Award Number: W81XWH-10-1-0404

TITLE: Receptive Vocabulary Knowledge in Low-Functioning Autism as Assessed by Eye Movements, Pupillary Dilation, and Event-Related Potentials

PRINCIPAL INVESTIGATOR: Barry Gordon, M.D., Ph.D.

CONTRACTING ORGANIZATION: Johns Hopkins University
Baltimore, MD 21218

REPORT DATE: June 2012

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
### Abstract

We have been testing the hypothesis that relatively implicit measures of cognitive processing (eye movements, pupillary dilation monitoring, and the N400 component of event-related potentials) will prove sensitive to receptive vocabulary knowledge, even in the absence of more traditional behavioral responses. We have sought to first demonstrate the use of these measures in three populations in whom behavioral responses are expected to be reliable: normal adults, normally developing children, and higher-functioning individuals with autism. In all three groups, the implicit measures differentiated known from unknown words: eye movements were faster to a named picture for known words; pupillary dilation from baseline was greater in the unknown condition; and an N400 congruency effect was observed for known (but not unknown) words. Our results also suggest that these measures similarly differentiate known from unknown words in lower-functioning individuals with autism, even in the absence of a behavioral response. These results suggest that these measures may be used as valid measures of comprehension, even in nonverbal, non-responding individuals.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>Body</td>
<td>5</td>
</tr>
<tr>
<td>Key Research Accomplishments</td>
<td>18</td>
</tr>
<tr>
<td>Reportable Outcomes</td>
<td>19</td>
</tr>
<tr>
<td>Conclusion</td>
<td>20</td>
</tr>
<tr>
<td>References</td>
<td>21</td>
</tr>
</tbody>
</table>
INTRODUCTION

Approximately 50% of individuals affected by autism fail to develop useful speech, and many of these individuals never learn to communicate in any functional way. An important scientific as well as practical question about such individuals, as well as in those with other diagnoses and a similar inability to express themselves, is whether this lack of expressive ability is necessarily accompanied by an equally severe deficit in knowledge of receptive language. Little rigorous research has been directed at this possibility, both because of the difficulty of working with such low-functioning subjects, and because of the lack of sensitivity of most traditional behavioral methodologies. Recently, however, several experimental methodologies have been developed and refined to the point that they may prove sensitive enough to provide reliable evidence of comprehension, even in the absence of more traditional behavioral responses such as speech and gesturing, and even at the individual subject level. We have been developing the use of three such research methods to attempt to detect receptive vocabulary knowledge – eye movement recording, pupillary dilation monitoring, and event-related brain potentials. We have been testing whether these relatively implicit measures of comprehension actually do reflect single-word comprehension in participants in whom we expect reliable behavioral responses to serve as comparison measures (normal adults, normally developing children, and high-functioning individuals with autism), as well as in low-functioning, nonverbal individuals with autism, for whom overt behavioral responses might be unreliable or even impossible.
BODY

Experiment 1: Validating the use of the eye movement monitoring (EM), pupillary dilation monitoring (PD), and event-related potential (ERP) techniques for the measurement of receptive vocabulary knowledge – normal adults

