Hyperbaric Oxygen Therapy’s Effects on Cognition and Narrative Discourse in Severe TBI: A Case Study

Ashley I. Anderson
University of Colorado Boulder, anderandersen@gmail.com

Follow this and additional works at: http://scholar.colorado.edu/slhs_gradetds

Part of the Cognitive Psychology Commons, Health Services Administration Commons, and the Speech and Rhetorical Studies Commons

Recommended Citation
Hyperbaric Oxygen Therapy’s Effects on Cognition and Narrative Discourse in Severe TBI: A Case Study

Ashley I. Andersen

University of Colorado Boulder

A thesis submitted to the Faculty of the Graduate School of the University of Colorado Boulder in partial fulfillment of the requirement for the degree of Master’s in Speech-Language Pathology Department of Speech, Language, & Hearing Sciences 2014
This thesis entitled:
Hyperbaric Oxygen Therapy’s Effects on Cognition and Narrative Discourse in Severe TBI: A Case Study
written by Ashley Andersen
has been approved for the Department of Speech, Language, & Hearing Sciences

__________________________________________
(Brenda Schick)

__________________________________________
(Gail Ramsberger)

Date ___ April 8, 2014 _______

The final copy of this thesis has been examined by the signatories, and we Find that both the content and the form meet acceptable presentation standards Of scholarly work in the above mentioned discipline.

IRB protocol # 13-0264
Discourse deficits following traumatic brain injury (TBI) have been found to negatively impact social reintegration and quality of life and are connected to underlying impairments in cognition. Hyperbaric oxygen therapy (HBOT) has been previously explored as a treatment for cognition in the TBI population, resulting in mixed outcomes. The present study examined the effect of HBOT on the cognitive and narrative discourse performance of an individual with chronic severe TBI. Multiple measurements of general cognition, receptive vocabulary, and discourse performance in the form of narrative storytelling were taken before and after HBOT. Hypotheses predicted that cognition would improve as a result of HBOT and facilitate enhanced narrative discourse performance; receptive vocabulary measures were not expected to improve on account of the participant’s lack of deficits in this area. Narratives were divided into T-units and assessed for organization, efficiency, and thoroughness. One-tailed t-tests indicated significant improvement in overall cognition but only in one aspect of the three narrative discourse measures. Further examination revealed that only improvements in attention explained the gains in cognition. For this study’s subject, HBOT likely increased attention but did not improve other areas of cognition measured, nor narrative discourse. Further studies incorporating more subjects and long-term outcome measurements are necessary for supporting these results and exploring HBOT’s effect on additional cognitive components and narrative discourse performance in a larger population of individuals with chronic, severe TBI.
Acknowledgements

The author would like to thank her advisor Professor Kathryn Hardin for bringing this research opportunity to her attention; her advisors Dr. Brenda Schick, Dr. Gail Ramsberger, and Professor Kathryn Hardin for their support throughout the process; Paula Messamer for her extensive assistance and encouragement; and my fellow graduate students Caroline Bowles, Terri Cusick, Molly Hines, Noelle Mitchell, and Kerri Mullen for generously volunteering their time to help with test administration and/or data analysis.
CONTENTS

CHAPTER

I. INTRODUCTION ........................................................................................................ 1

   Hyperbaric Oxygen Therapy ................................................................................ 3

   Present Study ........................................................................................................ 6

Method ......................................................................................................................... 6

   Participant: Case History ...................................................................................... 6

Measures ...................................................................................................................... 8

Procedures ................................................................................................................. 11

Results ....................................................................................................................... 13

Reliability ................................................................................................................... 13

Data Analysis ............................................................................................................. 15

Discussion .................................................................................................................. 20

   Cognition ................................................................................................................. 21

   Narrative Discourse ............................................................................................... 23

   Clinical Implications ............................................................................................. 25

REFERENCES .............................................................................................................. 27

APPENDIX

   A. RAW DATA .......................................................................................................... 32

   B. SAMPLE T-UNIT TRANSCRIPT ANALYSIS ..................................................... 34
TABLES

Table

1. Narrative Discourse Measurements—Raw Data and Statistical Comparison .............................................................. 19
FIGURES

1. PPVT-IV Scores................................................................. 16
2. RBANS Total Percentile Scores........................................... 17
3. RBANS Average Index Scores............................................. 18
4. Narrative Discourse Scores............................................... 20
INTRODUCTION

For individuals who have sustained a traumatic brain injury (TBI), reintegration into society presents a significant challenge. Quality of life may decline due to isolation, and more limited, less supportive social networks have been commonly reported among those recovering from a brain injury (Johnson & Davis, 1998; McCabe et al., 2007). General communication impairments among this population include “word finding difficulties, excessive talkativeness, difficulty staying on topic, difficulties thinking of questions to sustain a topic, tactlessness, repetitiveness, and difficulties keeping track of topics in group situations” (Togher, McDonald, Code, & Grant, 2004; p. 314). Such communication deficits can lead to problems forming and maintaining social relationships (Coelho, Liles, & Duffy, 1991; Coelho et al., 2013; Galski, Tompkins, & Johnston, 1998) and have also been linked to poorer quality of life and higher depression measures (Galski, Tompkins, & Johnston, 1998). Youse and Coelho (2009) stated that conversation partners often must take a more proactive role in facilitating the verbal interaction to compensate for deficits found in individuals with TBI. Specifically, individuals with TBI have demonstrated difficulty with staying on topic, providing sufficient information to meet the conversation partner’s needs, conveying information efficiently, and interrupting their conversational partner to an inappropriate extent, among other factors (Davis & Coelho, 2004; Galski et al., 1998; Marini et al., 2011; McDonald, 1993; Youse & Coelho, 2009). As a result, conversation partners have reported less satisfaction in communicative exchanges with individuals with TBI (Bond & Godfrey, 2000; Youse & Coelho, 2009).

