The Pennsylvania State University

The Graduate School

Department of Communication Sciences and Disorders

INTEGRATION OF TOP-DOWN AND BOTTOM-UP INFORMATION: AN INVESTIGATION OF SCENE AND WORD LIST PROCESSING

A Thesis in

Communication Sciences and Disorders

by

Brandon Alexander Henken

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2019

The thesis of Brandon Alexander Henken was reviewed and approved* by the following:

Diane L. Williams
Professor and Head, Program Director of Communication Sciences and Disorders
Thesis Adviser

Carol A. Miller
Professor of Communication Sciences and Disorders

Michael P. Robb
Professor of Communication Sciences and Disorders

*Signatures are on file in the Graduate School.

ABSTRACT

Utilizing a scene-viewing task and false memory list task, this study investigated how individuals with typical development (TD) integrate top-down semantic information with bottom-up visual information. Participants reported what they saw in household scenes and completed a recall and recognition task for the Deese-Roediger-McDermott (DRM) false memory lists. Responses from the scene-viewing task were analyzed for agreement and divergence between participants. Responses from the DRM lists were scored for veridical and false recall as well as hits, misses, correct rejections of nonlures and critical lures, and false alarms for nonlures and critical lures. Scene-viewing data revealed items that many individuals with TD found salient in each scene but also indicated heterogeneity between participants. On the DRM tasks, individuals with TD falsely recalled and recognized critical lures, indicating activation of conceptual knowledge. These responses will be used as a comparison group in a future study with individuals with autism spectrum disorder (ASD) to elucidate how individuals with ASD process scenes and recall lists of semantically-related words.

TABLE OF CONTENTS

LISTS OF TABLES

ACKNOWLEDGMENTS

For running participants in the lab, thanks are due to the following students in the Cognition and Language Learning Laboratory: Laura Anderson, Alexis Basciano, Jamie Bui, Kathleen Coggins, Breanna Corle, Jordyn Freed, Taylor Freed, Erika Harrison-Ellis El, Stephanie Harvey, Rachel Hemphill, and Nafea Shaikh. I am deeply appreciative for Kirk Wydner’s coding knowledge that he used to help with data analysis. I also appreciate the dedication exhibited by Alexis Basciano and Breanna Corle in coding the data for the scene-viewing task and coming to consensus on responses. Finally, I owe special thanks to my supervisor Dr. Diane L. Williams for her assistance with this project.

Portions of this thesis were presented at a poster session at the 2018 American Speech-Language-Hearing Association Convention. Travel costs were generously provided by the Department of Communication Sciences and Disorders at the Pennsylvania State University. The Department of Communication Sciences and Disorders also provided funding for a graduate assistantship for the investigator that helped make this thesis possible. Additional support for the project was provided through research funds made available to Dr. Williams by the Department of Communication Sciences and Disorders.

Chapter 1

Introduction

1.1 The Complex Information Processing Theory of ASD

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by “persistent deficits in social communication and social interaction” and “restricted, repetitive patterns of behavior, interests, or activities” (American Psychiatric Association, 2013). Another cluster of cognitive features of ASD explored in the literature is the processing of complex information. The Complex Information Processing Theory of ASD (Minshew, Goldstein, Muenz, & Payton, 1992) is a psychological theory of ASD that marks a difference between different types of information processing demands. Complex information processing is conceptualized as processing that requires comprehension, integration, or strategic organization of multiple pieces of information, reflecting the underlying need for coordination of multiple brain regions to accomplish such tasks, whereas simple information processing does not (Minshew, Goldstein, Siegel, 1997; Williams, Minshew, and Goldstein, 2015). Less efficient processing results from this poor coordination between neural systems, as described by the neural underconnectivity theory of ASD (e.g., Just, Cherkassky, Keller, & Minshew, 2004; Just, Cherkassky, Keller, Kana, & Minshew, 2007). Researchers operating under the complex information processing theory of ASD have found that individuals with ASD show intact performance on simple attention, word recall, and paired associates tasks but poorer performance than individuals with typical development (TD) on tasks involving problem-solving, abstraction, and using context to aid in recall of sentences (Minshew & Goldstein, 2001; Minshew et al., 1992).

In the seminal study proposing the Complex Information Processing Theory of ASD, individuals with ASD experienced selective difficulty with tasks that required them to use information-organization strategies such as categorization by semantic category but not on tasks assessing lower-level processing skills such as associative memory or attention (Minshew et al., 1992). Additional research has shown that children and adults with ASD experience greater difficulty with tasks across cognitive domains that require integration and organization of information. Minshew, Goldstein, and Siegel (1997) found that adults and young adults with ASD performed worse than individuals with TD on tasks of complex memory and complex language but not on tasks of attention, simple memory, and simple language. Complex memory tasks were those that required an individual to organize information to facilitate recall, whereas simple memory tasks were those that only involved remembering paired associates or three-word pairs. Complex language tasks were those that required the interpretation of information such as passage and metaphor comprehension, whereas simple language tasks only required word-reading and spelling skills. Thus, the tasks that were difficult for individuals with ASD were those that required implementation of a strategy and integration of material. Williams, Goldstein, and Minshew (2006) replicated this pattern of results and found the same pattern of performance in children with ASD.

Williams, Minshew, and Goldstein (2015) proposed that difficulty with complex information processing is related to poor conceptual knowledge in individuals with ASD. In their study, conceptual knowledge is defined as “the ability to use and organize knowledge into general concepts which can be applied to solve novel problems” (p. 860). In a factor analysis of a test battery measuring conceptual knowledge, language, and attention, the authors found that the complex tasks had high loadings on the conceptual reasoning factor and that loadings on this factor discriminated between the group with ASD and the group with TD. Less accurate conceptual knowledge may be the result of less robust categories (Gastgeb, Rump, Best, Minshew, & Strauss, 2009; Gastgeb, Strauss, & Minshew, 2006; Klinger & Dawson, 2001; Strauss et al., 2012; see Discussion).

The present study was designed to establish a comparison group of individuals with TD for future comparison to individuals with ASD. Two tasks thought to assess integration of material (scene-viewing and word list recall and recognition) were completed. To understand the rationale for the use of these two tasks, the following sections discuss how individuals with TD perform on tasks involving scenes, schemas, and word lists.

1.2 Schemas and scene processing

Scene processing is a task in which individuals must integrate multiple pieces of information. In scene processing, individuals must integrate top-down semantic information (conceptual knowledge) with bottom-up perceptual information (Bar, 2004). One type of cognitive construct in which atypical semantic organization may manifest is the schema. Schemas are "sets of expectations that can facilitate perception" and serve as "prototypical contexts . . . that integrate information about the identity of the objects that are most likely to appear in a specific scene with information about their relationships" (Bar, 2004, p. 618). In other words, a schema is a representation of semantic knowledge regarding typical scenes. While acknowledging that a scene is difficult to define, Henderson and Hollingworth (1999) propose a tentative definition of a scene as "a semantically coherent (and often nameable) view of a real-world environment comprising background elements and multiple discrete objects arranged in a spatially licensed manner" (p. 244). For examples of scenes, the reader is referred to Öhlschläger and Võ (2017) and the database at https://www.scenegrammarlab.com/research/scegram-database/. Because of their weaker conceptual knowledge and difficulty integrating multiple pieces of information, individuals with ASD may have atypical schemas that are less robust than those of individuals with TD as well as atypicalities in scene processing.

1.3 Schemas in visual tasks in individuals with TD

Evidence for the organizing effect of schemas on cognition in individuals with TD is supported by the results of studies of visual processing tasks. For example, participants produced more incorrect responses when choosing a verbal label for visually jumbled scenes (photographs that were divided into sections and rearranged) than for non-jumbled scenes and when recalling an object from a jumbled scene than from a non-jumbled one (Biederman, Rabinowitz, Glass, & Stacy, 1974). These findings suggest that the spatial arrangement of the scene provided a structure that facilitated the access of semantic information by participants. In addition, when the background was contextually congruent with an object, object recognition was facilitated (Boyce, Pollatsek, & Rayner, 1989).

Visual false memory experiments also support schemas as a construct used by individuals with TD in information processing. In the critical condition of Palmer (1975), a prototypical scene was presented followed by a contextually incongruent object that physically resembled a contextually congruent object. For example, a kitchen scene was presented followed by a mailbox (which physically resembles a loaf of bread). In this condition, relative to a condition in which a scene was followed by a contextually congruent object, participants more often incorrectly identified the contextually incongruent but physically similar object (a mailbox) as the object toward which the visual scene biased them (a loaf of bread). These findings suggest that the schema for “kitchen” was activated and then exerted a top-down influence on the visual information and altered participants’ perception of it. Miller and Gazzaniga (1998) found that participants falsely recognized critical lures (e.g., a beach ball) not present in prototypical scenes (e.g., a beach scene). According to schema theory, the context or schema exerted a top-down influence on visual perception, resulting in false recall or false recognition.[1]

1.4 Schemas in language tasks in individuals with TD

Another way to investigate semantic (conceptual) knowledge is to assess associations between concepts in memory. One paradigm for doing so utilizes the Deese-Roediger-McDermott (DRM) lists. DRM lists used in Deese (1959) and Roediger and McDermott (1995) consist of words that are strongly associated with each other according to behavioral free association data from Russell and Jenkins (1954). Missing from each list is a critical lure, a word that Russell and Jenkins (1954) used to prompt free association. For example, a list with *table*, *sit*, *legs*, *seat*, *couch*, *desk*, and so forth is designed to prompt false recall of the critical lure *chair* which is not present in the list. Deese (1959) and Roediger and McDermott (1995) found that presenting such lists to participants elicited false recall of the critical lures. Roediger and McDermott (1995) explain their results in terms of spreading activation. Spreading activation is the associative process by which the processing of one concept activates the processing of another concept (Collins & Loftus, 1975). When a DRM list is presented, each word in the list activates the critical lure in a participant’s semantic network; after viewing the entire list, that critical lure is sufficiently activated or primed to cause the participant to believe the critical lure was present in the list itself (Collins & Loftus, 1975; Roediger & McDermott, 1995).