Our first experiment was designed to validate the use of the three implicit methodologies for the detection of receptive vocabulary knowledge in normal adults, a participant population in whom overt behavioral responses would be expected to be reliable (and thus capable of serving as a measure of comparison). Participants were asked to engage in two separate tasks using the same set of 160 words and pictures. Eighty of the words were very high frequency and were expected to be very familiar to all of the adults; these included words such as airplane and camera. The remaining 80 words were low frequency, relatively unfamiliar words that were not expected to be known by many of the participants (as confirmed by prior pre-testing). Examples of words in this set included agouti and cainito. All words were concrete and highly imageable. High-resolution, color digital pictures were selected to represent each word. In the forced-choice recognition task, participants were asked to use the mouse to select one of four pictures presented simultaneously on a computer screen after hearing one of the objects named. We simultaneously collected eye movement and pupillary dilation data, using an ASL Model 504 eye-tracking system. In the congruity task, a picture was presented on the computer screen, accompanied by the auditory presentation of a single word, which either matched (congruous condition) or did not match (incongruous condition) the pictured item. Participants were asked to push a button to indicate whether the auditory word and the picture matched. Simultaneously, ERPs were recorded using Electrical Geodesics Inc.’s 256-channel Hydrocel Geodesic Sensor Nets. Finally, normal adult participants were asked to participate in a word familiarity posttest, in which they were asked to rate their familiarity with the 160 words used in the experiment, on a scale from 1 (very unfamiliar) to 9 (extremely familiar), with an additional option of 0 (no familiarity whatsoever).

At the end of our first year, we had run 55 participants. In our second year, we have focused most of our efforts on refining our data analytic techniques for this group of individuals, under the belief that perfecting these techniques with this group will make data analysis with our other participant populations easier. The development of a metric by which we can pool the known and unknown words on an individual basis for each participant has been a crucial step in this process (and has been encouraged by our peers in response to presentation of our results at conferences). Most of our previous analyses had been done using a priori classifications of whether a word was “known” or “unknown;” that is, we grouped the items in our analyses based on our prior determinations (and pre-testing, as described previously). However, especially for the group of normal adult participants, it is possible that some participants may have been familiar with some of the words that we had estimated to be generally “unknown.” In other words, it would not be surprising if some adults had idiosyncratic knowledge of the less familiar words. To the extent that this was the case, the effects that we had observed would have been muted (due to the inclusion of what are truly known items in our analyses of “unknown” words). It was therefore to our benefit to try to base the classification of “known” and “unknown” items on what truly was known or unknown to each individual participant. To this end, we developed a way of incorporating the various behavioral responses that we have collected from each participant into a metric that allows an estimate of the likelihood that each item was known to
that participant. These behavioral responses include accuracy on the forced choice; accuracy on the congruity task; and the participant’s rating on the familiarity post-test. Based on these calculations, we then re-categorized the experimental items for each participant into new groups of “known” and “unknown” items, and re-analyzed the data accordingly.

As a result of this re-analysis, three of the twenty participants that we had previously deemed “good” (that had enough useable trials in all conditions for both ERPs and EM) were excluded from further analysis (all now had too few useable trials in at least one condition, primarily due to having known too many of the “unknown” words). We were thus forced to recruit new normal adult participants for additional testing. We tested seven new normal adults in order to find three with enough useable “good” data to include in our final count of twenty participants. (Reasons for data exclusion have included participant attrition; difficulty with eye-tracker calibration, which usually results from excessive participant movement; and/or loss of ERP data due to difficulty lowering scalp impedances, excessive motion artifact, or excessive eye movements/blink.)

Behaviorally, participants were faster and more accurate at both the forced-choice task and the congruity task for known words than for unknown words. Our results for the normal adults were also as predicted for all three of the implicit measures (See Figure 1). Specifically, eye movements to the picture that matched the auditory word were faster for known than for unknown words. End-of-trial fixations were also on the named picture more frequently for known words. Pupillary dilation from baseline was greater in the unknown condition (evidencing the greater engagement of cognitive resources when the word is unknown). Additionally, an N400 congruency effect was observed in the event-related potentials for known words, but not for unknown words. Thus, all three implicit measures (EM, PD, and ERPs) were able to distinguish the processing of known from unknown words in this participant population.

