INSIGHT DURING DEVELOPMENT, AND ITS STRUCTURAL CORRELATES

WEIPENG JIH

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

September, 2018

Abstract

We investigate whether adolescents and adults differ in their use of common cognitive processes in solving insight problems. We also investigate whether performance on insight problems is associated with brain structure, and whether these insight-structure associations are distinct or consistent across the two age groups. Common cognitive processes (operationalized by IQ scores) showed a positive trending correlation with insight (operationalized by accuracy in solving verbal riddles) in adults, but not in adolescents. However, these correlations were not significantly different. Thus, we failed to find a cognitive difference between adolescents and adults with regard to insight problem solving. Voxel based morphometry revealed that insight and gray matter volume are related in both age groups. Tract-based spatial statistics revealed that insight and fractional anisotropy values are related in adults. We could not determine whether insight-structure relationships are age-unique or age-consistent.

Table of Contents

List of Tables

List of Figures

Introduction

This study investigates insight from adolescence to early adulthood. Across this age range, it aims to examine behavioral measures and brain structures associated with insight.

Insight has previously been defined cognitively as "...any sudden comprehension, realization, or problem solution that involves reorganization of the elements of a person's mental representation of a stimulus, situation, or event to yield a nonobvious or nondominant interpretation" (Kounios & Beeman, 2014). Behaviorally, insight problem solving may be characterized by impasse, followed by fixation, followed by a disengaged incubation period, followed by the sudden obtaining of a problem solution often accompanied by a feeling of "aha" (Duncker & Lees, 1945).

The task that has been chosen to represent insight problems in this study is the verbal riddle. Riddles are a type of insight because the initial interpretation or mental representation of the riddle typically does not lead to comprehension. In order for comprehension to occur, a change in the riddle's interpretation needs to occur, after which the correct interpretation may seem almost trivial. Riddles are a type of *problem* because there is a gap between the initial incoherent interpretation of the riddle and the final coherent interpretation of the riddle. For example, one of the riddles to be included in this study, the "Professor Bumble" riddle, goes: " Professor Bumble, who is getting on in years is growing absent minded. On the way to a lecture one day he went through a red light and turned down a one way street in the wrong direction. A policeman observed the entire scene but did nothing about it. How could Professor Bumble get away with such behavior?" The initial interpretation is likely to be that Professor Bumble was driving in the wrong direction. This interpretation does not produce adequate comprehension of the riddle. However, once a shift in interpretation occurs in which Professor Bumble is walking, the riddle is suddenly comprehended. Behaviorally, riddles fit our aforementioned criteria of suddenness in insight. Support for this comes from Metcalfe (1986), which found that: participants, who successfully solved the riddle used in the study, indicated little change in self-reported progress prior to a sudden realization of the solution. Furthermore, in the neuropsychological literature, verbal riddles have been used to study insight (e.g., Luo & Niki, 2003; Mai, Luo, Wu, & Luo, 2004; Qiu et al., 2008a; Goel, Eimontaite, Goel, & Schindler, 2015).

If we accept riddles as a form of insight, what kind of a cognitive model can we specify for it? Here, the view will be adopted that riddles begin as normal verbal comprehension of sentence or multi-sentence discourse that is then interrupted and sidetracked by other processes. The crux of the work done in solving a riddle is in comprehending the riddle correctly. Once the riddle is comprehended correctly, the solution to the riddle is automatically inferred. Consistent with this view, Weisberg (1995) says that "riddles hinge on interpreting the language in the problem".

Cognitive Model of Riddle Insight

We define a mental representation as any abstraction of the external world within our internal world (the brain). Our task of building a comprehensible mental representation of a riddle begins with linguistic perception, where we categorize physical sounds as discrete phonemes (i.e., abstract linguistic categories of sounds based on pitch and formant information). Following linguistic perception, lexical retrieval for each word occurs, which according to the Cohort Model of word retrieval (Marslen-Wilson, 1990) involves a word's initial phonemes briefly cueing and activating a group (i.e., cohort) of several potential words simultaneously before accruing phonological or sentence context information suppresses all but one word. We then do this for each word we come across in a phrase or sentence.

Once we settle on a particular word, we retrieve its concept. Concepts are combined to form higher-level meaning units, known as propositions, via syntactic transformation rules. Concepts can also be combined to form another type of representation called the situational model. If propositions and situational models are determined to have local and global coherence (which we will revisit), then the riddle is comprehensible.

Before continuing, it is worth mentioning that so far we have described linguistic processing as proceeding in a single direction from units of sounds to units of meanings, in a bottom-up fashion. However, linguistic processing can also proceed in the other direction, where units of meaning influence perception of sounds, in a top-down process.

Let us now overview three levels of mental representation which are essential for solving riddles, starting with the notion of concepts. A concept is a mental representation which refers to some category of things in the world (Medin & Rips, 2005, p37). A concept can be directly activated by a word. Such concepts are called "lexical concepts." It is assumed that once a word activates a lexical concept, activation will spread outward to other first-order related concepts and then to second-order concepts and third-order concepts etcetera, with decreasingly smaller activation the further activation spreads. This description of concepts is consistent with the semantic network perspective adopted by Bock and Levelt (1994), and earlier by Collins and Loftus (1975) (figured below). Though research into exactly what a concept entails has historically been contentious (for a review see Medin & Rips, 2005), for practical reasons we will settle for our current description.