In the word list task, the activation of the critical lure reflects the structure of conceptual or semantic knowledge. Although false recall may be considered an error in cognition, being

---

[1] Note that schema theory relates to the perceptual schema model of object identification in scenes—there are other theories of object identification that differ in the relationship between top-down and bottom-up processing, as reviewed by Henderson and Hollingworth (1999).

susceptible to false recall reflects an intricate network of related concepts that allows for the application of context to information processing. If such networks were incomplete or impaired, more effortful processing would be required to associate one concept with another, to integrate several concepts at once, or to apply top-down context of information to a processing task.

1.5 Experimental tasks and hypotheses

The overall purpose of the program of research was to determine whether individuals with ASD have less robust semantic knowledge and poorer integration of top-down and bottom-up information. However, before conducting studies with individuals with ASD, the tasks needed to be validated by administration to individuals with TD. The purpose of the current study was to validate two tasks of integrated semantic processing: recall of items from a scene and recall and recognition of words using a false memory word list task.

The first task was a scene-viewing task that was based on the scene-viewing studies described earlier. In this task, participants were shown scenes from rooms in a house from the database developed by Öhlschläger and Võ (2017). After viewing the scenes, they were asked to name five items that they saw. The responses of the participants with TD were considered to be a manifestation of what individuals with robust schemas extract as salient from a scene.

The second task was the Deese-Roediger-McDermott (DRM) false memory list task (Roediger & McDermott, 1995), which consists of the presentation of lists of words that are semantically-related. In this task, each list excludes a critical lure that is highly associated with those items. In a recall and recognition phase, participants were asked to recall the words they saw and to state if they recognized words from list presentation. If schematic knowledge was activated, more false recall and recognition of critical lures is expected.

Chapter 2

Method

2.1 Participants

Thirty participants (18 females) aged 18 to 29 (*M* = 22.43, *SD* = 2.84) participated in both tasks in the present study as part of a larger study on conceptual processing. Participants were recruited through fliers placed around department offices, an announcement on StudyFinder (an online recruitment system used by Penn State University), and a recruitment email. Participants were informed in recruitment materials that compensation would be offered for their participation. Inclusion criterion included being 18 to 35 years of age, having normal or corrected-to-normal vision, and speaking English as their primary language. Exclusionary criteria included lack of sentence-level English, hearing impairment or uncorrected visual impairment, history of significant learning disabilities or other neurological impairment, or a first degree relative with ASD. Testing was conducted over two sessions on two separate days to reduce participant fatigue, as each day of testing took approximately 1 to 1.5 hours.

On the first day of testing, participants completed characterization testing so that they could be more accurately compared to individuals with ASD in a future study. Participants completed a hearing and vision screening, the Social Responsiveness Scale–Second Edition (SRS–2; Constantino & Gruber, 2012), the Test of Nonverbal Intelligence–Fourth Edition (TONI–4; Brown, Sherbenou, & Johnsen, 2010), the Peabody Picture Vocabulary Test–Fourth Edition (PPVT–4; Dunn & Dunn, 2007), and the Expressive One-Word Picture Vocabulary Test–Fourth Edition (EOWPVT–4; Martin & Brownell, 2011). On the second day of testing, participants completed the experimental measures reported below as well as other experimental tasks that were part of a larger study. Demographic information and results of characterization testing for participants appear in Table 2-1.

Note that four participants had Total *T*-scores above 59 on the SRS–2; *T*-scores less than or equal to 59 are considered within normal limits. The scores of these participants were 62, 63, 63, and 64. These scores indicate mild social deficits but were not in the range that is considered consistent with a diagnosis of ASD (that is, a *T*-score of 66 or greater). These participants were included in the study, but their data may be inspected at a later time to determine if they differ from participants with Total *T*-scores less than or equal to 59.

Table 2-1. Demographic information for participants.

Measure	Mean	Standard Deviation	Minimum	Maximum	Range
Age in Years	22.42	2.84	18.83	29.75	10.92
TONI–4 Index Score	104.77	10.67	86	126	40
EOWPVT–4 Standard Score	96.93	12.8	70	124	54
PPVT–4 Standard Score	113.63	23.09	76	216	140
SRS–2 Total *T*-Score	50.17	6.84	38	64	26
SRS–2 Social Awareness *T*-Score	50.23	7.78	32	64	32
SRS–2 Social Cognition *T*-Score	50.47	7.38	42	72	30
SRS–2 Social Communication *T*-Score	47.97	7.07	37	65	28
SRS–2 Social Motivation *T*-Score	51.40	7.87	39	71	32
SRS–2 Restricted Interests and Repetitive Behavior *T*-Score	52.43	8.82	40	78	38
SRS–2 Social Communication and Interaction *T*-Score	49.60	6.72	37	64	27

Note. The TONI–4, EOWPVT–4, and PPVT–4 all have an average standard score of 100 and a standard deviation of 15. The SRS–2 provides a *T*-score that is compared to normative findings to determine an individual's profile of social behavior. A mean *T*-score is 50, and the standard deviation is 10. A *T*-score less than or equal to 59 is considered within normal limits. A *T*-score between 60 and 65 indicates mild social deficits. A *T*-score between 66 and 75 indicates "substantial interference with everyday social interactions" and is consistent with a diagnosis of

ASD. A *T*-score greater than or equal to 76 indicates “severe and enduring interference with everyday social interactions” and is consistent with a diagnosis of ASD (Constantino & Gruber, 2012, test form).

2.2 Design & materials

All experimental tasks were presented on a computer using E-Prime 3.0 software (Psychology Software Tools, 2016). Individual tasks are described in the following subsections.

2.2.1 Materials: Scene-viewing task

Scenes were selected from images available from the Scene Grammar (SCEGRAM) database (Öhlschläger & Võ, 2017) downloaded from https://www.scenegrammarlab.com/research/scegram-database/. This database consists of color photographs of household scenes taken with natural lighting and minimal editing. Photographs were sized by Öhlschläger and Võ to 1024 × 768 pixels and converted to PNG files. These images were selected because they depict common scenes that are likely familiar to individuals aged 18 to 35. Additionally, they do not contain people—deficits in processing social stimuli (American Psychiatric Association, 2013; Davis & Carter, 2014; McPartland, Tillman, Yang, Bernier, & Pelphrey, 2014) could be confounded with deficits in the use of scene context if scenes with people were presented to individuals with ASD in a future study. The images selected for the current experiment did not contain any of the critical objects used by Öhlschläger and Võ (2017). These images were from the condition in which the experimental objects were absent (ABS) from the scene; they were the counterpart to the scenes with a semantically consistent object was present in a control condition (CON). As such, the images used in the present study were labeled ABS_CON in the SCEGRAM database. Scenes 1 through 60 were used in the current experiment and are listed in Appendix A. Two practice trials used two images from Demiral, Malcolm, and Henderson (2012).

2.2.2 Materials: DRM lists task

The DRM lists experiment was based on Roediger and McDermott (1995) and was downloaded from the System for Teaching Experimental Psychology (STEP; MacWhinney, St. James, Schunn, Li, & Schneider, 2001) database at https://step.talkbank.org/scripts/Memory/RoedigerMcDermott1995.html. Minor edits to wording of instructions were made. During the recall phase, 6 word lists (*anger*, *black*, *bread*, *chair*, *cold*, and *doctor*) consisting of 15 words each were presented (see Appendix B). During the recognition phase, the following were presented: 6 lures, 18 words that were shown during the recall phase (3 from each list), and 72 nonlures that were not shown during the recall phase.

2.3 Procedure

2.3.1 Procedure: Scene-viewing task

Participants were provided with instructions presented on the computer prior to the task. In part, the instructions stated “Type what you saw” (rather than “Type something/an item that you saw”) in an effort not to bias participants to report only items in the scene (e.g., *refrigerator*, *stove*) instead of the name (global label) of the scene itself (e.g., *kitchen*). In each trial, a central fixation cross appeared for 1000 ms, and then a scene appeared for 700 ms. Following this presentation, participants typed the names of five items that they saw into five response boxes that appeared sequentially. Sixty scenes were presented. Two practice trials preceded the experimental trials.

2.3.2 Procedure: DRM lists task

Participants were provided with instructions that were presented on the computer prior to the task. Then, 6 lists of 15 semantically-related words (e.g., *table, sit*, *legs*, *seat*, *couch*, *desk,* and so forth) were shown with a critical lure absent (e.g., *chair*). Each word was presented one at a time and appeared for 1500 ms. Following the presentation of each list, a recall phase was completed in which participants were asked to recall as many words as they could by typing into consecutively appearing boxes that appeared on the screen. Finally, in a recognition phase after all the lists were presented, participants were shown words from the lists, critical lures, and nonpresented words. They selected whether they recognized words by typing a yes (1) or no (0) response on a keyboard. Each participant was shown a total of 96 words during the recognition phase: 6 lures, 18 words that were shown during the recall phase (3 from each list), and 72 nonlures that were not shown during the recall phase.

Chapter 3

Results

3.1 Results for the scene-viewing task

E-Prime .dat files containing each participant's data were merged into a common file using E-Merge and then exported to Microsoft Excel (Microsoft Corporation, 2016). In the Excel spreadsheet, responses were sorted by scene, then by participant. Because responses were typed by the participants, a number of idiosyncrasies in the individual responses occurred due to misspellings, typographical errors, differences in naming (e.g., *jacket* vs. *coat*), or the use of descriptors (e.g., *wooden floor*) rather than a single noun (e.g., *floor*). Therefore, the data needed to be regularized so that the responses could be collated across participants. The data were inspected by the investigator and two undergraduate students. Responses were regularized according to the following general rules:

1. Fix obvious misspellings (e.g., in a kitchen scene, *tove* → *stove*).
2. Delete adjectives (e.g., *wooden floor* → *floor*; *red jacket hanging* → *jacket*; *food cooking* → *food*).
    a. During the consensus process, it was decided that some adjectives were to be retained if they were part of an item's usual name or provided distinguishing details (e.g., *colored pencils*; *wire baskets* in a scene with several kinds of containers).
3. Delete articles (e.g., *a table* → *table*).
4. Delete question marks or extraneous symbols (e.g., *plant?* → *plant*; *sink'* → *sink*).
5. Delete responses such as "I don't know"; they were considered equivalent to a non-response.

6. If participants called the same item by different names (e.g., *clothes* and *clothing*, *counter* and *countertop*), responses were collapsed (e.g., *clothes/clothing*, *counter/countertop*). The same rule applied to singular and plural forms of nouns. For example, *shoe* and *shoes* were collapsed to a single response for coding purposes: *shoe/shoes*.
7. Delete repeated items. For example, if the participant wrote *sink* two times, the second instance of it was deleted.

