778000450)
provided by Asaf Gilboa to match those used in the original study of hippocampal patients. Additional familiar items were taken from the internet. Pilot work using the same paradigm as the current study but using a set-size of eight items with a small group (n = 13 DS) showed that participants could not correctly recognize the target item significantly above chance across conditions: FM (M = 48.08%; p = .80) and EE (M = 59.62%; p = .16).

Therefore, we decreased the set-size to four items per list to maximize the levels of encoding performance as to be able to obtain an assessment of retention across long-term delays. Questions regarding the perceptual detail of items, like the ones used in the Sharon et al. (2011) study (e.g., “is the numbat’s tail pointed up?”), were too language demanding for individuals with DS based on some piloting conducted on a small set of individuals to determine if they could map labels onto new referents using this method. Therefore, participants were asked to touch the target item instead of answering a perceptually based question similar to a paradigm used by Spiegel and Halberda (2011).

Procedure

To ensure that participants encoded the items incidentally in the FM condition, every participant underwent the FM condition first, followed by the EE condition, which began on the third session. Each session was separated by a week, equating to four sessions per participant (Figure 2). All participants learned both lists A and B, one using the FM encoding procedure and the other using the EE procedure. The two lists were counterbalanced between the two encoding conditions. The neuropsychological and IQ assessments were administered between session 2 and 3.

Fast Mapping Condition

Practice encoding phase

In session 1, all participants began with the same five non-randomized practice trials (three familiar and two novel targets) to ensure participants understood the task. Participants were told that they would be playing a pointing game (Figure 3). Participants used a pointing stick to make their selection. Each trial began with a fixation cross and then participants heard the target name while being shown the target item and a comparison item from the same category, for 5000 ms. After the exposure, participants were instructed to point to the target item (“Now point to the target”) within 5000 ms. The selection-screen timed out after 5000 ms to ensure that every participant was exposed to the stimuli for

Fig. 3. Fast mapping, explicit encoding, and recognition phases.

Fig. 2. Study design. Fast mapping condition: sessions 1 and 2. Explicit encoding condition: sessions 3 and 4. The category test consisted both lists from the FM and EE condition.
the same amount of time. A feedback screen was provided after their selection for 2000 ms (Figure 3). Having separate exposure and stimuli selection displays allowed for eye tracking collection without contamination from a motor response (a procedure we used uniformly across conditions).

**Encoding**

After the practice encoding trials, participants were presented with the FM encoding phase (four novel and two familiar targets). This is the learning phase in which participants infer and select the name of the target item they heard. This followed the same presentation as described for the practice trials. They received two randomized blocks of 6 trials, for a total of 12 trials.

**Recognition test**

Immediately after the FM encoding phase, participants were presented with three of the four target items and asked to point to the object that matched the auditorily presented label (Figure 3). We counterbalanced the location of the target item and tested across three blocks to allow variability in the data with 12 responses instead of three (three blocks of four trials). Participants were tested twice more: (1) at a 5-min delay, and (2) after a 1-week delay, without feedback. Each block was randomized for each participant, and the blocks were randomized for each of the three recognition tests. Participants watched a short cartoon video during the 5-min delay.

**Familiarity test**

After the 1-week delayed recognition test, participants were tested on their familiarity for FM items versus distractors (novel items). During this test, participants were asked “Did you see this in the game we just played with the pointing stick?” This task was used to assess whether participants showed memory of the target items from the FM task.

**Explicit Condition**

**Encoding phase**

After another week, participants returned for the third session to participate in the EE procedure. This time participants were told that they would be playing a “remembering game,” for which they would have to remember the names of the items (Figure 3). They were presented with just one item on each trial and were asked to point to the object that matched the auditorily presented label. This procedure was conducted to equate the two learning conditions as much as possible, aside from having a familiar item presented or not and whether the task was incidental or not. Since there was only one option, participants did not receive feedback.

**Recognition test**

Like the FM condition participants were tested: immediately, after a 5-min delay, and after a 1-week delay, without feedback. The number of target items, recognition test design, and total test trial numbers were equated in the FM and EE tasks.

**Familiarity test**

After the 1-week delayed recognition test, participants were tested on their familiarity for EE items versus distractors (novel items).