We are revising our initial draft of the manuscript of these results; we expect to submit the manuscript for publication by the end of July 2012.
Figure 1. EM, PD, and ERP results from the group of 20 normal adult participants. Top panel: Sample EM results (from one participant) to known (left) and unknown (right) words. Size of the dot indicates the length of the fixation. Middle panel: Average PD results for known and unknown trials. Bottom panel: Grand-average ERPs for congruent vs. incongruent trials for known (left) and unknown (right) words.
Experiment 2: Validating the use of the EM, PD, and ERP techniques for the measurement of receptive vocabulary knowledge – normally developing children

Our second experiment was designed to validate the use of the three implicit methodologies for the detection of receptive vocabulary knowledge in normally developing children (ages 5 – 17), another participant population in whom overt behavioral responses would be expected to be reliable (and thus capable of serving as a measure of comparison). The child participants were tested on the Peabody Picture Vocabulary Test (PPVT; [1]), the Kaufman Brief Intelligence Test (KBIT; [4]) and the Autism Spectrum Screening Questionnaire (ASSQ; [2]), the latter of which was used to ensure that none of the normally developing children exhibited excessive behaviors associated with autism. All of the children were asked to complete the forced-choice recognition task and the congruity task described above. Older children (those old enough to understand and properly perform the task; generally those ages 10 and above) were also asked to complete the familiarity post-test described above.

At the end of our first year, we had run 68 child participants, and we thought (based on our analytic techniques at the time) that we had succeeded in capturing “useable” data from 20 of those. The further refinement of our data analytic techniques (described above for normal adults in Experiment 1), however, led to a re-analysis of our child data that resulted in the exclusion of data from seven of these children. We have been recruiting to replace these participants; in the past few months, we have tested an additional seven child participants, out of whom we have been able to retain one complete set of useable data, for a total of 14 children to date.

Recruitment of the children has proven difficult. The recruitment avenues that have been successful for us in the past yield fewer participants as time goes on. We are exploring other ways of contacting the families of children who might be interested in participating. We also have experienced difficulty with the data of many of the children that we have recruited: between the eye tracker and ERP paradigms, it seems that we lose many of the trials for an individual child, to that point that we cannot keep the data of those that we run. Anecdotally, our research assistants have noted that children seem to be good at one or the other paradigm; we often see that for a given child, their eye tracking (EM and PD) data is good but not their ERP data, or vice versa. We have strived to fill our participant quota with children from whom we have collected good data on both measures. However, we are considering running analyses over a set of children with good ERP data, and separately over a set of children with good eye tracking data, because we would certainly have enough data on each measure (but just not from the same children on both). This remains an option, as our conclusions do not necessarily depend on having data on both measures from the same children (although we do think our conclusions would be strengthened if we could do so).

Our analyses on the 14 children continue to demonstrate that the results for the child participants are very similar to those observed for the normal adults (see Figure 2). Behaviorally, children were faster and more accurate at both the forced-choice task and the congruity task for known words than for unknown words. Eye movements to the picture that matched the auditory word were faster for known than for unknown words. End-of-trial fixations were also on the named picture more frequently for known words. Pupillary dilation from baseline was greater in the unknown condition. Additionally, an N400 congruency effect was observed in the event-related potentials for known words, but not for unknown words. Thus, all three implicit measures (EM, PD, and ERPs) were able to distinguish the processing of known from unknown words for the normally developing children that we have tested.
Figure 2. EM, PD, and ERP results from the group of 14 normal child participants. Top panel: Sample EM results (from one participant) to known (left) and unknown (right) words. Size of the dot indicates the length of the fixation. Middle panel: Average PD results for known and unknown trials. Bottom panel: Grand-average ERPs for congruent vs. incongruent trials for known (left) and unknown (right) words.
Experiment 3: Validating the use of the EM, PD, and ERP techniques for the measurement of receptive vocabulary knowledge – high-functioning individuals with autism

Our third experiment was designed to validate the use of the three implicit methodologies for the detection of receptive vocabulary knowledge in high-functioning individuals with autism, another participant population in whom overt behavioral responses would be expected to be reliable (and thus capable of serving as a measure of comparison), but which also offers a more closely-matched comparison group to the low-functioning individuals with autism. Participants were administered the Kaufman Brief Intelligence Test (KBIT; [4]), the Autism Diagnostic Observation Schedule (ADOS; [5]), and the Autism Diagnostic Interview – Revised (ADI-R; [6]) to confirm diagnosis and to determine level of functioning/verbal ability.