When there was a disagreement on how to regularize a response, agreement was reached through consensus. During the consensus process, a number of differences between coders were found that required resolution. The process of regularizing the written responses and the resolution of coding differences required substantial time and effort. Due to the time constraints imposed by the thesis process (see Limitations of the Current Study), 20 of the original 60 scenes were selected for coding. Therefore, the results described below are preliminary.

After the data were regularized, responses were compiled across participants by scene. The data were coded such that the order in which participants listed their responses was maintained. In this way, the first response each participant made could be coded and analyzed as a measure of what was selected as most salient in the scene. The results were inspected for measures of agreement and heterogeneity between participants. Norms regarding recall of items in the scenes developed by Öhlschläger and Võ (2017) have not been developed, and a literature review did not reveal studies utilizing the present methodology for comparing individuals with TD to individuals with ASD. Therefore, the investigator utilized the following measures as attempts to quantify what participants selected as salient and how much variability existed between participants:

1. The number of responses made by at least 50% of participants for each scene

2. The number of unique responses (i.e., responses made by only one participant) for each scene
3. The number and proportion of participants who named the most frequently occurring response for each scene
4. The number and proportion of participants who named the most frequently occurring *first* response for each scene
5. The number and proportion of scenes for which the most frequently occurring response matched the most frequently occurring *first* response
6. The proportion of all responses for a scene represented by the most frequently occurring response
7. The proportion of all responses for a scene represented by the three most frequently occurring responses

Because duplicate responses by the same participant for the same scene were deleted, the number of times a given response was produced was equivalent to the number of participants who produced it. For example, if 24 participants named *shoes*, it was named 24 times. Table 3-1 contains a sample of how data were collated for Scene 11 (which depicted a shoe rack scene).

Table 3-1. An example of how the data were collated for Scene 11 (shoe rack).

<table>
<tr><th colspan="4">Scene 11</th></tr>
<tr><td colspan="3">1. Number of Responses Made by at Least 50% of Participants</td><td>4</td></tr>
<tr><td colspan="3">2. Number of Unique Responses</td><td>20</td></tr>
<tr><td colspan="3">Most Frequently Occurring Response</td><td>shoes</td></tr>
<tr><td colspan="3">Most Frequently Occurring *First* Response</td><td>shoes</td></tr>
<tr><td colspan="3">3. Number and Proportion of Participants Who Named Most Frequently Occurring Response</td><td>24 (80%)</td></tr>
<tr><td colspan="3">4. Number and Proportion of Participants Who Named Most Frequently Occurring *First* Response</td><td>12 (40%)</td></tr>
<tr><td colspan="3">5. Proportion of All Responses Represented by the Most Frequently Occurring Response</td><td>0.18</td></tr>
<tr><td colspan="3">6. Proportion of All Responses Represented by the Three Most Frequently Occurring Responses</td><td>0.52</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td>Named Item</td><td>Number of Participants Who Named Item</td><td>Proportion of Responses that Consisted of This Item</td><td>Proportion of Participants Who Named Item</td></tr>
<tr><td>shoes</td><td>24</td><td>0.18</td><td>0.80</td></tr>
</table>

rack/shoe rack/shoes rack/stand	23	0.17	0.77
coat/coats/jacket/jackets	23	0.17	0.77
door	16	0.12	0.53
floor/floors/flooring	10	0.078	0.33
boots	5	0.04	0.17
wall/walls	4	0.03	0.13
mudroom	3	0.02	0.10
jeans	2	0.02	0.07
bowl	2	0.02	0.07
shelf/shelves	2	0.02	0.07
rubber bands	1	0.01	0.03
floormat	1	0.01	0.03
hook	1	0.01	0.03
sneakers	1	0.01	0.03
closet	1	0.01	0.03
storage	1	0.01	0.03
sandals	1	0.01	0.03
table	1	0.01	0.03
socks	1	0.01	0.03
white	1	0.01	0.03
black	1	0.01	0.03
clothes	1	0.01	0.03
sweatshirt	1	0.01	0.03
carpet	1	0.01	0.03
towels	1	0.01	0.03
hoodie	1	0.01	0.03
shoe laces	1	0.01	0.03
plant	1	0.01	0.03
washroom	1	0.01	0.03
storage unit	1	0.01	0.03

The number of responses made by at least 50% of participants for each scene was calculated by first dividing the number of participants who made a given response by the total number of participants. If that proportion was greater than or equal to 0.5 for a given response, it was tallied as a response made by at least 50% of participants. The number of responses that met this criterion for each scene was the number of responses that were made by at least 50% of participants. For the scenes coded, the average number of responses made by at least 50% of participants was 2.2 per scene (*SD* = 1.28).

The number of unique responses (i.e., responses made by only one participant) for each scene was calculated by counting how many responses were made only one time. The average number of unique responses across the 30 participants was 16.60 per scene (*SD* = 6.32).

The most frequently occurring response for a scene was identified as the response produced by the highest number of participants. The number of times this response was produced was counted. On average, the number of the 30 participants who reported the most frequently occurring response was 23.2 (*SD* = 5.41). The proportion of participants who produced this response was calculated by dividing the number of participants who produced the most frequently occurring response for a scene by the number of participants (30). On average, 0.77 (*SD* = 0.18) of the participants reported the most frequently occurring item.

Because the data were coded such that the order in which participants listed their responses was maintained, the most frequently occurring *first* response was found by isolating the *first* response made by each participant and then counting to determine which first response occurred most frequently. The average number of same first responses per scene was 11.5 (*SD* = 3.15). For each scene, the number of participants who made the most frequently occurring first response was divided by the total number of participants to determine the proportion of participants who made the same first response for a given scene. An average of 0.38 of the participants (*SD* = 0.11) of participants reported the same first item per scene.

By counting the number of scenes for which the most frequently occurring response matched the most frequently occurring *first* response, it was determined that, in 13 of 20 scenes (65% of the scenes), the most frequently occurring response was the same as the most frequently occurring *first* response.

The proportion of all responses for a scene represented by the most frequently occurring response was calculated by dividing the number of times the most frequently occurring response was produced by the total number of responses for each scene. The most frequently occurring response for each scene represented an average of 0.17 of all responses (*SD* = 0.04).

The proportion of all responses for a scene represented by the three most frequently occurring responses was calculated by dividing the number of times the three most frequently occurring responses were produced by the total number of responses for each scene. The three most frequently occurring responses represented an average of 0.42 of all responses (*SD* = 0.08).

The data for the measures of agreement and heterogeneity of responses discussed above are presented in Appendices C, D, and E.

Next, the scenes were analyzed to determine commonalities between scenes with high and low levels of agreement between participants' responses. Scenes were divided within each measure (the number of responses made by at least 50% of participants, the number of unique responses, the proportions of participants who named the most frequently occurring response and the most frequently occurring *first* response, the proportion of all responses represented by the most frequently occurring response, and the proportion of all responses represented by the three most frequently occurring responses) based on whether their value for the given measure was 1) less than or equal to the average or 2) greater than the average. Informal comparison by the investigator for number of items in the scene and the "nameability" of the scene (e.g., *kitchen* or *bathroom* are highly nameable, but *art table* or *place where you put your keys when you arrive home from work* are less nameable) did not reveal clear patterns that differentially indicated group membership for any measure. The number of items in the scene and nameability warrant further systematic study; see Discussion.

In summary, the data reflected heterogeneity in the participants' responses, but there were signs of a shared means of assigning salience to objects in the scenes. There was an average of 16.6 unique responses for each scene, and only an average of 2.2 objects per scene were reported by at least 50% of participants; these findings suggest that participants made many idiosyncratic responses and were in agreement regarding salience for only a few items per scene. Although there was some agreement in terms of common first responses, even these responses were not usually made by a majority of the participants. Nevertheless, these items tended to be visually prominent in the scenes and often provided information about the scene as a whole (e.g., a common response of *dishwasher* indicated that the scene was a kitchen). This similarity suggested that the participants had a similar method of determining salience for at least some of the scenes. Additionally, the most frequently occurring response in each scene was reported by an average of 77% of participants. Sixty-five percent of scenes had the same value for the most frequently occurring response and the most frequently occurring *first* response, but only an average of 38% of participants reported the same first item—this finding suggests that the most frequently reported response was not always made first, indicating that measures other than the first response were needed to characterize agreement among participants. Moreover, the most frequently occurring response encompassed approximately 17% of all responses, and the top three most frequently occurring responses encompassed approximately 42% of all responses. The proportion of data encompassed by these most frequently occurring responses again indicated some common notion of salience across participants.[2]

---

[2] Note that participants tended to report items in the scene rather than a global label for the scene. Scene 18 ("Kitchen") was the only scene for which the most common first response was the

3.2 Results for DRM lists task

Summary statistics for performance on the recall and recognition phases of the DRM list task were calculated by lists and by participants using SPSS 25.0 (IBM Corp., 2017). All 30 participants were included in the analyses. Before summary statistics were calculated, the data were cleaned. That is, duplicate responses from the raw data were deleted. For example, if a participant recalled *night* two times, the second instance was deleted. Due to these deletions, the number of times a given word was typed was equal to the number of participants who typed it. For example, if *cold* was typed 9 times, it was typed by 9 participants. In addition, obvious spelling errors were corrected. For example, *physcian* was corrected to *physician*.

3.2.1 Results for the recall phase of the DRM lists task

Performance during the recall phase by lists is shown in Table 3-2. For each list, correct recall was scored as a raw count of the total number of correct responses that each list elicited by all participants and as a proportion of 450 (15 words per list seen by each of 30 participants). For example, across participants, list 3 (*chair*) elicited 189 correct recalls, which means that 42% of all the words presented were recalled (189/450=0.42). Lures were scored as a raw count of the total number of lures that each list elicited by all opportunities for lure recall for that list and as a proportion of opportunities in which it was recalled for each list. For example, list 5 (*cold*) elicited a total of 9 lure recalls, which means that it was recalled in 9 of 30 opportunities (i.e., by 9 of 30 participants; 0.30). Intrusions were defined as recalled words that were neither presented words nor lures. Intrusions were scored as a raw count and as a proportion of all the words each

---

name of the scene rather than an item in it. This finding suggests that participants focused on reporting the objects in the scene rather than the gist of the scene.

list elicited. For example, list 6 (*doctor*) elicited 18 intrusions; participants recalled a total of 252 words for that list, so 7.14% (18/252) of those words were intrusions.