**Category test**

At the end of session 4, participants received a category test where they were asked to sort pictures of the novel target items learned in both conditions into three baskets (animal, fruit, and other). This was to assess whether participants categorized each item as the intended category.

**Eye Tracking**

A SensoMotoric Instruments (SMI) eye tracking system was used to track the participant’s eye gaze throughout the FM and EE retrieval trials of the experiment. The SMI eye tracking system rested at the bottom of a desktop monitor (510 mm width, 250 mm height) at an angle of 20. The tracker monitored the participant’s eye position at 60 Hz.

Calibration was performed before beginning the experiment using SMI iView RED-m software. Participants were seated in front of the monitor, in a desk chair for the older participants or a car seat for the younger participants. The experimenters adjusted the position of the monitor until the eye-tracker could accurately detect the eye position. The distance between participant’s head and the monitor was approximately 50–70 cm. During the calibration, participants were asked to follow a dot, which moved to five different locations on the screen.

After the participants finished the FM and EE encoding tasks, respectively, they were immediately tested on their ability to recognize the four novel target items. Eye tracking was collected throughout the study but analyzed at the test recognition phases across all delays. The fixation cross preceded every trial, focusing the attention of the participant and ensuring that the eye tracker would pick up the eye movements starting from the beginning of each trial. The first display of each trial, included the onset of the three–item display while the audio played the novel word, lasted 5000 ms. The second display of each trial, when the participant was prompted to point to the target item, lasted until their response.

Three areas of interest (AOI) (top item, left item, right item) were applied to each trial assigned as Target, Distractor 1, or Distractor 2. Using SMI BeGaze software, net dwell time of eye gaze on each AOI and the background white space was computed. Focusing on only the AOs (disregarding the white space net dwell time), the proportions of net dwell time for Target, Distractor 1, and Distractor 2 were calculated by summing the net dwell times of only the AOs given from BeGaze and dividing by the total looking time of all three.
The data were divided into four groups: average looking time of target for correct trials, average looking time of target for incorrect trials, average looking time per distractor for correct trials, and average looking time per distractor for incorrect trials. We focus our analyses of the time course of these effects on correct trials only.

**RESULTS**

**Behavioral Results**

Results were considered significant for behavioral tasks at \( p < .01 \) to lower the threshold for multiple comparisons (based on five tests central to our hypotheses and Bonferroni correction). Post hoc tests, additional exploratory analyses, and tests of above chance performance were evaluated at \( p < .001 \). Both the TD (\( M = 91.35; SD = 16.09 \)) and DS (\( M = 86.88; SD = 17.99 \)) groups performed significantly above chance (\( p < .001 \)) and not significantly different from each other (\( t(50) = .94; p = .35 \)) during the FM encoding task, suggesting they had enough word knowledge for the well-known category exemplars to accurately map the label to the novel object. The groups did not perform significantly different from each other on the familiarity test (\( t(50) = .36; p = .35 \)).

A \( 2 \times 2 \times 3 \) repeated measures analysis of variance (Group \( \times \) Condition \( \times \) Delay) was conducted to compare the effect of delays on the number of words learned in each condition between the two groups (Figure 4, Table 1). There were no main group [\( F(1,50) = 0.06; p = .80 \)], delay [\( F(2,49) = 1.32; p = .28 \)], or condition [\( F(1,50) = 0.84; p = .36 \)] effects. There was a similar pattern for the two groups across conditions; no Condition \( \times \) Group effect was found [\( F(1,50) = 0.26; p = .61 \)], against our original prediction. There also was not a Condition \( \times \) Delay \[ F(2,49) = 1.29; p = .29 \] interaction. However, there was a non-significant trend toward a Group \( \times \) Delay interaction [\( F(2,49) = 4.39; p = .02 \)]. This interaction was more evident between the 5-min and 1-week delay and reflected a tendency for the controls to improve from 5-min to 1-week while the DS group showed maintenance. Therefore, we computed gain/loss scores between the last two delays for each group by condition. In the EE condition, the DS group had a mean loss of 8.58%, while the TD group had a mean gain of 2.63%. These gains and losses for the DS and TD groups were significantly different [\( t(50) = -2.79; p = .007; d = .78 \)]. In the FM condition, the DS group had a mean loss of 1.35%, while the TD group had a mean gain of 3.65%. These changes were not significantly different between groups [\( t(50) = 1.04; p = .31 \)] (Figure 5).