An additional 12 participants with autism were recruited for participation. Of these, 7 were determined to be higher-functioning. These participants were asked to complete the forced-choice recognition task and the congruity task described above. Of the 7 higher-functioning individuals with autism recruited, 4 participants supplied a complete set of “useable” data from the three implicit measures, in addition to the 4 participants that had completed testing at the time of last year’s report, for a total of 8 high-functioning individuals. Reasons for data exclusion were similar to those described above for the normal adults, but again were more pronounced for the autism participants.

We have been much more successful in the past year at recruiting participants with autism that we had been in our first year. We now have a steady enrollment of autism participants (we currently have another 10 participants at various stages in the enrollment or pre-testing stages of the experiment). We have also been much more successful in terms of participant retention; we have learned ways to encourage participants to remain interested and excited about the project over the various testing sessions. Conducting the screening evaluations at participants’ homes (when desired by the participants or their families) has been a huge help in easing the burden of the many sessions for the participants.

We try to analyze the data for the individuals with autism as each participant is run, so we are able to report the results for the 8 higher-functioning individuals with autism that have completed testing to date. The results for the high-functioning individuals with autism are very similar to what we have seen for the normal adults and the normally developing children (see Figure 3). The individuals in this group were able to make reliable behavioral responses. Behaviorally, they were faster and more accurate at both the forced-choice task and the congruity task for known words than for unknown words. Eye movements to the picture that matched the auditory word were faster for known than for unknown words. End-of-trial fixations were also on the named picture more frequently for known words. Pupillary dilation from baseline was greater in the unknown condition. Additionally, an N400 congruency effect was observed in the event-related potentials for known words, but not for unknown words. Thus, all three implicit measures (EM, PD, and ERPs) were able to distinguish the processing of known from unknown words for the high-functioning individuals with autism that we have tested.
Figure 3. EM, PD, and ERP results from the group of 8 high-functioning individuals with autism. Top panel: Sample EM results (from one participant) to known (left) and unknown (right) words. Size of the dot indicates the length of the fixation. Middle panel: Average PD results for known and unknown trials. Bottom panel: Grand-average ERPs for congruent vs. incongruent trials for known (left) and unknown (right) words.
Experiment 4: Extending the use of the EM, PD, and ERP techniques for the assessment of receptive vocabulary knowledge to low-functioning individuals with autism

Our fourth experiment was designed to extend the use of the three implicit methodologies for the detection of receptive vocabulary knowledge to a population in whom behavioral responses are generally less reliable (or absent altogether) – low-functioning, low- or non-verbal individuals with autism. Participants were administered the Kaufman Brief Intelligence Test (KBIT; [4]), the Autism Diagnostic Observation Schedule (ADOS; [5]), and the Autism Diagnostic Interview – Revised (ADI-R; [6]) to confirm diagnosis and to determine level of functioning/verbal ability. Of the 24 participants recruited, seven are currently scheduled for initial screening, or are currently at some stage in the initial screening process.

Of the additional 12 participants with autism recruited this year (see above for Experiment 3), 5 were determined to be low-functioning. All were minimally verbal to nonverbal. Stimuli for the low-functioning group were drawn from the larger pool of 160 words and pictures, but were individualized for each participant based on parental/caregiver report of items that were expected to be known receptively by the child. Parents/caregivers were asked to complete the MacArthur-Bates Communicative Development Inventory – Words and Gestures ([3]), plus a similar experiment-specific inventory that covered those words from our set of 160 that were not included on the MacArthur-Bates. These measures thus provided information about what words were likely to be known (and unknown) receptively by the child. The number of stimuli tested were determined for each individual to maximize signal-to-noise ratio while minimizing experiment length. For some individuals, for whom the pool of known words was small, repetition of items within a testing session, or the repeated testing across multiple testing sessions, was necessary to adequately assess their receptive knowledge.