Table 3-2. Recall performance on DRM lists analyzed by lists.

Measure	Mean	Standard Deviation	Minimum	Maximum	Range
Correct Recall (Proportion)	0.46	0.04	0.42	0.51	0.09
Lure Recall (Proportion)	0.29	0.1	0.20	0.43	0.23
Intrusions (Proportion)	0.1	0.02	0.07	0.13	0.05
Correct Recall (Count)	208.83	17.35	189	228	39
Lure Recall (Count)	8.67	2.88	6	13	7
Intrusions (Count)	23.17	3.92	18.00	28	10

Analyses of variance (ANOVAs) were performed to determine whether the lists differed in how many correct recalls, lures, or intrusions they elicited. The obtained results were not significant at the $p < 0.05$ level, indicating that the lists did not differ significantly in these properties. Results of the ANOVAs appear in Tables 3-3, 3-4, and 3-5.

Table 3-3. Results of ANOVA for correct recall by lists.

	Sum of Squares	Degrees of Freedom	Mean Square	*F* Statistic	*p*-value
Between Groups	50.16	5	10.03	2.07	0.07
Within Groups	842.57	174	4.84		
Total	892.73	179			

Table 3-4. Results of ANOVA for lure recall by lists.

	Sum of Squares	Degrees of Freedom	Mean Square	*F* Statistic	*p*-value
Between Groups	1.38	5	0.28	1.35	0.25
Within Groups	35.60	174	0.21		
Total	36.98	179			

Table 3-5. Results of ANOVA for intrusions by lists.

	Sum of Squares	Degrees of Freedom	Mean Square	*F* Statistic	*p*-value
Between Groups	2.56	5	0.51	0.12	0.99
Within Groups	757.10	174	4.35		
Total	759.66	179			

Table 3-6 shows performance on the recall phase by participants. Correct recall was measured as a raw count and as a proportion of correct recall of the 90 total words presented. Lure recall was measured as a raw count and as a proportion of the 6 total lures presented. Intrusions were measured as a raw count and as a proportion of all the words that each participant recalled.

Table 3-6. Performance on DRM lists analyzed by participants.

Measure	Mean	Standard Deviation	Minimum	Maximum	Range
Correct Recall (Proportion)	0.46	0.09	0.26	0.62	0.37
Lure Recall (Proportion)	0.29	0.21	0	0.83	0.83
Intrusions (Proportion)	0.07	0.13	0	0.74	0.74
Correct Recall (Count)	41.77	8.33	23	56	33
Lure Recall (Count)	1.73	1.26	0	5	5
Intrusions (Count)	4.63	11.58	0	65	65

In summary, the lists were not significantly different in terms of elicitation of correct recall, critical lures, and intrusions. Participants recalled slightly less than half of each list and less than 10% of responses were intrusions, indicating that they were attending to the task. The lists also elicited critical lures, suggesting that the lists worked as intended by eliciting false recall of a semantically-related word.

3.2.2 Results for the recognition phase of the DRM lists task

To analyze the recognition phase, the following values were calculated; hits (participant correctly recognized a presented word), misses (participant did not recognize a presented word), false alarms for lures (participant falsely recognized a lure), false alarms for nonlures (participant falsely recognized a word not presented), correct rejections of lures (participant rejected a lure), and correct rejections of nonlures (participant rejected a nonlure). In all, 18 items that appeared in the recall phase were presented in the recognition phase, so the proportions of hits and misses are each a proportion of 18. The proportions of false alarms and correct rejections of lures are proportions of 6 because 6 lures were presented. The proportion of false alarms and correct rejections of nonlures are proportions of 72 because 72 nonlures were presented. As indicated by the raw counts and proportions for false alarms of lures, the lures appeared to have had their intended effect of eliciting false recognition—apparently, the presentation of semantically-related words activated the lures sufficiently in participants' semantic systems such that they were later falsely recognized. The summary statistics appear in Table 3-7 below.

Table 3-7. Summary statistics for the recognition phase of DRM list task.

	Mean	Standard Deviation	Minimum	Maximum	Range
Hit (Proportion)	0.73	0.18	0.33	1	0.67
Miss (Proportion)	0.27	0.18	0	0.67	0.67
False Alarm for Lure (Proportion)	0.73	0.24	0	1	1
Correct Rejection for Lure (Proportion)	0.27	0.24	0	1	1
False Alarm for Nonlure (Proportion)	0.09	0.10	0	0.40	0.40
Correct Rejection for Nonlure (Proportion)	0.91	0.10	0.60	1	0.40
Hit (Count)	13.07	3.17	6	18	12
Miss (Count)	4.93	3.17	0	12	12
False Alarm for Lure (Count)	4.40	1.43	0	6	6
Correct Rejection for Lure (Count)	1.60	1.43	0	6	6
False Alarm for Nonlure (Count)	6.30	7.35	0	29	29
Correct Rejection for Nonlure (Count)	65.70	7.35	43	72	29

The *d'* statistic is a measure "reflecting the discrimination of true listed items versus the non-listed lure items, pooled across lists" (Tewolde, Bishop, & Manning, 2018, p. 512–513). It was calculated by converting hit and lure false alarm rates to *z*-scores, then subtracting the lure false alarm *z*-score from the hit *z*-score for each participant. *d'* for each participant appears in Appendix F. This measure will be compared to individuals with ASD in a future study.

Finally, variables related to recall and recognition were examined for correlations with scores on the SRS–2, EOWPVT–4, PPVT–4, and TONI–4. Pearson correlation coefficients were calculated. A positive correlation ($r(29) = 0.716$, $p < 0.01$) was found between PPVT–4 and EOWPVT–4 standard scores, reflecting a positive correlation between receptive and expressive vocabulary, as would be expected. A negative correlation significant at $p < 0.05$ was found between EOWPVT–4 standard scores and intrusions (as a raw count and as a proportion of words recalled), indicating that, as expressive vocabulary increased, intrusions decreased. Similarly, a negative correlation significant at $p < 0.05$ was found between PPVT–4 standard scores and intrusions (as a proportion of words recalled), indicating that, as receptive vocabulary increased, intrusions decreased. The correlation between PPVT–4 scores and intrusions as a raw count was not significant ($p = 0.057$). Negative correlations significant at $p < 0.05$ between TONI–4 standard scores and false alarms for lures and correct rejections for lures (as raw counts and proportions) were found as well, which might suggest there was some relationship between general cognitive ability and performance on these measures. The correlation table appears in Tables 3-8 and 3-9.

To summarize, participants were accurate in their recognition of words previously presented and their rejection of non-lures not previously presented, but they often showed false recognition for critical lures. These findings suggest that participants recognized the words previously presented (i.e., were performing the task as intended). They also demonstrate that the presentation of semantically-related words elicited false recall of critical lures, suggesting that the semantically-related words activated another semantically-related word, thereby supporting the spreading activation theory of semantic organization (Collins and Loftus, 1975). Correlations suggested that, as expressive and receptive vocabulary increased, intrusions tended to decrease, perhaps suggesting that stronger vocabulary skills helped participants organize the initially presented words into coherent gists or schemas that aided in reducing unrelated intrusions.

Table 3-8. Correlation table for variables measured during the recall phase.

			Variables measured during recall phase					
		Age (months)	Lure Recall (Count)	Lure Recall (Proportion)	Intrusion (Count)	Intrusion (Proportion)	Correct Recall (Count)	Correct Recall (Proportion)
Age (months)	Pearson Corr.	1	-0.098	-0.098	-0.197	-0.227	0.001	0.001
	Sig. (2-tailed)		0.606	0.606	0.296	0.228	0.996	0.996
TONI–4 Index Score	Pearson Corr.	0.141	0.224	0.224	-0.272	-0.25	-0.03	-0.03
	Sig. (2-tailed)	0.457	0.234	0.234	0.146	0.184	0.876	0.876
SSRS–2 Total T Score	Pearson Corr.	0.094	0.234	0.234	-0.114	-0.053	0.085	0.085
	Sig. (2-tailed)	0.622	0.214	0.214	0.55	0.781	0.656	0.656
EOWPVT–4 Standard Score	Pearson Corr.	0.1	0.147	0.147	-.444*	-.455*	-0.028	-0.028
	Sig. (2-tailed)	0.6	0.439	0.439	0.014	0.012	0.882	0.882
PPVT–4 Standard Score	Pearson Corr.	0.018	0	0	-0.351	-.365*	0.044	0.044
	Sig. (2-tailed)	0.924	1	1	0.057	0.048	0.816	0.816

Note. * denotes a correlation significant at $p < 0.05$.

Table 3-9. Correlation table for variables measured during the recognition phase.

		Variables measured during recognition phase											
		Hit (Count)	Hit (Proportion)	False Alarm for Lure (Count)	False Alarm for Lure (Proportion)	False Alarm for Nonlure (Count)	False Alarm for Nonlure (Proportion)	Miss (Count)	Miss (Proportion)	Correct Rejection for Lure (Count)	Correct Rejection for Lure (Proportion)	Correct Rejection for Nonlure (Count)	Correct Rejection for Nonlure (Proportion)
Age (months)	Pearson Corr.	-0.266	-0.266	-0.121	-0.121	-0.276	-0.276	0.266	0.266	0.121	0.121	0.276	0.276
	Sig. (2-tailed)	0.155	0.155	0.525	0.525	0.139	0.139	0.155	0.155	0.525	0.525	0.139	0.139
TONI–4 Index Score	Pearson Corr.	0.029	0.029	.400*	.400*	-0.246	-0.246	-0.029	-0.029	-.400*	-.400*	0.246	0.246
	Sig. (2-tailed)	0.879	0.879	0.029	0.029	0.19	0.19	0.879	0.879	0.029	0.029	0.19	0.19
SSRS–2 Total T Score	Pearson Corr.	-0.056	-0.056	0.159	0.159	-0.308	-0.308	0.056	0.056	-0.159	-0.159	0.308	0.308
	Sig. (2-tailed)	0.768	0.768	0.402	0.402	0.098	0.098	0.768	0.768	0.402	0.402	0.098	0.098
EOWPVT–4 Standard Score	Pearson Corr.	0.039	0.039	0.237	0.237	-0.104	-0.104	-0.039	-0.039	-0.237	-0.237	0.104	0.104
	Sig. (2-tailed)	0.837	0.837	0.207	0.207	0.585	0.585	0.837	0.837	0.207	0.207	0.585	0.585
PPVT–4 Standard Score	Pearson Corr.	0.075	0.075	0.18	0.18	-0.096	-0.096	-0.075	-0.075	-0.18	-0.18	0.096	0.096
	Sig. (2-tailed)	0.695	0.695	0.341	0.341	0.615	0.615	0.695	0.695	0.341	0.341	0.615	0.615

Note. * denotes a correlation significant at $p < 0.05$.