Discourse analysis has been used to measure the subtle communication deficits that result from brain injury because it involves an interaction between language and cognition (Coelho,
After reviewing the literature on narrative discourse analysis in TBI, Coelho et al. (2013) suggested assessing the following aspects of narrative discourse: productivity, efficiency, content accuracy and organization, story grammar, and coherence. Coelho (2002) argues for the clinical value of story narratives as a specific type of discourse elicitation because they require the speaker to organize language in a manner that transmits the logical, temporal, and causal relationships between people and events in a story. Story narratives rely particularly on such cognitive processes as attention, processing speed, memory, and executive functioning, which are some of the frequently reported sequelae resulting from TBI (Youse & Coelho, 2005; Sohlberg & Mateer, 2001). Executive function impairments have been directly correlated to shortcomings in discourse production (Coelho, Liles, & Duffy, 1995; Coelho, 2002; Coelho et al., 2013; Lê, Coelho, Mozeiko, Krueger, & Grafman, 2012; Rousseaux, Vérighteaux, & Kozlowski, 2010; Ylvisaker & Szekeres, 1989), but in some studies no significant correlations existed (Marini et al., 2011). Similar mixed results have been documented with correlations between memory and aspects of narrative discourse. In particular, Youse and Coelho found a significant correlation between immediate memory and narrative discourse measurements but not working memory (2005). Lê et al. (2012) found that certain measurements of narrative discourse correlated moderately with working memory and moderately high with immediate declarative memory. Different methods of measuring narrative discourse could have contributed to the opposing correlational outcomes of these studies. For example, Coelho et al. (2002) correlated executive function tasks with the narrative discourse measurements of sentence production, cohesive adequacy, and story grammar. On the other hand, Marini et al. (2011) correlated executive function tasks to speech rate, cohesive errors, global coherence errors, lexical
informativeness, and ratio of thematic density. Both the matters of type of measurement and method used in the study affect comparisons of outcomes.

As many of the aforementioned studies have affirmed that cognition underpins narrative discourse, it can be hypothesized that treatments to improve cognition may enhance narrative discourse skills. Lê, Mozeiko, and Coelho (2011) suggest that, since cognitive processes play such a significant role in discourse, treating cognition may be more beneficial than targeting discourse and pragmatics in therapy. However, few studies have examined the interrelationship between cognition and narrative discourse as a result of intervention. Cannizzaro and Coelho (2002) demonstrated this in their pilot study of story grammar treatment with an individual with TBI. Although the thoroughness of the individual’s story narratives improved immediately after treatment, these gains were not maintained. Youse and Coelho (2009) tested the effect of attention training alone compared to a combination attention/communication treatment on the conversational discourse performance of two individuals with TBI. No gains were observed in either case; however, poor motivation of the participants was observed to contribute to these outcomes.

**Hyperbaric Oxygen Therapy**

One emerging medical intervention that has been used to treat cognitive and other deficits prevalent in TBI is hyperbaric oxygen therapy (HBOT). HBOT is the application of concentrated oxygen at atmospheric pressures (ATMs) that exceed normal pressure levels. Patients undergoing HBOT are contained in an air-sealed environment while administered high concentrations of oxygen ranging from 94% to 100% at 1 to 2 ATMS across numerous sessions (Hardy et al., 2007). Normal air pressure usually contains 21% oxygen. HBOT is generally
administered in 20 to 60 treatment sessions that range from 30 to 90 minutes long, but the exact protocol for HBOT dosing is still a matter of controversy (Masel, 2011).

The hyper-oxygenated, hyperbaric environment is thought to increase oxygen perfusion in the blood plasma to the extent that it can reach hypoxic brain tissue (Golden, 2006; Hardy et al., 2007). With TBI, HBOT theoretically functions in two different ways, although the exact mechanisms are still a matter of debate. In the acute stage, HBOT may prevent the sequence of events that alter metabolism in brain tissue (Masel, 2011). During the postacute stage (six or more months after injury), HBOT is thought to bring oxygen to and thus reactivate damaged, idle neurons that would otherwise remain dormant long-term (McDonagh, 2004). In this manner, HBOT is theorized to help neuronal functioning in general, including sensory, motor, and cognitive processes.

To date, there is scant research exploring the impact of HBOT on discourse skills in individuals with TBI. Previous studies assessing HBOT’s effect on individuals with chronic traumatic brain injuries ranging from mild to severe have documented cognitive and language gains (Boussi-Gross et al., 2013; Golden, 2006; Harch et al., 2012; Hardy, 2007; and Wright, 2009). These studies’ results indicated improvements in these specific cognitive processes as a result of HBOT: attention, processing speed, immediate and delayed memory, working memory, executive function, and verbal fluency. Only one of these studies (Hardy, 2007) included a subject with chronic severe TBI, which matches the injury severity level sustained by the subject of the present study. The other studies involved subjects with chronic mild or moderate brain injury, or did not specify the level of severity (Golden, 2006). Hardy’s (2007) case study on neurocognitive functioning in severe TBI found improvements in attention, working memory, and receptive and expressive language after HBOT. However, Hardy’s research participant
presented with severe mixed aphasia pre-treatment, whereas many individuals with TBI do not exhibit extensive deficits in receptive or expressive language.

It is important to consider that several of these studies’ methodologies have been criticized. For example, Wortzel et al. (2012) argued that the majority of Harch et al.’s (2012) subject pool had concomitant psychosocial disorders that detracted from its true representation of the mild TBI population, the majority of which experience a favorable prognosis. Thus, these psychiatric disorders could have been the primary contributors to the persistence of symptoms rather than their remote mild TBIs. In addition, Wortzel et al. asserted that the design of Harch et al.’s study did not control enough to negate a placebo effect, and that the gains observed post-HBOT were not substantial enough to draw them out of the clinical range of their pre-HBOT psychiatric illnesses. Wolf et al. (2012) also critiqued designs of previous studies of HBOT among the TBI population, pointing to a lack of randomization, blinding, and control for a placebo response.