In addition to the assessments provided by the ADOS and the ADI-R, each low-functioning participant received a series of behavioral assessments designed to evaluate his/her ability to successfully participate in our language testing. We assessed potential participants on things such as their ability to sit still for extended periods of time; their ability to look at the computer screen; their ability to tolerate the eye tracking and ERP equipment; and their likelihood to exhibit adverse behaviors (such as hitting, biting, or other aggressive behaviors). Based on these assessments, an individual determination was made as to the appropriateness of participation and the need for further individualized training to acclimate the participant to the eye-tracking and ERP equipment and experiment procedures. Such training was then conducted as needed over a period of days or weeks, sometimes in our testing space, and often at the participant’s home.

After training, participants completed the same forced-choice and congruity tasks as described for Experiments 1-3. However, they were not required to make any overt behavioral response (using the mouse or pressing a button). (Some low-functioning individuals with autism are very familiar with computer programs of the type used in our experiments, and would like to engage in some kind of task during the experiment. For these participants, we allow them to make responses as they wish. However, importantly, the successful analysis of the implicit measure data in this experiment does not depend upon the behavioral completion of these tasks.)

Of the 5 lower-functioning individuals with autism recruited, 3 participants supplied a complete set of “useable” data from the three implicit measures. Reasons for data exclusion were similar to those described above for the normal adults. Additionally, even with acclimation training, low-functioning participants have a much harder time engaging in the tasks for
extended periods of time, and therefore all of the eye-tracking and ERP artifacts for this group are very pronounced. The behavior of individuals in this population is quite variable, so that on some days, they are unwilling to participate at all. Also, participant attrition is a problem, given the large time commitment required of the participants and their families for successful acclimation training and testing.

We have completely analyzed the data for 1 of the low-functioning individuals with autism tested this year (we are in the process of finalizing the analyses on the other two). When combined with our data from the first year, we are now able to report results from four low-functioning individuals with autism. To date, there is a fair amount of individual variability amongst the results of the low-functioning individuals with autism; for this reason, we have chosen to present the data from each individual separately, instead of averaging across individuals (see Figures 4-7). For most trials for all four participants, eye movements were faster and more accurate for known words than for unknown words. For three of the four participants, changes in pupillary dilation were greater to unknown than to known words. Finally, each of the four LFA participants showed evidence of an N400 congruency effect, with a larger amplitude in the N400 time range in the incongruent condition relative to the congruent condition, but only for the words that were expected (based on parental report) to be known by the individuals.

Figure 4. EM, PD, and ERP results from one low-functioning individual with autism. Top panel: Sample EM results to known (left) and unknown (right) words. Size of the dot indicates the length of the fixation. Middle panel: Average PD results for known and unknown trials. Bottom panel: Average ERPs for congruent vs. incongruent trials for known (left) and unknown (right) words.
Figure 5. EM, PD, and ERP results from a second low-functioning individual with autism. Top panel: Sample EM results to known (left) and unknown (right) words. Size of the dot indicates the length of the fixation. Middle panel: Average PD results for known and unknown trials. Bottom panel: Average ERPs for congruent vs. incongruent trials for known (left) and unknown (right) words.
Figure 6. EM, PD, and ERP results from a third low-functioning individual with autism. Top panel: Sample EM results to known (left) and unknown (right) words. Size of the dot indicates the length of the fixation. Middle panel: Average PD results for known and unknown trials. Bottom panel: Average ERPs for congruent vs. incongruent trials for known (left) and unknown (right) words.
Experiment 5: Using the EM, PD, and ERP techniques to study new word learning in low-functioning individuals with autism

Our fifth experiment was designed to examine changes in EM, PD, and ERP measures in nonverbal, low-functioning individuals with autism that accompany repetitive exposure to new words during a learning period.