Chapter 4

Discussion

4.1 Discussion of results of the present study

Participants with TD performed two tasks designed to assess the integration of top-down and bottom-up information: scene-viewing and list recall and recognition. The scene-viewing task results are preliminary because 20 of the 60 scenes for which data were collected were coded for analysis. Norms for recall do not yet exist for the scenes developed by Öhlschläger and Võ (2017), so the results of the present study may be regarded as normative in the sense that they preliminarily establish what individuals with TD recall from the scenes. This performance may be interpreted as a measure of what the participants extracted as salient. On the scene-viewing task, participants demonstrated variability in their responses—only an average of 2.2 responses were provided by the majority of participants for each scene. There was an average of 16.60 unique responses (responses made by only one participant) per scene, again indicating diversity of responses. However, although only an average of 38% of participants reported the same first item per scene, these items were often informative of the nature of the scene (e.g., *sink* for a bathroom scene, *bed* for a bedroom scene) and were often visually prominent in the scene. The nature of these responses indicates that participants selected the first response using similar mechanisms at least for some of the scenes. Moreover, the most frequently occurring response in each scene was reported by an average of 77% of participants, indicating general agreement on the most salient item in the scenes. In addition, the most frequently occurring response for each scene represented an average of 17% of the data for each scene, and the three most frequently occurring responses for each scene represented an average of 42% of the data for each scene. These frequently occurring responses again suggest a common means of determining saliency among participants. While an informal inspection of nameability and number of objects in the scene did not reveal clear patterns that indicated whether scenes were above or below the mean of the measures (the number of responses made by at least 50% of participants, the number of unique responses, the proportions of participants who named the most frequently occurring response and the most frequently occurring *first* response, the proportion of all responses represented by the most frequently occurring response, and the proportion of all responses represented by the three most frequently occurring responses), more formal and systematic analysis should be incorporated into future studies. For example, the number of objects in each scene could be counted and used as a variable. Ratings of nameability as well as global labels (names) for each scene could be provided by participants in a separate study.

The scene-viewing task did not produce expected results that gave clear insight into what participants interpreted as salient. There was more variability than expected in the participants with TD. Although the results of participants with TD are not as clear as expected, perhaps by comparing these results with the responses of individuals with ASD in a future study, clearer patterns in the responses provided by individuals with TD may emerge, granting insight into the means by which individuals with TD assign salience. With regard to specific scenes, some elicited more consistent responses, while others elicited more unique responses, but it is important to have scenes of both types because they can reveal patterns of performance when, for example, the scene does not represent an easily named schema. In the present study, 20 scenes were coded, but when the remaining scenes are coded, a pattern of performance that varies depending on, for example, nameability and number of objects in the scene may emerge.

Furthermore, the data coding process proved more challenging than anticipated. As the investigator compared results with the undergraduate coders, differences in applying the coding rules emerged. One challenge was how to collapse responses. Inspection of the scene helped determine whether two different words referred to the same object (e.g., *jacket* and *coat*). Another challenge was removing adjectives—the rule needed to be refined so that it was not applied too liberally. For example, removing *colored* from the commonly used term *colored pencils* in Scene 16 would remove important information that indicates that the participant saw these objects as colored pencils, not just as pencils. Similarly, in Scene 42 removing *metal* and *wire* from *metal baskets* and *wire racks* removes the information that makes these responses different from *bins/box/boxes/crates* because the latter series of words appear to be referring to the wooden bins, not the wire racks, in the scene. Visual inspection of the scenes helped clarify the referents of participants' responses. In spite of these challenges, instructing participants to respond only with single nouns or to type the labels for the scenes as well as the objects would yield inauthentic responses, bias their reporting, and potentially diminish the variability in typical responses. In future studies, refinement of the coding rules and consistent inspection of the scenes will be useful.

Additional means of coding may be useful as well. For example, participants with ASD and TD may differ on amount or type of information provided, providing insight into their adherence to or violation of the Gricean maxims of Quantity and Manner when providing descriptions of what they saw (e.g., Ghaziuddin & Gerstein, 1996; Surian, Baron-Cohen, & Van der Lely, 1996; and see anecdote in de Marchena & Eigsti, 2016, p. 45). It may also be important to inventory each item in each scene so that measures of accurate and false visual recall may be obtained.

On the DRM task, participants falsely recalled and recognized lures, indicating that the semantically-related words activated a semantic network of concepts (Collins & Loftus, 1975;

Roediger & McDermott, 1995). Although this false recall could be considered an error, it reflects the organization of semantic or conceptual knowledge that individuals with TD possess. In addition, as receptive and expressive vocabulary scores increased, the number of intrusions decreased, perhaps indicating that stronger vocabulary skills (reflections of semantic knowledge) helped participants organize the information presented during the initial list presentation. With this cognitive organization in place, they may have been less susceptible to recognition of unrelated words.

In the current study, the six DRM lists that were used did not elicit different amounts of correct recall. In a future study, it may be useful to include additional DRM lists to examine whether different lists elicit correct recall, lure recall, and intrusion recall differently. Participants with ASD may perform less consistently across lists than participants with TD given that individuals with ASD may have vocabulary deficits and/or atypicalities rooted in lexical acquisition difficulties (e.g., Arunachalam & Luyster, 2016; Charman, Drew, Baird, & Baird, 2003; Kover, McDuffie, Hagerman, & Abbeduto, 2013; Volden, et al., 2011) and atypical category knowledge (e.g., Gastgeb et al., 2009; Gastgeb et al., 2006; Klinger & Dawson, 2001; Strauss et al., 2012). Therefore, it will be important to have a variety of lists to assess whether performance differs between groups.

Previous studies using the DRM lists with individuals with TD have been conducted. However, these studies often reported statistics related to serial position effects and other psycholinguistic factors related to recall and recognition (Deese, 1959; Gallo, McDermott, Percer, & Roediger, 2001; Roediger & McDermott, 1995; Roediger, Watson, McDermott, & Gallo, 2001; Stadler, Roediger, & McDermott, 1999), utilized different stimuli (Metzger et al., 2008), and did not report statistics that can be compared to those of the present study. Other studies have compared participants with TD and ASD on DRM tasks. However, differences in stimuli used and age group (e.g., Tewolde et al., 2018; Wojcik et al., 2018) as well as response mode and statistics reported (e.g., data reported according to certainty responses in Beversdorf and colleagues [2000] and Bowler, Gardiner, Grice, and Saavalainen [2000]) make it difficult to compare the data of the participants with TD in these studies to the participants with TD in the present study. Studies also varied on how many lists were presented, which lists were presented, and whether presentation was auditory, visual, or both. Therefore, the comparison shown in Table 4-1 should be considered with caution. Table 4-1 displays the statistics available from the studies cited; some of the studies cited did not report the statistics utilized in the present study and therefore do not appear in the table. The table shows some similarities, many differences, and many gaps in the data provided.

Table 4-1. Comparison of previous studies using a DRM paradigm to the present study.

	Correct recall (Prop)	False recall of lure (Prop)	Recall of intrusions (Prop)	Hit (Prop)	Miss (Prop)	False Alarm for Lure (Prop)	Correct Rejection for Lure (Prop)	False Alarm for Nonlure (Prop)	Correct Rejection for Nonlure (Prop)
Present study	0.46 (0.09)	0.29 (0.21)	0.07 (0.13)	0.73 (0.18)	0.27 (0.18)	0.73 (0.24)	0.27 (0.24)	0.09 (0.10)	0.91 (0.10)
Bowler and colleagues (2000)	0.40 (0.05)	0.43 (0.21)	0.10 (0.13)						
Deese (1959)		0.24 (0.12)							
Gallo and colleagues (2001)	0.60 for auditory presentation, 0.58 for visual presentation	0.46 for auditory presentation, 0.38 for visual presentation	0.45 for auditory presentation, 0.32 for visual presentation, 0.39 average between conditions					0.07 average between conditions	
Metzger and colleagues (2008)	0.59 for college-age group	0.16 for college-age group		0.8 for college-age group		0.72 for college-age group			
Stadler and colleagues (1999)						0.66			
Wojcik and colleagues (2018)	0.74 (0.08)	0.69 (0.19)	0.02 (0.01)			0.27 (0.21)	0.74 (0.20)		

Note. When available, means appear followed by standard deviations in parentheses.

4.2 Limitations of the current study

The present study is limited by several factors. First, four of the participants had slightly elevated *T*-scores; their data could be analyzed separately or even compared as a small-*N* group to participants with typical *T*-scores and participants with ASD. There were also limitations in the analysis of the scene-viewing data. Twenty of the 60 scenes presented in the scene-viewing task were coded, so the results discussed in this thesis are preliminary. Differences in the application of coding rules to the scene-viewing data indicate the need to make more exceptions to the rules and to require careful visual inspection of the scenes prior to coding and collapsing responses that have the same referent. Coding each scene for nameability (as rated by participants in a separate experiment) and number of objects may reveal response patterns that differ on these factors.

4.3 Relation of the current study to ASD

The data from individuals with TD who participated in this study will be utilized as a comparison group for a future study with individuals with ASD. As the studies discussed below indicate, there are strong theoretical reasons to believe that individuals with ASD will perform differently on these tasks. Therefore, a comparison in performance with individuals with TD may help elucidate how individuals with ASD process complex information (Minshew et al., 1992; Minshew, Goldstein, Siegel, 1997; Williams, Minshew, and Goldstein, 2015). It is suspected that these individuals will have more heterogeneity in items selected as salient in scenes and less critical lure false recall and recognition for DRM lists due to less efficiently organized or less robust conceptual knowledge. The evidence for these hypotheses about performance by individuals with ASD on the tasks used in the current study is reviewed below.