Other studies have not found HBOT to be an effective treatment when compared to a control treatment. One study observed no significant difference in cognitive gains as a result of HBOT between a mild TBI treatment group and a sham control group (Wolf et al., 2012). In this case, a sham control group comprised of individuals without a history of TBI was administered room air at 1.3 ATMS while the treatment group received 100% oxygen at 2.4 ATMs. However, Boussi-Gross et al. (2013) question the placebo element of the study, claiming that even the sham control group received increased oxygenation by receiving the room air at a higher than normal pressure. Cifu et al. (2014) attempted to control for this possibility by decreasing the amount oxygen received at 2 ATAS (atmospheres absolute; equivalent to ATMS when above water) to below room air concentrations (10.5%). The actual treatment groups of Cifu’s study,
comprised of military personnel who had sustained a mild TBI within three years and experienced persistent post-concussion symptoms for at least 3 months, received either 75% or 100% oxygen at 2 ATMS. Results indicated no significant difference between control and treatment groups on the outcome measures. Thus, overall the literature on HBOT for subacute TBI is inconclusive regarding its benefits.

**Present Study**

The present study examined the effect of 40 sessions of HBOT on cognitive and narrative discourse measures of an individual with chronic severe TBI. It was predicted that HBOT would improve measures of neurocognition and narrative discourse but not receptive vocabulary. During clinical interactions with the research participant of this study before HBOT treatment was implemented, the first author observed deficits in the client’s narrative efficiency, content accuracy and organization, and coherence in conversation. Specifically, the client’s narratives were difficult to follow and comprehend due to the presence of maze words, off-topic instances, lack of cohesive adequacy (i.e., specificity and clearness of reference), and inaccurate and/or insufficient information. Thus, these were some of the aspects of narrative discourse measured in this study. The primary objectives of the study were to see if and to what extent HBOT improves cognitive processes and narrative discourse performance in an individual with chronic severe traumatic brain injury.

**Method**

**Participant: Case History**

The single subject of this study, AB, is a 20-year-old male who sustained a severe TBI in a pedestrian-motor vehicle accident in 2011. AB underwent an initial craniectomy to evacuate a large subdural hematoma. He later received a cranioplasty that was complicated by
hydrocephalus; a VP shunt was placed to relieve the pressure. He received acute rehabilitation care for five months and then was moved to a long-term outpatient rehabilitation care facility for the next seven months. At 1.5 years post-injury, AB’s physiatrist noted that he had mild hemiataxia, mild dysarthria, difficulty at times with sustained attention to conversation, and decreased psychosocial functioning when fatigued.

AB participated in speech-language therapy at the University of Colorado Boulder’s Speech, Language, & Hearing Clinic for one semester prior to receiving HBOT. Clinical observation noted deficits in memory, attention, executive function, off-topic responses in conversation, impaired insight, and some moments of confabulation. Standardized cognitive measurements from January 2013 (1.5 years post-injury) on the Woodcock-Johnson III Tests of Cognitive Abilities and Tests of Achievement (Woodcock, McGrew, & Mather, 2001a; Woodcock, McGrew, & Mather, 2001b) revealed processing speed cluster scores in the 1st percentile, immediate memory (story recall task) scores in the 60th percentile, and delayed memory (also a story recall task) scores in the 13th percentile. AB’s performance on the story recall task corresponded with a discrepancy percentile rank of 0.2%, meaning delayed memory was below expected recall when considering his immediate recall performance. On testing three months later (one year, eight months post-injury), AB’s processing speed cluster score had not improved significantly (3rd percentile), and his immediate memory score actually decreased significantly (11th percentile). In comparison to this immediate recall score, his delayed memory had jumped to the 58th percentile, demonstrating a discrepancy percentile rank of 99.9% that was significantly above expected performance. As is evident by these test scores, AB demonstrated variability in his performance on these tasks.
Measures

**Receptive vocabulary.** AB was administered the *Peabody Picture Vocabulary Test, Fourth Edition* (PPVT-4; Dunn & Dunn, 2007). This assessment, which measures auditory receptive vocabulary, was not expected to improve from HBOT treatment.

**Cognitive.** The participant completed the *Repeatable Battery for the Assessment of Neuropsychological Status* (RBANS; Randolph, 2012) for measures of general cognitive functioning. The RBANS contains subtests that measure skills of immediate memory (list learning and story memory), visuospatial (figure copy and line orientation), language (expressive vocabulary and semantic fluency), attention (focused and alternating attention), and delayed memory (list learning recall and recognition, story memory recall, and figure recall). The RBANS has been found to have moderate to strong clinical validity and reliability as a screening tool for moderate to severe TBI and high convergent validity for its subtests (McKay, Casey, Wertheimer, & Fichtenberg, 2006). It was also chosen based on its short duration and four equivalent forms. Since it only takes approximately 20 to 30 minutes to complete, the RBANS could be used as a cognitive screener combined with other testing procedures without exceeding three hours of total testing time and thereby reduce the possibility of inducing fatigue in the participant. Research indicates that individuals who have experienced a brain injury are more susceptible to fatigue than the general population; approximately 73% of the TBI population complains of fatigue up to five years after their head injury (Cantor et al., 2008; Ziino & Ponsford, 2006).

**Narrative discourse task.** To obtain narrative discourse measurements, AB provided oral story retells of four *I Love Lucy* episodes. The *I Love Lucy* story tell task was originally developed and implemented by Ramsberger and Rende (2002) to assess the conversational
transactional success of persons with aphasia. Ramsberger and Rende’s method of utilizing the *I Love Lucy* episodes was not followed in this study; rather, only the stimuli and analysis of the episode stories were used. It is important to note that Ramsberger and Rende identified two of the episode stimuli as more complex and the other two as more simple. Consequently, the authors calculated coefficients of equivalence between the complex and two less complex episodes on the composite scores of the measures taken. The results indicated moderately high coefficients of equivalence. In the present study, the discrepancy in complexity of stimuli was balanced by coupling one simple episode with one complex episode at both pre and post-treatment testing sessions.