Performance for all delays and conditions for the two groups were significantly above chance (\( p < .001 \)). Contrary to our hypothesis that individuals with DS would learn novel arbitrary associations better through FM than EE, we found no significant differences at immediate test [\( t(25) = -1.44; p = .16 \)], at the 5-min delay [\( t(25) = -0.82; p = .42 \)], or at the 1-week delay [\( t(25) = -0.55; p = .59 \)] (Figure 4). Similarly, we did not find significant differences in the TD group between the two conditions at immediate test [\( t(25) = -0.84; p = .41 \)], the 5-min delay [\( t(25) = 0.53; p = .60 \)], or at the 1-week delay [\( t(25) = -0.36; p = .72 \)].

**Sleep Results**

\( T \) tests for group differences in sleep were conducted on the actigraphy data with a critical \( p \)-value set at 0.01 (0.05/5 = 0.01). Contrary to our hypothesis, individuals with DS only showed a significant difference in average sleep time [\( t(42) = -2.66; p = .01; d = 0.80 \)] compared to TD children. There were no significant differences in sleep efficiency scores [\( t(42) = -2.02; p = .05 \)], average wake after sleep

---

**Table 1. Mean percentage and standard deviation of recognition performance across groups, conditions, and delays**

<table>
<thead>
<tr>
<th>Delay</th>
<th>DS group</th>
<th></th>
<th></th>
<th></th>
<th>TD group</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td></td>
<td></td>
<td>EE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate</td>
<td>26</td>
<td>56.65</td>
<td>24.11</td>
<td>26</td>
<td>56.65</td>
<td>21.26</td>
<td></td>
</tr>
<tr>
<td>5-minute</td>
<td>26</td>
<td>58.62</td>
<td>26.07</td>
<td>26</td>
<td>60.69</td>
<td>25.99</td>
<td></td>
</tr>
<tr>
<td>1-week</td>
<td>26</td>
<td>57.27</td>
<td>27.27</td>
<td>26</td>
<td>64.35</td>
<td>23.51</td>
<td></td>
</tr>
</tbody>
</table>

---

![Fig. 4. Recognition performance.](image-url)

![Fig. 5. Explicit and fast mapping encoding gains/loss between the 5-min test and 1-week test.](image-url)
onset \( t(42) = 0.81; \ p = .42 \), average wake percentage \( t(42) = 1.87; \ p = .068 \), or sleep fragmentation \( t(42) = 1.37; \ p = .18 \) (Table 2). In an exploratory analysis, sleep efficiency for the DS group was not correlated with change in scores in either the FM \( t(22) = .01; \ p = .98 \) or EE \( t(22) = -.18; \ p = .41 \) condition. This was also the case for the TD group in both the FM \( t(22) = .18; \ p = .42 \) and EE \( t(22) = .23; \ p = .29 \) condition. While not reported in full, similar results were found for other sleep variables.

**Eye Tracking Results**

To determine the course of preferential looking to target words, we examined the proportion of looking time across eight 250 ms time bins totaling to the first 2-s (2000 ms) of the trial. The viewing scores to the target were compared to chance performance of 0.33. At immediate test, there were no group or condition differences \( F(1,40) = 12.19, \ p < .001 \) for both and both groups did not view the target item significantly above chance in any bin. Multiple \( t \) tests were conducted on the eye movement data with a critical \( p \) value set at 0.006 (0.05/8 = 0.006 to control based on the number of bins tested; Figure 6).

One-sample \( t \) tests showed that the looking preference to the target item during the FM condition at the 5-min delay occurred at 1000–1250 ms for DS \( t(23) = 3.15; \ p = .004 \). Similarly, at the 5-min delay for the EE condition, preferential looking occurred at 1000–1250 ms for DS only \( t(22) = 3.42; \ p = .002 \). There were no significant overall condition or group differences across the 2-s window at 5-min (Group \( F(1,50) = 0.11; \ p = .74 \); Condition \( F(1,50) = 2.57; \ p = .12 \); Figure 7).