4.4 Atypical categories in individuals with ASD

Individuals with ASD show atypicalities in abstracting prototypes (forming a category based on many exemplars) in experimental tasks involving categorization of faces, objects and animals, and novel characters (Gastgeb et al., 2009; Gastgeb et al., 2006; Klinger & Dawson, 2001; Strauss et al., 2012). Difficulty extracting prototypes could lead to less robust categories in these individuals. Converging evidence of less robust categories comes from electrophysiological research. Fiebelkorn, Foxe, McCourt, Dumas, and Molholm (2013) found that deficits in category knowledge manifested as a weaker response to members of the same category in individuals with ASD.

4.5 Schemas in visual tasks in individuals with ASD

The top-down information-processing component of scene perception involves knowledge of the world. If individuals with ASD have less robust prototypes and categories, they may have impaired top-down resources. That is, if they have atypical knowledge of the world (including atypical schemas), they bring atypical top-down processing to a scene processing task. Moreover, if they have difficulty integrating multiple pieces of information, they may have difficulty integrating top-down information with bottom-up information. Therefore, scene-viewing tasks are ideal for investigating these cognitive differences.

Studies have found that individuals with ASD have impaired scene gist perception (the extraction of the global or general nature of a scene rather than its individual parts (e.g., Vanmarcke and colleagues, 2016a, 2016b, 2016c). Loth, Gómez, and Happé (2011) found that individuals with ASD integrated top-down story information with bottom-up visual information in scenes less effectively than individuals with TD. Benson, Castelhano, Au-Yeung, and Rayner (2012) found that individuals with ASD and individuals with TD had similar reaction times for a simple visual processing task requiring them to spot the difference between two scenes. However, the participants with ASD had longer reaction times and different eye tracking behavior on the complex which-one's-weird task that involved spotting an incongruous "weird" object in a scene. The authors suggest that their findings indicate that knowledge of what typically appears in a scene (schematic knowledge) is impaired in individuals with ASD and that visual processing is impaired for complex tasks that involve integrating top-down and bottom-up information but not for simple tasks that rely on bottom-up information processing.

4.6 Schemas in language tasks in individuals with ASD

If individuals with ASD have atypical conceptual networks, spreading activation (Collins & Loftus, 1975) may proceed in a less effective manner. One behavioral manifestation of such processing would be less susceptibility to falsely remembering of critical lures. Beversdorf and colleagues (2000) found that individuals with ASD were less susceptible to false memories after being presented with the DRM lists, reflecting impaired semantic networks or impaired use of context. Computational modeling supports this finding (Beversdorf, Narayanan, Hillier, & Hughes, 2007). However, Bowler, Gardiner, Grice, and Saavalainen (2000) did not replicate these findings with DRM lists. Recently, Wojcik and colleagues (2018) found that adults with ASD had less recall of critical lures on a DRM-style task, but Tewolde, Bishop, and Manning (2018) found that children with ASD did not differ from children with TD in discriminating words presented from critical lures.

Tasks other than the DRM lists have been used to assess recall of organized verbal information in individuals with ASD. Kamio and Toichi (2007) did not find differences in false recall of sentences with varying numbers of propositions between individuals with TD, individuals with Asperger's syndrome, or individuals with high-functioning autism. However, other list-learning experiments have demonstrated reduced use of semantic context in individuals with ASD (Bowler, Gaigg, & Gardiner, 2008, 2009), and adults with ASD have been found to show reduced use of gist recall when recalling narratives (Williams, Minshew, Goldstein, & Mazefsky, 2017).

Another way to measure semantic knowledge is to ask individuals to list members of a category. As Carmo, Duarte, Pinho, Marques, and Filipe (2015) summarize, asking individuals with ASD to generate words that belong to the same category (category or semantic fluency) or that start with the same letter (letter or phonemic fluency) has yielded mixed results—while some studies have shown impairments relative to individuals with TD, others have not. In their own study, Carmo and colleagues (2015) found that individuals with ASD did not differ significantly from individuals with TD in their use of clustering or switching.[3] They only differed in terms of timing measures—the individuals with ASD had difficulty getting started, which could reflect less robust semantic networks. However, the authors interpret their findings as a demonstration of a timing deficit in individuals with ASD. Supporting these findings, Carmo and colleagues (2017) found that individuals with ASD perform similarly to individuals with TD in a visual category identification task but showed slower reaction times. Alternatively, these findings could indicate poorer semantic organization in individuals with ASD—if networks are less robust in individuals with ASD, accessing them could take longer even though the response itself is similar.

---

[3] Clustering is naming several items from a subcategory within a superordinate category. Switching is changing from one subcategory to another. For example if the superordinate category were *animals*, a participant may cluster by naming three jungle animals and then switch to household pets by naming three in a cluster as in *tiger*, *lion*, *rhino*, *cat*, *dog*, *hamster*.

Boucher (1988) found that individuals with ASD named the same number of exemplars when a category name was provided (cued recall) but not when one was *not* provided (miscellaneous word generation), suggesting poorer semantic organization becomes apparent when no explicit structure is provided. This finding supports Klinger and Dawson (2001), who found improved performance in prototype abstraction given scaffolding in the form of an explicit rule. Minshew and Goldstein (1993) found that individuals with ASD showed poorer clustering and more unrelated intrusions when they recalled lists of words from the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987), which supports the claim of poorer semantic organization in individuals with ASD. On a different test of verbal fluency (the Verbal Expression Subtest of the Illinois Test of Psycholinguistic Aptitudes; Kirk, McCarthy, & Kirk, 2004), Pastor-Cerezuela, Fernández-Andrés, Feo-Álvarez, and González-Sala (2016) found that individuals with ASD produced fewer clusters and switches, smaller clusters, and fewer words generated, corroborating the findings of Minshew and Goldstein (1993).

The evidence related to semantic knowledge as revealed by visual and language tasks suggests that individuals with ASD have difficulty extracting prototypes, less robust category knowledge, and difficulty integrating top-down semantic knowledge or context with bottom-up information. The tasks in the present experiment (recall from scene-viewing and word list recall and recognition) are particularly suited for a future study that assesses how individuals with ASD and individuals with TD differ in their semantic knowledge and their integration of top-down and bottom-up information.

Chapter 5

Conclusion

In the current study, preliminary results indicate that individuals with TD demonstrated variability in their responses to a scene-viewing task. However, they appeared to exhibit a common means of extracting salient items in scenes at least for some of the scenes. In a false memory list task, individuals with TD recalled and recognized critical lures that were semantically-related to words presented, indicating the activation of conceptual knowledge. The two experimental tasks in the present study are thought to involve complex information processing in the form of integration of top-down and bottom-up information. They are therefore appropriate tasks for comparing the complex information processing skills of individuals with TD and individuals with ASD. If individuals with ASD have less robust top-down category knowledge and/or impaired integration of top-down and bottom-up information, their performance will likely differ from that of individuals with TD in the present study. It is predicted that individuals with ASD will show less consistency than individuals with TD in their responses to the scene-viewing task, reflecting a different means of determining salience in a scene. It is also predicted that they will recall and recognize fewer critical lures, suggesting less robust or less efficient semantic networks or atypically organized semantic knowledge. The completion of the tasks developed in this thesis by individuals with ASD will help shed light on how these individuals integrate top-down and bottom-up information and how they organize their semantic knowledge.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association. https://dx.doi.org/10.1176/appi.books.9780890425596.dsm01

Arunachalam, S., & Luyster, R. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research, 9,* 810–828. https://dx.doi.org/10.1002/aur.1590

Bar, M. (2004). Visual objects in context. *Nature Reviews*, *5*, 617–629. https://dx.doi.org/10.1038/nrn1476

Benson, V., Castelhano, M. S., Au-Yeung, S. K., & Rayner, K. (2012). Eye movements reveal no immediate "WOW" ("which one's weird") effect in autism spectrum disorder. *The Quarterly Journal of Experimental Psychology*, *65*, 1139–1150. https://dx.doi.org/10.1080/17470218.2011.644305

Beversdorf, D. Q., Narayanan, A., Hillier, A., & Hughes, J. D. (2007). Network model of decreased context utilization in autism spectrum disorder. *Journal of Autism and Developmental Disabilities*, *37,* 1040–1048. https://dx.doi.org/10.1007/s10803-006-0242-7

Beversdorf, D. Q., Smith, B. W., Crucian, G. P., Anderson, J. M., Keillor, J. M., Barrett, A. M., . . . Heilman, K. M. (2000). Increased discrimination of "false memories" in autism spectrum disorder. *Proceedings of the National Academy of Sciences of the United States of America, 97,* 8734–8737. https://dx.doi.org/10.1073/pnas.97.15.8734

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*, 597–600. http://dx.doi.org/10.1037/h0037158

Boucher, J. (1988). Word fluency in high-functioning autistic children. *Journal of Autism and Developmental Disabilities*, *18*, 637–645. https://dx.doi.org/10.1007/BF02211881

Bowler, D. M., Gaigg, S. B., & Gardiner, J. M. (2008). Effects of related and unrelated context on recall and recognition by adults with high-functioning autism spectrum disorder. *Neuropsychologia*, *46*, 993–999. https://dx.doi.org/10.1016/j.neuropsychologia.2007.12.004

Bowler, D. M., Gaigg, S. B., & Gardiner, J. M. (2009). Free recall learning of hierarchically organised lists by adults with Asperger's Syndrome: Additional evidence for diminished relational processing. *Journal of Autism and Developmental Disorders*, *39*, 589–595. https://dx.doi.org/10.1007/s10803-008-0659-2

Bowler, D. M., Gardiner, J. M., Grice, S., and Saavalainen, P. (2000). Memory illusions: False recall and recognition in adults with Asperger's Syndrome. *Journal of Abnormal Psychology*, *109*, 663–672. http://dx.doi.org/10.1037/0021-843X.109.4.663

Boyce, S. J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 556–566. http://dx.doi.org/10.1037/0096-1523.15.3.556

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2010). *Test of Nonverbal Intelligence–Fourth Edition*. Austin, TX: Pro-Ed.

Carmo, J. C., Duarte, E., Pinho, S., Marques, J. F., & Filipe, C. N. (2015). Verbal fluency as a function of time in autism spectrum disorder: An impairment of initiation processes?