The episodes provided stimuli for four different narratives, thereby minimizing the influence of a practice effect. Narratives were transcribed verbatim and divided into T-units for analysis. The T-unit, derived from Hunt (1985), has been employed as a narrative discourse measurement in numerous studies of the TBI population (Body & Perkins, 2004; Cannizzaro & Coelho, 2002; Coelho et al., 2013; Coelho et al., 2003; Jones & Turkstra, 2011; Lê et al., 2011; Lê et al., 2012). A T-unit compromises an independent clause and any of its associated dependent clauses. After the narratives transcripts were assigned into T-units, they were analyzed for the following measures:

1. **Maze words.** According to Loban (1976), mazes consist of “a series of words (or initial parts of words), or unattached fragments which do not constitute a communication unit and are not necessary to the communication unit” (p. 10). They include repetitions of words, word phrases, sounds, or syllables (“he just passes her passes her by”); revisions (“they get jobs as candy in the candy factory”); filled pauses (“he pulls out his dollar and um checks it”); and false starts (“and you guys just go I mean and this initially the guys were
sitting there”). Maze words are found in the oral speech of all individuals (deJoy & Gregory, 1985; Starkweather, 1987), but use of excessive maze words can detract from the flow of language and impair comprehension on the part of the listener. Thus, measures of maze words in speech can be indicative of communicative efficiency.

2. **Cohesive adequacy.** Cohesion can be considered a measure of discourse organization at a macrolinguistic level (Marini et al., 2011) and reflects a measure of clear versus vague or erroneous cohesive ties given in a narrative. Inadequate cohesive ties fail to establish solid referential information, thereby invoking more confusion for the listener.

Cohesive adequacy was assessed using procedures outlined by Coelho et al. (2013) and established in earlier publications (Liles, 1985; Liles & Coelho, 1998). All cohesive markers were first identified. As defined by Liles (1985), cohesive markers included references such as pronouns (“they,” “his dollar”); noun phrases (“the other girl”), demonstratives (“the wife,” “they cannot do that”); and conjunctions (because,” “and then,” “in which”). Liles further limits an item as cohesive “only when it cues the listener that the information is recoverable outside the sentence” (p. 133).

Each identified cohesive marker was further attributed to the category of complete, incomplete, or erroneous. Ties were marked complete if they clearly connected to other information elsewhere in the narrative; incomplete if no referring information for that cohesive element was found in the narrative; and erroneous if the tie indicated a vague or inaccurate referent. Maze words were excluded from being counted as cohesive markers.

3. **Story completeness.** To assess the thoroughness and accuracy of the storyline of each retell of *I Love Lucy*, a measurement of story completeness, as labeled by Lé, Coelho, Mozeiko, and Grafman (2011), was obtained. In both of these studies, narratives of individuals with TBI
were compared to those of a control group. Two steps were followed to find a story completeness measure. First, the authors determined the main ideas that were present in over 80% of the narratives of the non-brain-injured control group. Next, a completeness score was given based on the number of these critical story components present in the narratives of the individuals with TBI.

For the purposes of this study, which only looked at the narratives of one individual, the method of Lê et al. (2011) was not followed. Instead, the main ideas for each *I Love Lucy* episode determined by the four judges of the Ramsberger and Rende (2002) study were used as the critical story component criteria for judging the completeness score. These main ideas were formulated based on Hedberg and Westby’s (1993) outline of story grammar analysis. Each narrative retell was evaluated for the percentage of these main ideas it contained.

**Procedures**

At one year, 10 months post-injury, AB commenced HBOT under the supervision of his physiatrist in Louisville, CO. AB had forty individual 90-minute sessions of HBOT five days a week over the course of eight weeks. Each session occurred at an atmospheric pressure of 1.5 ATMs and included 60 minutes of actual pressurized oxygen intake; the initial 20 minutes comprised a slow compression procedure while the final 10 minutes were devoted to decompression.

This study employed a within subject pre-post treatment design. Originally, pre-, mid-, and post-treatment testing sessions occurred: four baselines prior to HBOT treatment, two days of testing between the 20th and 21st sessions of HBOT, and two days of testing post-treatment. Unfortunately, components of the intended mid-treatment testing sessions were not completed, compromising the validity of these measurements to the final results. Consequently, only the
baseline and post-treatment data were available for analysis. Four baseline measurements were taken prior to the participant initiating HBOT. The four measurements were repeated nine days after completing the 40th session of HBOT.

The participant was administered the RBANS and PPVT-4 on each testing day (four times pre-HBOT, four times post-HBOT). Overall, AB was tested on each of the four equivalent forms of the RBANS twice and each of the two equivalent PPVT-4 forms four times total.

Narrative elicitation tasks occurred only once at each pre- and post-treatment session. During pre-treatment testing, AB viewed the following two episodes on the same day: *Lucy is Enceinte* (Oppenheimer, Pugh, & Carroll, 1989a) and *Pioneer Women* (Oppenheimer, Pugh, & Carroll, 1990). After a two-hour delay, during which AB completed the RBANS and PPVT, he was instructed to retell the story to a confidant (in this case, a graduate student) who had not previously seen these episodes. The confidant was instructed to offer only comments that would not influence the course of the retell, such as “Oh, I see” or “so that’s what happened.” This procedure was again repeated post-treatment with the two remaining episodes: *Job Switching* (Oppenheimer, Pugh, & Carroll, 1989b) and *Bonus Bucks* (Oppenheimer, Pugh, & Carroll, 1991). Episodes were delegated for testing in such a manner as to allow for one complex episode to be shown along with one simpler episode. All story retells were videotaped and later transcribed and analyzed by the author and two other trained graduate student judges.

All cognitive, language, and narrative discourse testing was performed by two graduate students. For baseline measurements, one graduate student administered the first two testing sessions and the other the second two. During post-treatment testing, they each administered two testing sessions but in reverse order. The same graduate student administered the narrative
discourse assessment in one pre-treatment session and one post-treatment session to keep procedures consistent. Equivalent forms were presented in a mixed order.