In the past year, we have been working with one low-functioning individual with autism as a pilot participant for this phase of the research. (The other individual who we had identified at the end of last year chose not to participate further.) This individual had participated in the fourth experiment successfully, and has proven available and amenable to a long-term training study. We have determined a appropriate exposure and non-exposure sets of words for this individual, and have done extensive behavioral (and additional EM, PD, and ERP testing) to ensure that the words we have chosen were all unknown to the individual at baseline. We have piloted three different training methods with this individual; all three had been developed in collaboration with our speech-language pathologist and our autism specialist to try to maximize word learning over a short period of time. The methods differ in their intensity (number of sessions per day and per week) and in their scope (number of words to be trained), as well as in
the exact method of instruction. We are currently analyzing the data from our pilot sessions with this individual to determine which of the methods seems to result in the largest learning effects.

We have also identified two additional low-functioning individuals with autism who we believe will be suitable candidates for this phase of our research. Both of them participated in the fourth experiment successfully, and are available for a long-term training study. Based on our analysis of their data from Experiment 4, we have recently begun to determine the appropriate exposure and non-exposure sets for these two individuals.

Finally, we have been working to replace our research assistant who will assist with the training phase of the study (visiting participants’ homes for multiple sessions per week), as our former research assistant has left for graduate school.
KEY RESEARCH ACCOMPLISHMENTS

- Finalization of determination of best practice data analytic procedures for normal adult participants
- Re-analysis of all normal adult data using new best practices; recruitment of new participants to replace those eliminated by new data analysis (recruited 7, retained 3)
- Draft of manuscript describing normal adult data
- Completion of child data analysis and re-analysis using new best practices; recruitment of new participants to replace those eliminated by new data analysis (recruited 7, retained 1)
- Recruitment and testing of 7 new high-functioning individuals with autism, of whom 4 provided good data; analysis of all high-functioning data using best practices
- Recruitment, training, and testing of 5 new low-functioning individuals with autism, of whom 3 provided good data; analysis of all low-functioning data using best practices
- Pilot testing of the word training methodologies to be used in our long-term training study
- Presentation of data at The Neurobiology of Language Conference, November 2011; the Annual Meeting of the Cognitive Neuroscience Society, April 2012; and the International Meeting for Autism Research, May 2012
REPORTABLE OUTCOMES


CONCLUSION

In all three populations capable of reliable behavioral responses (normal adults, normally developing children, and high-functioning children with autism), our results indicate that the implicit techniques provide valid measures of receptive knowledge. All three measures (eye movement monitoring, pupillary dilation monitoring, and ERPs) appear capable of differentiating known from unknown words, as compared to the criterion of behavioral responses. Specifically, eye movements were faster to, and eyes fixated longer on, the matching picture in the forced-choice task for known words relative to unknown words; pupillary dilation was greater from baseline for unknown words, relative to known words; and the N400 congruency effect was observed for known words, but not for unknown words. Thus, these data support the validity of these techniques and our experimental design.

Our testing of low-functioning participants with autism also suggests that these implicit techniques provide valid measures of receptive knowledge even in the absence of behavioral responses. All three appear capable of differentiating known from unknown words within this group, in that the results for low-functioning individuals with autism were remarkably similar to those observed for normal adults, normally-developing children, and high-functioning individuals with autism.

The demonstration of receptive abilities in nonverbal individuals would lay a foundation upon which we might better understand their baseline abilities for communication and for comprehension. Knowing that an individual can understand language even when he or she does not speak might support the development of more intensive speech and language therapies, using a broader range of modalities, to capitalize on that individual’s functional preferences or strengths. We believe that our results provide an initial demonstration of such abilities in a group in which such knowledge has traditionally been difficult to assess, and which has been severely under-represented in the studies of cognitive processing and abilities.
REFERENCES