Journal of Clinical and Experimental Neuropsychology, *37*, 710–721. https://dx.doi.org/10.1080/13803395.2015.1062082

Carmo, J. C., Souza, C., Gonçalves, F., Pinho, S., Filipe, C. N., & Lachmann, T. (2017). Effects of categorical representation on visuospatial working memory in autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology*, *39*, 131–141. https://dx.doi.org/10.1080/13803395.2016.1207754

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, *30*, 213–236. https://dx.doi.org/10.1017/S0305000902005482

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428. http://dx.doi.org/10.1037/0033-295X.82.6.407

Constantino, J. N., & Gruber, C. P. (2012). *Social Responsiveness Scale, Second Edition*. Torrance, CA: Western Psychological Services.

Davis, N. O., & Carter, A. S. (2014). Social development in autism. In F. R. Volkmar, S. J. Rogers, R. Paul, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, and brain mechanisms* (4th ed., Vol. 1, pp. 212–229). Hoboken, NJ: John Wiley & Sons, Inc.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22. http://dx.doi.org/10.1037/h0046671

Delis, D.C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *CVLT: California Verbal Learning Test Manual, Adult version*. San Antonio, TX: The Psychological Corporation.

de Marchena, A., & Eigsti, I.-M. (2016). The art of common ground: Emergence of a complex pragmatic skill in adolescents with autism spectrum disorders. *Journal of Child Language*, *43*, 43–80. https://dx.doi.org/10.1017/S0305000915000070

Demiral, Ş. B., Malcolm, G. L. and Henderson, J. M. (2012). ERP correlates of spatially incongruent object identification during scene viewing: Contextual expectancy versus simultaneous processing. *Neuropsychologia*, *50*, 1271–1285. https://dx.doi.org/10.1016/j.neuropsychologia.2012.02.011

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–Fourth Edition.* Minneapolis, MN: Pearson.

Fiebelkorn, I. C., Foxe, J. J., McCourt, M. E., Dumas, K. N., & Molholm, S. (2013). Atypical category processing and hemispheric asymmetries in high-functioning children with autism: Revealed through high-density EEG mapping. *Cortex*, *49*, 1259–1267. https://dx.doi.org/10.1016/j.cortex.2012.04.007

Gallo, D. A., McDermott, K. B., Percer J. M., & Roediger, H. L. (2001). Modality effects in false recall and false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 339–353. http://dx.doi.org/10.1037/0278-7393.27.2.339

Gastgeb, H. Z., Rump, K. M., Best, C. A., Minshew, N. J., & Strauss, M. S. (2009). Prototype formation in autism: Can individuals with autism abstract facial prototypes? *Autism Research*, *2*, 279–284. https://dx.doi.org/10.1002/aur.93

Gastgeb, H. Z., Strauss, M. S., & Minshew, N. J. (2006). Do individuals with autism process categories differently? The effect of typicality and development. *Child Development*, *77*, 1717–1729. https://dx.doi.org/10.1111/j.1467-8624.2006.00969.x

Ghaziuddin, M., & Gerstein, L. (1996). Pedantic speaking style differentiates Asperger syndrome from high-functioning autism. *Journal of Autism and Developmental Disorders*, *26*, 585–595. https://dx.doi.org/10.1007/BF02172348

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50,* 243–271. https://dx.doi.org/10.1146/annurev.psych.50.1.243

IBM Corp. (2017). IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex*, *17,* 951–961. https://dx.doi.org/10.1093/cercor/bhl006

Just, M.A., Cherkassky, V. L., Keller, T. A., & Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain*, *127,* 1811–1821. https://dx.doi.org/10.1093/brain/awh199

Kamio, Y., & Toichi, M. (2007). Memory illusion in high-functioning autism and Asperger's disorder. *Journal of Autism and Developmental Disorders, 37*, 867–876. https://dx.doi.org/10.1007/s10803-006-0214-y

Kirk, S., McCarthy, J., & Kirk, W. (2004). *Test Illinois de Aptitudes Psicolingüísticas* [*Illinois Test of Psycholinguistic Aptitudes*]. Sahagun: TEA Ediciones.

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology*, *13*, 111–124. http://dx.doi.org/10.1017/S0954579401001080

Kover, S. T., McDuffie, A. S., Hagerman, R. J., & Abbeduto, L. (2013). Receptive vocabulary in boys with autism spectrum disorder: Cross-sectional developmental trajectories. *Journal* *of Autism and Developmental Disorders*, *43*, 2696–2709. https://dx.doi.org/10.1007/s10803-013-1823-x

Loth, E., Gómez, J. C., & Happé, F. (2011). Do high-functioning people with autism spectrum disorder spontaneously use event knowledge to selectively attend to and remember context-relevant aspects in scenes? *Journal of Autism and Developmental Disorders, 41*, 945–961. https://dx.doi.org/10.1007/s10803-010-1124-6

MacWhinney, B., St. James, J., Schunn, C., Li, P., & Schneider, W. (2001). STEP—A system for teaching experimental psychology using E-Prime. *Behavior Research Methods, Instruments, & Computers*, *33*, 287–296. https://dx.doi.org/10.3758/BF03195379

Martin, N.A., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test–4*. Novato, CA: Academic Therapy Publications.

McPartland, J. C., Tillman, R. M., Yang, D. Y.-J., Bernier, R. A., & Pelphrey, K. A. (2014). The social neuroscience of autism spectrum disorder. In F. R. Volkmar, S. J. Rogers, R. Paul, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, and brain mechanisms* (4th ed., Vol. 1, pp. 482–496). Hoboken, NJ: John Wiley & Sons, Inc.

Metzger, R. L., Warren, A. R., Shelton, Price, J. T., Reed, A. W., & Williams, D. (2008). Do children "DRM" like adults? False memory production in children. *Developmental Psychology*, *44*, 169–181. http://dx.doi.org/10.1037/0012-1649.44.1.169

Microsoft Corporation (2016). Excel. Redmond, WA: Microsoft Corporation.

Miller, M. B., & Gazzaniga, M. S. (1998). Creating false memories for visual scenes. *Neuropsychologia*, *36,* 513–520. https://dx.doi.org/10.1016/S0028-3932(97)00148-6

Minshew, N. J., & Goldstein, G. (1993). Is autism an amnesic disorder? Evidence from the California Verbal Learning Test. *Neuropsychology*, *7*, 209–216. http://dx.doi.org/10.1037/0894-4105.7.2.209

Minshew, N. J., & Goldstein, G. (2001). The pattern of intact and impaired memory functions in autism. *Journal of Child Psychiatry, 42*, 1095–1101. https://dx.doi.org/10.1111/1469-7610.00808

Minshew, N. J., Goldstein, G., Muenz, L. R., & Payton, J. B. (1992). Neuropsychological functioning in nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology*, *14*, 749–761. https://dx.doi.org/10.1080/01688639208402860

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society, 3,* 303–316.

Öhlschläger, S., & Võ, M. L. (2017). SCEGRAM: An image database for semantic and syntactic inconsistencies in scenes. *Behavior Research Methods, 49,* 1780–1791. https://dx.doi.org/10.3758/s13428-016-0820-3

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3,* 519–526. https://dx.doi.org/10.3758/BF03197524

Pastor-Cerezuela, G., Fernández-Andrés, M. I., Feo-Álvarez, M., & González-Sala, F. (2016). Semantic verbal fluency in children with and without autism spectrum disorder: Relationship with chronological age and IQ. *Frontiers in Psychology*, *7,* 1–12. https://dx.doi.org/10.3389/fpsyg.2016.00921

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from https://www.pstnet.com

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology*, *21*, 803–814. http://dx.doi.org/10.1037/0278-7393.21.4.803

Roediger, H. L, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*, 385–407. https://dx.doi.org/10.3758/BF03196177

Russell, W. A., & Jenkins, J. J. (1954). *The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test*. (Tech. Rep. No. 11, Contract N8 ONR 66216, Office of Naval Research). University of Minnesota.

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*, 494–500. https://dx.doi.org/10.3758/BF03211543

Strauss, M. S., Newell, L. C., Best, C. A., Hannigen, S. F., Gastgeb, H. Z., & Giovannelli, J. L. (2012). The development of facial gender categorization in individuals with and without autism: The impact of typicality. *Journal of Autism and Developmental Disorders*, *42,* 1847–1855. https://dx.doi.org/10.1007/s10803-011-1428-1

Surian, L., Baron-Cohen, S., & Van der Lely, H. (1996). Are children with autism deaf to Gricean maxims? *Cognitive Neuropsychiatry*, *1*, 55–71. https://dx.doi.org/10.1080/135468096396703

Tewolde, F. G., Bishop, D. V. M., & Manning, C. (2018). Visual motion prediction and verbal false memory performance in autistic children. *Autism Research*, *11*, 509–518. https://dx.doi.org/10.1002/aur.1915

Vanmarcke, S., Mullin, C., Van der Hallen, R., Evers, K., Noens, I., Steyaert, J., & Wagemans, J. (2016a). In the eye of the beholder: Rapid visual perception of real-life scenes by young adults with and without ASD. *Journal of Autism and Developmental Disorders*, *46*, 2635–2652. https://dx.doi.org/10.1007/s10803-016-2802-9

Vanmarcke, S., Van Der Hallen, R., Evers, K., Noens, I., Steyaert, J., & Wagemans, J. (2016b). Ultra-rapid categorization of meaningful real-life scenes in adults with and without ASD. *Journal of Autism and Developmental Disorders*, *46*, 450–466. http://dx.doi.org/10.1007/s10803-015-2583-6

Vanmarcke, S., van Esch, L., Van der Hallen, R., Evers, K., Noens, I., Steyeart, J., & Wagemans, J. (2016c). Gist perception in adolescents with and without ASD: Ultra-rapid categorization of meaningful real-life scenes. *Research in Autism Spectrum Disorders*, *29–30*, 30–47. https://dx.doi.org/10.1016/j.rasd.2016.05.007

Volden, J., Smith, I. M., Szatmari, P., Bryson, S., Fombonne, E., Mirenda, P., . . . & Thompson, A. (2011). Using the Preschool Language Scale, Fourth Edition to characterize language in preschoolers with autism spectrum disorders. *American Journal of Speech-Language Pathology*, *20*, 200–208. https://dx.doi.org/10.1044/1058-0360(2011/10-0035)