**Results**

**Reliability**

All PPVT tests were scored by the author and one other graduate student test administrator. Agreement between the judges was 100%. Each RBANS test was scored by the first author and the graduate student who administered that particular form of the RBANS. Thus, the only judge blind to the stage of treatment (pre vs. post) was the first author. Inter-rater reliability for the 16 total score items resulted in a Cronbach’s alpha of 0.9814, demonstrating strong reliability. At the subtest level, inter-rater reliability was only calculated for the visuospatial index, as the scoring criteria for this index is fairly subjective. The author of the RBANS (Randolph, 2012) assessed inter-rater reliability for this subtest at 0.85. For this study, its inter-rater reliability did not meet that criteria (16 items; $\alpha = 0.83$). Thus, it is important to view this study’s results of the RBANS with some caution due to the substandard inter-rater reliability of scoring on this particular subtest, which could have occurred as a result of the subjectivity of the scoring procedure or possibly as a result of bias in scoring on the part of the stage-informed judge.

**Narrative transcripts.** Each story retell was transcribed verbatim by the first author. Four weeks later the first author transcribed the story retells again. To determine intra-rater reliability, words that differed between transcripts, as well as any additional words found in one transcript but not in the other, were counted as disagreements. The total number of disagreements was then divided by the total number of words of all four retells, taken from the smaller of the two transcripts. Agreement between the transcripts was at 98%. A graduate student who was an
unfamiliar listener then helped to resolve the discrepancies between transcripts by viewing the episode retells and deciding between the word discrepancies in the transcripts.

The final episode retell transcripts were then divided into T-units by the first author and two trained graduate students who were blind to which stage of the study (pre/post) the episode retells came from. Inter-rater reliability for T-unit assignment among the transcripts across the three judges was determined using a Cronbach’s alpha coefficient and demonstrated moderate reliability (3 items; $\alpha = 0.774$). Different T-unit allocations were resolved via discussion to agree upon a finalized T-unit allocation transcript for each episode (see Appendix B for an example of a finalized T-unit allocation transcript). Intra- and inter-rater reliability for the three analyses taken from the finalized T-unit transcripts were as follows:

1. Maze words. The first author counted the number of maze words for each *I Love Lucy* episode retell transcript at two separate occasions. The total number of maze words for each transcript was divided by its total number of T-units, providing an index of maze words per T-unit. Intra-rater reliability for the maze words per T-unit index, based off of the first author’s scores from the two different occasions, was determined using a two-tailed paired $t$ test; the differences between scores were not significant [$t(6) = 1.87, p = 0.16$]. The two trained graduate student judges also counted the number of mazes per T-unit for each transcript at one occasion. Their scores were compared to the average scores calculated by the first author. Inter-scorer agreement was found to have high reliability (12 items; $\alpha = 0.984$). All discrepancies were resolved through discussion. The final agreed upon maze words per T-unit index increased from the individual judges’ initial counts due to consensus on inclusion of certain phrases as maze words (e.g., “I mean” and “basically”) that the research participant used more as fillers than clarifiers.
2. Cohesive adequacy. The first author analyzed each transcript for the total number of cohesive items on two separate occasions to obtain a calculation of intra-rater reliability (8 items; $\alpha = 0.98$). Inter-rater reliability between the three judges for the total number of cohesive items in each transcript resulted in very high reliability as well (12 items; $\alpha = 0.999$). All discrepancies were resolved through discussion. Next, all judges coded the cohesive markers as complete, incomplete, or erroneous on two separate occasions. The number of complete cohesive ties out of the number of total cohesive ties was used to provide an index of cohesive adequacy. Comparison of the judges’ cohesive adequacy index scores demonstrated low reliability (12 items; $\alpha = 0.57$).

Because of the discrepancy between these scores, all differences were resolved via discussion to agree upon a final cohesive adequacy index score for each transcript. The pre-discussion averages for the cohesive adequacy index across the three judges for the four episode retells were as follows: 41.96%, 57.4%, 51.59%, and 58.64%. Corresponding post-discussion values concluded as 44.19%, 61.40%, 50.22%, and 69.19%. As these numbers demonstrate, three of the cohesive adequacy indexes increased upon discussion and one decreased.

3. Story completeness. The first author evaluated all narratives for story completeness twice and had an intra-rater reliability of 100%. All three judges’ completeness scores were compared and demonstrated high inter-rater reliability (12 items; $\alpha = 0.953$).

Data Analysis

All statistical calculations were obtained using IBM SPSS Statistics 22 (2013) software. Measurements taken at baseline were compared to those observed after the participant completed HBOT. A significance level of $p < 0.05$ was set for all analyses performed.
Receptive Vocabulary. Because each PPVT test form (A and B) was administered two times during baseline testing and an additional two times during post-treatment testing, average Form A scores and Form B scores from pre-testing were compared to average Form A scores and Form B scores from post-HBOT observations. A paired samples two-tailed \( t \)-test was used due to the expectation that HBOT would not improve receptive vocabulary skills. As predicted, these scores did not differ significantly after treatment \[ t(3) = 1.26, \ p = 0.4682 \] (see Table A1 in Appendix A for raw data).

Figure 1. PPVT percentile scores across all forms at baseline and post-treatment.

Cognitive. Total score. It was hypothesized that AB’s cognitive measurements would increase significantly as a result of HBOT. RBANS scores before treatment were compared with the scores on the same corresponding form after treatment using a one-tailed paired samples \( t \)-test. Thus, the baseline Form A results were paired with the post-HBOT Form A results, etc. This calculation indicated a significant increase in the RBANS total score from pre-treatment \( (M = 14, SD = 7) \) to post-treatment \( (M = 42.38, SD = 27.78) \), \( t(3) = 2.37, p = 0.0495 \). AB’s overall
cognition as measured by the RBANS did improve after HBOT (see Table A2 in Appendix A for raw data); however, post-HBOT scores demonstrated large variability and warranted further examination at the index level.

*Figure 2.* RBANS total percentile scores across all forms at baseline and post-treatment.