For FM at 1-week, a looking preference emerged at 1000–1250 ms in TD \( t(25) = 3.74; \ p = .001 \), which stayed above chance for the following 750 ms. Greater than chance-level viewing started at 1250–1500 ms for DS \( t(23) = 3.40; \ p = .002 \), continuing across 500 ms for FM. Similar results were found for EE. At 1-week, preferential looking emerged at 1250–1500 ms for TD \( t(25) = 2.97; \ p = .006 \) continuing 500 ms, and at 1500–1750 ms for DS \( t(22) = 4.75; \ p < .001 \) lasting 250 ms. A repeated-measures analysis of variance showed a main effect of Delay \( F(1,40) = 12.19, \ p = .001 \) between the 5-min and 1-week. No group or condition effects were found (Group \( F(1,50) = 0.02; \ p = .90 \); Condition \( F(1,50) = 0.65; \ p = .42 \). The delay main effect shows that both groups looked longer at the target items during the 1-week delay than at 5-min for both conditions across both groups (Figure 8).

**DISCUSSION**

Overall, our hypotheses regarding better integration via encoding with FM in DS and sleep effects on retention were not supported. Therefore, we first detail the main finding of interest in the study, the word-learning retention in both

---

**Table 2. Sleep measures for the DS and TD groups**

<table>
<thead>
<tr>
<th></th>
<th>DS group</th>
<th></th>
<th>TD group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
<td>( M )</td>
<td>( SD )</td>
<td>( n )</td>
</tr>
<tr>
<td>Sleep Efficiency</td>
<td>22</td>
<td>82.82</td>
<td>10.16</td>
<td>22</td>
</tr>
<tr>
<td>Sleep Time</td>
<td>22</td>
<td>422.96</td>
<td>95.41</td>
<td>22</td>
</tr>
<tr>
<td>Wake After Sleep Onset</td>
<td>22</td>
<td>67.08</td>
<td>35.62</td>
<td>22</td>
</tr>
<tr>
<td>Wake Percentage</td>
<td>22</td>
<td>13.93</td>
<td>6.86</td>
<td>22</td>
</tr>
<tr>
<td>Sleep Fragmentation</td>
<td>22</td>
<td>28.14</td>
<td>9.03</td>
<td>22</td>
</tr>
</tbody>
</table>

**Average sleep efficiency (percent of time spent asleep from sleep onset to offset), average sleep time (time spent asleep minus any periods of wake), average wake after sleep onset (time spent awake), average wake percentage (percent of time spent awake from sleep onset to offset), and average sleep fragmentation (an index of restlessness based on the sum of mobile time and immobile time that last less than a minute during the night). Averages were taken across all nights of sleep collected.**

---

---

**Fig. 6. Proportion looking time to the target word across FM and EE conditions and groups at immediate test.**
groups. Specifically, we measured long-term memory retention of small sets of novel words in both groups after just two exposures. Surprisingly, we found that the DS group maintained what they had learned over a long-term delay, whereas typically developing controls showed a small increase in performance from the 5-min delay to 1-week. Eye tracking data showed that preferential viewing for both groups emerged at 1-week, when these effects were not present at the immediate or 5-min delay. This memory related eye movement seems to suggest a consolidation process across one week, and one that operates similarly in typical children and DS across both encoding conditions. These study results are consistent with findings regarding long-term retention for 1 novel item in children with DS in Chapman et al. (1990).

However, Bird, Chapman, and Schwartz (2004) and Ashworth, Hill, Karmiloff-Smith, and Dimitriou (2015) found that individuals with DS had more difficulty with learning 8 to 10 novel words after a delay. A larger set-size appears to be more difficult for individuals with DS both at short and long-term retention intervals. This finding was also supported by our pilot study; DS participants performed at below chance level during an immediate recognition test when remembering 8 novel items. Although decreasing the set-size in the current study improved initial encoding performance, the task of remembering four items may have been too simple to engage the hippocampal memory system for the EE condition. Additionally, Yonelinas (2013) argued that simple associations, potentially similar to the current study’s task demands, are less affected by hippocampal damage. More work is needed to determine if this pattern of results is due to working memory difficulties at encoding and possibly the effects of sleep on different set-sizes (as we discuss in depth below).