Williams, D. L., Goldstein, G., & Minshew, N. J. (2006). Neuropsychological functioning in children with autism: Further evidence for disordered complex information-processing. *Child Neuropsychology*, *12*, 279–298. https://dx.doi.org/10.1080/09297040600681190

Williams, D. L., Minshew, N. J., & Goldstein, G. (2015). Further understanding of complex information processing in verbal adolescents and adults with autism spectrum disorders. *Autism*, *19*, 859–867. https://dx.doi.org/10.1177/1362361315586171

Williams, D. L., Minshew, N. J., Goldstein, G., & Mazefsky, C. A. (2017). Long-term memory in older children/adolescents and adults with autism spectrum disorder. *Autism Research*, *10*, 1523–1532. https://dx.doi.org/10.1002/aur.1801

Wojcik, D. Z., Díez, E., Alonso, M. A., Martín-Gilleros, M. V., Guisuraga-Fernández, Z., . . . & Fernandez, A. (2018). Diminished false memory in adults with autism spectrum disorder: Evidence of identify-to-reject mechanism impairment. *Research in Autism Spectrum Disorders*, *45*, 51–57. https://dx.doi.org/10.1016/j.rasd.2017.11.001

Appendix A

Brief description of scenes from Öhlschläger and Võ (2017) used in the current study

1: Kitchen	21: Bathroom	41: Table
2: Bathroom	22: Kitchen	42: Bathroom
3: Kitchen	23: Living room	43: Desk
4: Shoe rack	24: Kitchen	44: Kitchen
5: Kitchen	25: Bathroom	45: Kitchen
6: Living room	26: Bedroom	46: Living room
7: Bedroom	27: Shoe rack	47: Kitchen
8: Kitchen	28: Bathroom	48: Living room
9: Kitchen	29: Kitchen	49: Storage cubes
10: Kitchen	30: Desk	50: Kitchen
11: Shoe rack	31: Kitchen	51: Bathroom
12: Bathroom	32: Shoe rack	52: Kitchen
13: Cabinet	33: Living room	53: Bathroom
14: Bathroom	34: Dining room	54: Kitchen
15: Kitchen	35: Bathroom	55: Bedroom
16: Desk	36: Desk	56: Kitchen
17: Bedroom	37: Kitchen	57: Living room
18: Kitchen	38: Bathroom	58: Kitchen
19: Living room	39: Kitchen	59: Desk
20: Kitchen	40: Bedroom	60: Kitchen

Appendix B

DRM lists used in the current study

Lure	Stimulus Words for Each List														
anger	mad	fear	hate	rage	temper	fury	ire	wrath	happy	fight	hatred	mean	calm	emotion	enrage
black	white	dark	cat	charred	night	funeral	color	grief	blue	death	ink	bottom	coal	brown	gray
bread	butter	food	eat	sandwich	rye	jam	milk	flour	jelly	dough	crust	slice	wine	loaf	toast
chair	table	sit	legs	seat	couch	desk	recliner	sofa	wood	cushion	swivel	stool	sitting	rocking	bench
cold	hot	snow	warm	winter	ice	wet	frigid	chilly	heat	weather	freeze	air	shiver	arctic	frost
doctor	nurse	sick	lawyer	medicine	health	hospital	dentist	physician	ill	patient	office	stethoscope	surgeon	clinic	cure

Appendix C

The number of responses made by at least 50% of participants for each scene and the number of unique responses for each scene

Scenes Sorted by Responses Made by at Least 50% of Participants				**Scenes Sorted by Number of Unique Responses**		
Number of responses made by at least 50% of participants: *M* = 2.2, *SD* = 1.28				Number of unique responses: *M* = 16.60, *SD* = 6.32		
Scene	Responses Made by at Least 50% of Participants	Number of Unique Responses		Scene	Responses Made by at Least 50% of Participants	Number of Unique Responses
11	4	20		13	0	29
15	4	9		18	0	26
43	4	11		41	1	23
10	3	16		42	1	23
14	3	19		37	0	21
38	3	7		11	4	20
39	3	12		12	2	20
4	3	16		14	3	19
40	3	7		44	2	18
1	2	8		45	2	18
12	2	20		10	3	16
16	2	13		17	2	16
17	2	16		4	3	16
44	2	18		16	2	13
45	2	18		39	3	12
41	1	23		43	4	11
42	1	23		15	4	9
13	0	29		1	2	8
18	0	26		38	3	7
37	0	21		40	3	7

Appendix D

Data regarding the most frequently occurring responses and the most frequently occurring *first* responses

Scenes Sorted by Proportion of Participants Who Made Most Frequently Occurring Response					Scenes Sorted by Proportion of Participants Who Made Most Frequently Occurring *First* Response			
Number and proportion of participants who made most frequently occurring response: Count: M = 23.2, SD = 5.41 Proportion: M = 0.77, SD = 0.18					Number and proportion of participants who made most frequently occurring *first* response: Count: M = 11.5, SD = 3.15 Proportion: M = 0.38, SD = 0.11			
Scene	Most Frequently Occurring Response	Number of Participants who Named It	Proportion of Participants who Named It		Scene	Most Frequently Occurring *First* Response	Number of Participants who Named It	Proportion of Participants who Named It
45	banana/bananas	29	0.97		1	dishwasher	18.00	0.60
39	juice/orange juice	29	0.97		12	toilet	16.00	0.53
12	toilet	29	0.97		45	banana/bananas	16.00	0.53
14	sink	28	0.93		43	bottles/paint/paints/ paint bottles/paint tubes	15.00	0.50
1	dishwasher	28	0.93		15	pasta/spaghetti	14.00	0.47
43	bottles/paint/paints /paint bottles/paint tubes	27	0.90		4	shoe/shoes	13.00	0.43
38	sink	27	0.90		39†	banana/bananas & juice/orange juice	13.00	0.43
17	armoire/cabinet/closet/“cabinet/closet”/shelving unit/storage unit/“unit/wardrobe”/wardrobe	26	0.87		11	shoes	12.00	0.40
4	shoe/shoes	26	0.87		17	armoire/cabinet/closet/“cabinet/closet”/shelving unit/storage unit/“unit/wardrobe”/wardrobe	12.00	0.40
10	sink	26	0.87		44	carrot/carrots	12.00	0.40
40	bed	24	0.80		40	bed	11.00	0.37
11	shoes	24	0.80		38	sink	10.00	0.33
15	table	22	0.73		41	television/tv	10.00	0.33
44	carrot/carrots	21	0.70		10	sink	9.00	0.30
41	television/tv	21	0.70		13†	cactus & hat	9.00	0.30
16*	desk & color pencils/colored pencils/colorful pencils/pencil/pencils	21	0.70		14	sink	9.00	0.30
42	bins/box/boxes/crates	15	0.50		16	feather/feathers/feather painting/feather picture/picture	9.00	0.30
37	oil/olive oil	14	0.47		37	bell pepper/pepper/peppers	8.00	0.27
18	sink	14	0.47		18	kitchen	7.00	0.23
13	keys	13	0.43		42	shelf/shelves/shelving	7.00	0.23

Note. *For Scene 16, there were two responses that were equally most frequent: *desk* and *color pencils/colored pencils/colorful pencils/pencil/pencils*.

†For Scene 13 and Scene 39, there were two first responses that were equally most frequent: *cactus* and *hat* and *banana/bananas* and *juice/orange juice*.

Appendix E

Proportion of data represented the most frequently occurring responses

Sorted by Proportion of Total Responses Represented by the Most Frequently Occurring Response				Sorted by Proportion of Total Responses Represented by the Three Most Frequently Occurring Response		
Scene	Proportion of Total Responses Represented by the Most Frequently Occurring Response: *M* = 0.17, *SD* = 0.04	Proportion of Total Responses Represented by the Three Most Frequently Occurring Responses: *M* = 0.42, *SD* = 0.08		Scene	Proportion of Total Responses Represented by the Most Frequently Occurring Response: *M* = 0.17, *SD* = 0.04	Proportion of Total Responses Represented by the Three Most Frequently Occurring Responses: *M* = 0.42, *SD* = 0.08
12	0.23	0.47		39	0.21	0.54
1	0.22	0.5		11	0.18	0.52
14	0.22	0.49		38	0.2	0.52
17	0.22	0.41		1	0.22	0.5
39	0.21	0.54		43	0.2	0.49
45	0.21	0.46		14	0.22	0.49
38	0.2	0.52		10	0.18	0.48
43	0.2	0.49		12	0.23	0.47
10	0.18	0.48		45	0.21	0.46
11	0.18	0.52		4	0.18	0.45
4	0.18	0.45		40	0.18	0.45
40	0.18	0.45		15	0.16	0.42
15	0.16	0.42		17	0.22	0.41
16	0.16	0.4		16	0.16	0.4
41	0.15	0.34		44	0.15	0.35
44	0.15	0.35		42	0.13	0.34
42	0.13	0.34		41	0.15	0.34
37	0.11	0.3		37	0.11	0.3
13	0.1	0.29		13	0.1	0.29
18	0.1	0.27		18	0.1	0.27

Appendix F

z-scores and *d'* calculated for DRM list recognition data

Participant	*z*-score for Hits	*z*-score for Lure False Alarms	*d'*
2	1.56	1.12	0.44
5	-1.28	-3.08	1.80
6	1.24	1.12	0.12
7	-1.28	0.42	-1.70
8	-0.65	1.12	-1.77
13	-0.65	0.42	-1.07
14	0.92	1.12	-0.20
20	0.92	0.42	0.50
21	-0.02	-0.28	0.26
23	0.61	-0.98	1.59
25	-0.02	0.42	-0.44
26	0.29	-1.68	1.97
27	0.61	0.42	0.19
32	0.61	-0.98	1.59
33	0.61	1.12	-0.51
34	1.24	1.12	0.12
35	0.61	0.42	0.19
36	1.56	0.42	1.14
37	-1.60	0.42	-2.02
38	-0.65	-0.28	-0.37
41	-2.23	-0.98	-1.25
42	-0.34	-0.98	0.64
43	0.61	-0.28	0.89
44	-0.65	-0.98	0.33
39	-1.91	-0.28	-1.63
45	-0.65	-0.98	0.33
46	0.29	0.42	-0.13
47	0.29	0.42	-0.13
48	0.29	1.12	-0.83
49	-0.34	-0.28	-0.06