<table>
<thead>
<tr>
<th>Form</th>
<th>Pre-HBOT</th>
<th>Post-HBOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
<tr>
<td>B</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
<tr>
<td>C</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
<tr>
<td>D</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
</tbody>
</table>

**Subtests.** All pre-HBOT subtest scores were compared to post-HBOT subtest scores using one-tailed paired *t* tests. Only the attention index demonstrated a statistically significant increase for pre-treatment (*M* = 20.5, *SD* = 7.68) to post-HBOT (*M* = 56, *SD* = 17.15), *t*(3) = 3.65, *p* = 0.018. Pre-treatment averages for the remaining indexes were as follows: immediate memory (*M* = 23.75, *SD* = 12.58), visuospatial (*M* = 44, *SD* = 20.7), language (*M* = 33.75, *SD* = 15.97), and delayed memory (*M* = 7.56, *SD* = 12.98). Post-treatment, no significant gains were demonstrated for the indexes of immediate memory (*M* = 55.75, *SD* = 32.69), *t*(3) = 1.64, *p* = 0.10, visuospatial (*M* = 62.5, *SD* = 34.68), *t*(3) = 1.14, *p* = 0.169, language (*M* = 41.5, *SD* =
25.51), \( t(3) = 1.04, p = 0.187 \), and delayed memory (\( M = 20.75, SD = 21.98 \)), \( t(3) = 1.20, p = 0.158 \). As the raw data demonstrates, all the other subindex averages in addition to attention increased post-treatment, but their large standard deviation values inhibited achievement of these gains to statistically significant levels.

*Figure 3.* RBANS average subtest percentile scores across all forms at baseline and post-treatment.

---

**Discourse Analyses.** Separate statistical computations were performed for each of the three discourse analysis measures. All *t*-tests were one-tailed, as the hypotheses were specific with respect to the direction of change. It is important to note that level of complexity of the *I Love Lucy* episode retold was not factored into the analysis, as both pre and post-treatment testing sessions were balanced with one simple and one complex episode (see Table 1 for raw data).

1. *Maze words per T-unit.* Since two episodes were retold at each narrative elicitation
testing session, the maze words per T-unit values from the baseline testing session were compared with the post-treatment episode counts. Although these counts did improve (i.e., the number of maze words per T-unit decreased in the post-treatment episode retells), a one-tailed paired samples t-test revealed no statistically significant improvement \([t(1) = 1.36, p = 0.20]\).

2. **Cohesive adequacy index.** Pre-test cohesive adequacy index averages were compared to the corresponding post-test averages. Results indicated that the cohesive adequacy index measure significantly improved post-treatment \([t(1) = 7.85, p = 0.0401]\).

3. **Story completeness.** As above, story completeness measures were averaged for the two pre-test episode retells and two post-test retells and compared using a one-tailed paired samples t-test. Story completeness measures did not improve significantly after treatment \([t(1) = 0.77, p = 0.291]\).

*Table 1.* Statistical values of the comparison between pre- and post-treatment narrative discourse measurements of maze words per T-unit, cohesive adequacy index, and story completeness. One-tailed paired samples t-tests were used for all three measurement comparisons.

<table>
<thead>
<tr>
<th></th>
<th>Baseline Simple</th>
<th>Baseline Complex</th>
<th>Post-HBOT Simple</th>
<th>Post-HBOT Complex</th>
<th>Comparison of pre and post averages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maze words per T-unit</td>
<td>4.33</td>
<td>2.27</td>
<td>1.93</td>
<td>1.9</td>
<td>(p = 0.206)</td>
</tr>
<tr>
<td>Cohesive adequacy</td>
<td>61.40%</td>
<td>44.19%</td>
<td>69.19%</td>
<td>50.22%</td>
<td>(p = 0.041)</td>
</tr>
<tr>
<td>Story Completeness</td>
<td>88.89%</td>
<td>29.17%</td>
<td>83.33%</td>
<td>72.22%</td>
<td>(p = 0.291)</td>
</tr>
</tbody>
</table>
Figure 4. Narrative discourse measurements of maze words per T-unit, cohesive adequacy index, and story completeness. Maze words per T-units and story completeness measures did not show significant changes after treatment ($p > 0.05$), whereas cohesive adequacy improved significantly ($p < 0.05$).

**Narrative Discourse Scores**

<table>
<thead>
<tr>
<th>Maze words per T-unit</th>
<th>Cohesive Adequacy</th>
<th>Story Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Episode Retell</td>
<td>Complex Episode Retell</td>
<td></td>
</tr>
</tbody>
</table>

**Discussion**

This study investigated whether a series of 40 sessions of HBOT improved the cognitive and narrative performance of an individual with post-acute severe TBI. It was predicted that HBOT would increase general cognitive functioning and in turn facilitate increased measures of efficiency, thoroughness, and organization in storytelling.

Results indicated the benefits of HBOT included gains in general cognitive functioning, particularly in attention, and in one aspect of narrative discourse measured—cohesive adequacy. Receptive vocabulary test measures did not improve significantly after treatment, as was predicted based on the participant’s demonstrated lack of impairment in this area and scant research evidencing HBOT as an effective treatment for this linguistic capacity (Hardy et al., 2007).
Cognition

Cognitive gains made were clinically significant for the participant of this study. Before receiving treatment, AB’s average RBANS total scores were below normal limits (14th percentile). After treatment, his scores averaged well within normal limits and very close to the mean, at 42.375%. However, as indicated by individual index results, this gain was primarily restricted to attention. Closer scrutiny of the averages across each individual subindex comprising the RBANS revealed no significant increases across the indexes of immediate memory, delayed memory, attention, language, and visuospatial skills. Although some of the mean scores of these indexes were much higher post-HBOT (e.g., immediate memory, visuospatial, and delayed memory), the variability of scores was quite extensive.