The FM literature with young TD children also shows mixed results as to whether young children can retain FM words over a long-term delay. Some studies have shown rapid forgetting for fast mapped words (Horst & Samuelson, 2008; Vlach & Sandhofer, 2012), while others have supported long-term retention (Brady & Goodman, 2014; Carey & Bartlett, 1978; Markson & Bloom 1997; Waxman & Booth, 2000). One argument raised by Vlach and Sandhofer (2012) is that previous studies have varied in how they have incorporated memory supports (i.e., saliency, repetition, and generation). These supports include: presenting the label of objects in a way that make them more salient, labeling the object repeatedly, or requiring learners to verbally generate the label. Vlach and Sandhofer (2012) examined word-learning and retention in children without providing any memory support over a 1-week and 1-month delay and failed to find retention. Alternatively, Waxman and Booth (2000) showed 1-week retention for fast mapped words, with a saliency memory support.

In the current study, we found that both groups showed retention over a week delay. A potential reason for this retention could be attributed to testing effects (Roediger & Karpicke, 2006); testing can enhance later retention even
when feedback is not provided. In the current study, participants were tested after a 5-min delay, which could have contributed to retention at 1-week. Other studies have also failed to show consolidation deficits in DS when participants retrieved with multiple repetitions before a delay (Roberts & Richmond, 2015). More work is needed to examine how individuals with DS may retain information based on the number of repetitions and retrievals. In total, this study suggests two educational strategies that may be beneficial for supporting retention in DS, including (1) training on small item sets, and (2) repeated testing.

This study also investigated whether FM would benefit individuals with DS, a population with hippocampal dysfunction. Based on previous work, we expected that individuals with DS would learn novel arbitrary associations better through FM than EE. Inconsistent with findings from Sharon et al. (2011) and contrary to our hypothesis, our results showed no significant benefit in the FM condition for the DS group. This result is consistent with other recent work failing to measure a FM benefit in healthy older adults, memory-impaired patients (damage to the hippocampus), patients with amnesia (damage to the hippocampus), and patients with left temporal lobectomies (Greve et al., 2014; Smith et al., 2014; Warren & Duff, 2014; Warren et al., 2016).

Therefore, given these recent findings, there is mixed evidence regarding the benefit of FM for the retention of arbitrary labels in individuals that do not have amnesia. One explanation could be that patients with more severe amnesia may have developed alternative methods (i.e., fast mapping) for forming associations as a way to compensate for poor hippocampal function. Alternatively, individuals without amnesia may have some residual hippocampal function they can rely on, and the FM mechanism might not be operating in the same way (i.e., as may be the case in DS given these findings).

We further investigated whether sleep contributed to recognition performance given that sleep is known to play a role in memory consolidation (Diekelmann & Born, 2010). We expected that differences in sleep efficiency between the DS and TD groups would correlate with patterns of long-term retention. The DS and TD group showed a significant difference in average sleep time; however, a range of sleep parameters that relate to better long-term retention in groups with developmental disorders and DS. Given the difficulty executing studies including long-term memory delays in these populations, these studies are rarely conducted, and the large majority of studies of memory formation in developmental disorders focus on delays within short-term testing sessions. However, the examination of the factors influencing long-term retention may reveal educationally relevant, and potentially surprising, results.

**ACKNOWLEDGMENTS**

This research was supported by the LuMind Research Down Syndrome Foundation, the Southern Arizona Network for Down Syndrome, a National Institutes of Health R01 HD088409-01 (to JOE) and the Graduate and Professional Student Council at The University of Arizona. The authors have no conflicts of interest to report. We also thank our colleagues in the Down Syndrome Research Group: Dr. Caron Clark, Yating Liu, Casandra Nyhuis, Carlos Figueroa, Bianca Demara, Payal Anand, Alexandra Totillo, and Katharine Hughes. Our gratitude also goes to the wonderful families that dedicated their time to this study.
REFERENCES


Brady, K.W., & Goodman, J.C. (2014). The type, but not the amount, of information available influences toddlers’ fast mapping and retention of new words. American Journal of Speech-Language Pathology, 23(2), 120–133.


Long-term word retention in Down syndrome