The notable increases in attention warrant further discussion. Despite evidence pointing toward the strong validity and reliability of the RBANS among the TBI population for its total score and the majority of subtests, studies have also found the RBANS attention subtest to have weak internal reliability among the TBI population (McKay et al., 2007). Namely, as McKay et al. argue, the attention index of the RBANS comprises one test sensitive to symptoms of brain injury (coding) and one that is not (digit span). AB’s performance on these individual attention subtests aligned with this assertion. AB’s scores on the coding task increased significantly \( t(3) = 9.13, p = 0.0014 \), but the digit span scores did not \( t(3) = 1.58, p = 0.106 \) (see Table A3 in Appendix A). Thus, AB’s improvement in the facet of attention impacted by TBI was extensive enough to reach clinically significant gains for the whole index despite the lack of change of the other subcomponent.

Although a review of theoretical models of attention are beyond the scope of this paper, many of them encompass the concepts of “sustaining attention over time (vigilance), capacity for
information, shifting attention, and screening out nontarget information” (Sohlberg & Mateer, 2001; p. 126). As these processes are recruited for more complex, higher level thinking such as executive function tasks like planning and problem-solving, gains in attention are theorized to contribute a more solid foundation for other aspects of cognition. Gains from specific attention training have been shown to generalize to improvements in such areas as memory, learning, and executive function tasks in some studies (Neumann, Ruff, & Baser, 1990; Sohlberg et al., 2000; Sturm et al., 1997), but in other studies generalization did not occur (Park & Ingles, 2001).

It is plausible to consider, albeit with caution, the possibility that the significant improvements in attention that occurred in this study’s participant facilitated gains elsewhere. This bottom-up direction of gains could have been diluted as the cognitive processes moved up the hierarchy from simple to more complex thinking, as attention became one of multiple processes recruited. In this sense, attentional gains surpassed those of visuospatial skills, immediate and delayed memory, and language because these latter abilities implicate other cognitive processes in addition to attention.

However, it is important to consider this study’s design limitations when interpreting these results. First, the RBANS may not have been a tool sensitive enough to accurately detect more subtle changes in the cognitive processes it measures. However, as stated previously, the RBANS was chosen for its brevity, multiple equivalent forms, preestablished reliability, validity, and convergent validity. Thus, it offered advantages over other standardized tests. Second, the possibility of a practice effect cannot be excluded (Schiavetti & Metz, 2006). Although 61 days passed in the interim between pre and post-testing, the participant retook all four equivalent forms of the RBANS after receiving treatment. He improved on each individual form administered in post-treatment testing. Furthermore, since the original design of this study aimed
for administering two forms of the RBANS midway through treatment (between the 20th and 21st sessions of HBOT), the participant was exposed to Forms A and B a total of three times. As stated earlier, these mid-treatment testing sessions were invalidated and chosen to be repeated again upon conclusion of treatment to balance the number of testing sessions.

In hindsight, this design flaw could have been remedied by administering equivalent forms A through C for three baseline measurements and then repeating those same forms for three post-treatment measurements. Then, equivalent Form D could have been administered at post-treatment testing as an additional control mechanism.

**Narrative Discourse**

The narrative discourse analysis revealed significant improvement in only one measurement area: cohesive adequacy index. This suggests that after treatment the research participant established more specific and clear pronoun and demonstrative references in his story narratives, indicating progress in his ability to remember previously stated information and to organize and produce new information accordingly. This increase in organization theoretically translates to a less confusing and easier-to-follow narrative for the listener.

It is important to remember, however, that the cohesive adequacy index measurement correlated with low inter-rater reliability. Since the cohesive adequacy index measure has pre-established reliability (Coelho et al., 2013), in the specific case of this study more thorough training of the judges may have curtailed the reliability error. More specific training examples including samples of narratives from individuals with severe brain injury would have also abetted further consistency in this scoring process.

However, because significant progress was not made for the categories of maze words per T-unit and story completeness, confusing elements were still present in the post-treatment
narratives despite the advancements made in cohesive adequacy. This demonstrates how unidimensional progress can easily get lost in the dynamic complexities of language. The count of maze words per T-unit did decrease pre versus post-treatment, but due to the participant’s variability in performance, this decrease was not significant. Perhaps if more narrative data samples had been collected at pre and post-treatment testing sessions, these improvements would have achieved significance. Story completeness measures, on the other hand, did not increase after HBOT, meaning the participant did not get better at providing his listener with the integral components of the storyline.

Taken together, the data does not suggest overall improvement in narrative performance. Further testing with a greater number of subjects and more reliable measures of narrative discourse is required to determine the effect of HBOT on narrative discourse skills. More numerous measures would also be an asset for analyzing the subtle complexities of narrative discourse.

A further point of discussion involves consideration of the narrative discourse results in light of the cognitive improvements observed. All three narrative discourse measures incorporate several cognitive processes, including attention, immediate and delayed memory, and executive function. Consequently, it is somewhat surprising that cohesive adequacy improved but story completeness and maze words per T-unit did not. Regarding the cohesive adequacy index, generalization of attentional improvements may have fostered the research participant’s ability to attend to the specificity and clearness of references made in the story narrative provided. But why did attentional improvements not generalize to the other two narrative discourse measures? One explanation may be that the participant simply attended to his referencing better than the other aspects of his storytelling post-treatment. A more likely explanation may point towards the
validity of the cohesive adequacy measure in this study. As stated previously, the inter-rater reliability value for this measure was low; it is possible that the cohesive adequacy measure, if scored by another set of judges, would not have improved significantly and aligned more consistently with the outcomes of the other two narrative discourse measurements.

**Clinical Implications**

All together, the data from this particular study suggest that HBOT may be a treatment option for improving attention in this study’s particular individual with chronic severe TBI but would not be advised as a solitary intervention for narrative discourse. However, long-term follow-up is necessary to determine if the improvements seen in this study maintain, as other HBOT research has found benefits to diminish over time until additional HBOT sessions are administered (Hardy et al., 2007).

When considering HBOT primarily as an attention remediation program, comparison to existing attention treatments is necessary. In their review of the evidence regarding direct attention training for establishing practice guidelines, the Academy of Neurologic Communication Disorders and Sciences (ANCDS; Sohlberg et al., 2003) concluded that the efficacy data reveals clear improvement on trained tasks but little generalization to untrained tasks or activities of daily living. Furthermore, the ANCDS emphasized that, due to the heterogeneity of the TBI population in the literature, it is critical to consider individual client characteristics and desired outcomes when choosing an attention training program and how to implement it in order to procure the most value for that client. In the case of this study’s research participant and in view of the results, HBOT presented as a favorable alternative to a computerized or clinician-facilitated attention program. As a full-time undergraduate student, repetitive drill training on tasks unrelated to class or other real-life material would have been an
undesirable drain on valuable cognitive resources for this study’s participant. On the other hand, HBOT, although time consuming, presented opportunity for review of class materials or for cognitive rest while physically receiving treatment. And even though generalization of gains was not firmly established, results were suggestive of some improvement in one component of narrative discourse. Of further importance, anecdotal statements from the research participant indicated self-awareness of improvements, especially in memory, after treatment. Individual life factors of other clients, including severity of injury, financial resources, and transportation and other life supports, among others, may preclude the cost-benefit ratio for HBOT over other interventions. Further research involving numerous participants is needed to more confidently recommend or not recommend HBOT as a treatment for attention specifically, cognition generally, or even constituents of narrative discourse to a broader scope of the TBI population or even to this study’s participant in regards to functional and long-term outcomes.
References


Lê, K., Coelho, C., Mozeiko, J., Krueger, F., and Grafman, J. (2012). Predicting story goodness


### Appendix A

Table A1

**PPVT-IV Raw Scores**

<table>
<thead>
<tr>
<th>Testing Session</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>Percentile Score</td>
<td>75%</td>
<td>73%</td>
<td>42%</td>
<td>61%</td>
<td>66%</td>
<td>75%</td>
<td>82%</td>
<td>70%</td>
</tr>
</tbody>
</table>

*Note. Each testing session score reflects the average of the two judges’ scores.*

Table A2

**RBANS Raw Scores**

<table>
<thead>
<tr>
<th>Testing Session</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form</td>
<td>A</td>
<td>B</td>
<td>D</td>
<td>C</td>
<td>D</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>RBANS Subtest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Memory</td>
<td>25%</td>
<td>41%</td>
<td>16%</td>
<td>13%</td>
<td>85%</td>
<td>30%</td>
<td>25%</td>
<td>83%</td>
</tr>
<tr>
<td>Visuospatial</td>
<td>60%</td>
<td>19%</td>
<td>62%</td>
<td>35%</td>
<td>92%</td>
<td>92%</td>
<td>25%</td>
<td>41%</td>
</tr>
<tr>
<td>Language</td>
<td>30%</td>
<td>39%</td>
<td>52%</td>
<td>14%</td>
<td>61%</td>
<td>18%</td>
<td>66%</td>
<td>21%</td>
</tr>
<tr>
<td>Attention</td>
<td>12%</td>
<td>27%</td>
<td>27%</td>
<td>16%</td>
<td>65%</td>
<td>35%</td>
<td>50%</td>
<td>74%</td>
</tr>
<tr>
<td>Delayed Memory</td>
<td>27%</td>
<td>1%</td>
<td>2%</td>
<td>.25%</td>
<td>48%</td>
<td>5%</td>
<td>1%</td>
<td>29%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>43%</td>
<td>11.5%</td>
<td>17.5%</td>
<td>5.5%</td>
<td>79.5%</td>
<td>26%</td>
<td>17%</td>
<td>47%</td>
</tr>
</tbody>
</table>

*Note. Each testing session score reflects the average of the two judges’ scores.*
### Table A3
Attention Index Subcomponent Raw Scores, Means, and Standard Deviations

<table>
<thead>
<tr>
<th></th>
<th>Coding</th>
<th></th>
<th>Digit Span</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-HBOT</td>
<td>Post-HBOT</td>
<td>Pre-HBOT</td>
<td>Post-HBOT</td>
</tr>
<tr>
<td>Form A</td>
<td>44</td>
<td>56</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Form B</td>
<td>39</td>
<td>49</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Form C</td>
<td>45</td>
<td>52</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Form D</td>
<td>46</td>
<td>55</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Mean</td>
<td>43.5</td>
<td>53</td>
<td>12</td>
<td>13.75</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3.11</td>
<td>3.16</td>
<td>1.83</td>
<td>1.26</td>
</tr>
</tbody>
</table>

*Note.* Each testing session score reflects the average of the two judges’ scores.
Appendix B
Sample T-unit Transcript Analysis

Example of the first 10 T-units from the Lucy is Enceinte episode retell transcript. Maze words are italicized. Cohesive items are marked in color according to the following key: complete (blue), incomplete (green), erroneous (red).

1. **Lucy** got pregnant

2. **And um she** wanted to tell **him**

3. **But** she it was like how she she wanted to make it into her like how she wanted to make it into her like how she wanted to live how she dreamed that she was going to tell **him**

4. **And she kept getting it she** kept getting interrupted by many people like **her phone her um the her** friends

5. **I mean just like all kinds**

6. **And** then **I mean she** tries and tries and tries **but** is still unable to **cause** each **thing** keeps interrupting **her** when she’s about to say **this or that kind of thing**

7. **It’s extremely difficult**

8. **So in the end um ricky** is singing

9. **And um he** actually gets done with **the song and then** gets a little card that says **um there’s a man and woman in the audience who have who are just who just found out today that they’re having a baby in which um then he goes then it go would would you please sing your song of you’re my you’re having a baby my baby**

10. **and um it’s like and then he like** he’s like ok **who who’s the happy couple cause I want them to come up here**

*Note.* An item was only marked cohesive if the referential information was not contained within the same T-unit according to Liles (1985). Cohesive items repeated within the same T-unit (e.g., “But . . . she wanted to live how she dreamed”). Some items (e.g., “and then” and “the happy couple”) considered to be phrases were only marked as one cohesive item.