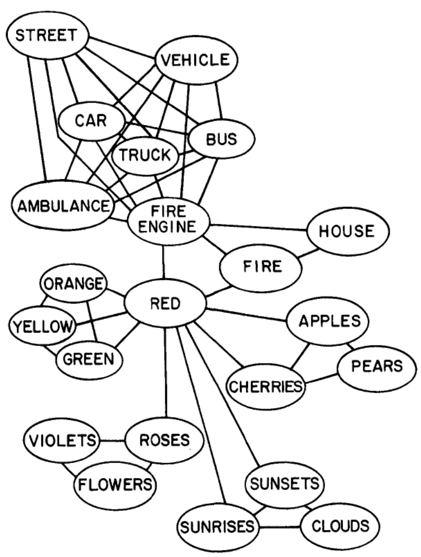


FIGURE 1. A schematic representation of concept relatedness in a stereotypical fragment of human memory (where a shorter line represents greater relatedness).

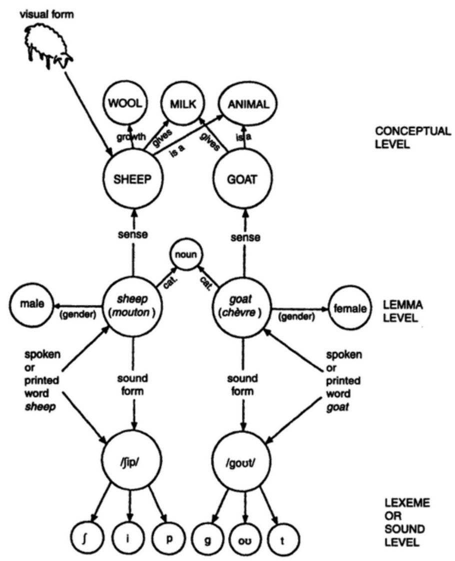


Figure 2 A part of the lexical network. Note that the arrows represent types of connections within the network, not the flow of information during production or comprehension.

Figure 1. (Left) The model of lexical access by Collins & Loftus (1975).
Figure 2. (Right) The model of lexical access by Bock & Levelt (1994).

The representational unit, proposition, which is built from concepts, can be viewed as the smallest unit of meaning which can carry truth value (Fletcher, 1994). Propositions are a higher level of representation than concepts; typically take the form of two or more concepts plus the relationship that binds them; and are embeddable within other propositions (Carroll, 2008, p154,168-170). An example of a proposition from the "Professor Bumble" riddle would be OBSERVED(POLICEMAN, SCENE). Furthermore, propositions can be abstracted to form summary or gist propositions, creating a propositional hierarchy with small detailed propositions at the bottom and overall summary propositions at the top (van Dijk & Kintsch, 1983 p191). Behavioral evidence for the reality of propositions coding for verbal material comes from Kintsch and Keenan (1973), who demonstrated that for sentences of the same word length, participants who were reading to remember took longer to read the sentences with more propositions than sentences with fewer propositions. Ratcliff and McKoon (1978) gave further evidence for the reality of propositions by demonstrating that words in a proposition were more primed (i.e., recognized faster) by another word in the same proposition than by a word in another proposition. This effect occurred even when the distance between test words and primes was controlled for.

Situational models (also known as mental models) are another level of text representation which riddles may be stored as. Like propositions, situational models are built up from the concepts activated by the riddle (some of which may not be explicit in the riddle, but rather, inferred). However, whereas propositions may be viewed as representations of statement meanings, situational models may be viewed as representations of actual events similar to the representation that may be formed from being in a situation (Fletcher, 1994). Evidence that situational models form another mental representation of text, distinct from propositions, comes from Bransford, Barclay, and Frank's (1972) study. In it, they found that participants were more likely to misidentify having heard test sentences that conveyed the same situational model as the prior target sentence than test sentences that were superficially similar to the target sentence but conveyed a different situational model from the target sentence. In either type of test sentence, the propositional representations could not have helped the reader to match test to target.

For an example of a situational model being formed of riddles, consider the following "Magician" riddle: "A magician claimed to be able to throw a ping pong ball so that it would go a short distance, come to a dead stop, and then reverse itself. He also added that he would not bounce the ball against any object or tie anything to it. How could he perform this feat?" In this riddle, a situational model initially formed could be a mental image of a magician throwing a ball horizontally, only for the ball to magically reverse direction mid-air and return to the magician's hand.

Before we move on to the mechanisms supporting the evaluation of our riddles' comprehensibility, let us consider a few alternative views on the types of mental representation that are thought to be active in riddle solving, other than the concepts, propositions, and situational models that we have described so far. One such mental representation, used by Luo and Niki (2003), characterizes riddles as "reference frames," and the finding of appropriate riddle solutions as being similar to the spatial navigational reorientation function of the hippocampus. According to this view, external cues trigger an appropriate reference frame (a mental representation of the environment), sometimes done so via a "context switch". However, the idea of a reference frame is actually very similar to that of a situational model; and the idea of a context switch is very similar to the idea of the "model updating" performed on situational models, as described by Zwaan and Radvansky (1998). The similarity between situational models and reference frames becomes even more apparent when we consider a study by Rinck, Hahnel, Bower, and Glowalla (1997). One of their findings that illustrated this was that: participants who formed situational models after reading a narrative about a person traversing rooms in a building, and who also studied a map of the same building, were afterwards able to judge which objects in the building were closer to a protagonist in the text based on travel distance calculated as a function of number of rooms traversed and length of room. This strategy of calculating distance is indistinguishable from the "dead-reckoning" navigation that is described to occur with reference frames. Furthermore, Redish (2001) described reference frames as being linked to goals, similar to how Zwaan and Radvansky (1998) described situational models as sometimes possessing a motivational dimension. Therefore, the notion of reference frames as a representation may simply be situational models in another guise.

Insight has also been characterized in the research literature as invoking a mental representation known as a "problem space" (Ollinger & Knoblich, 2009). The elements of a problem space include states (such as a start state, intermediate states, and goal states), legal operators that can be applied to transform one state to another, and strategies for applying those operators (either exhaustive strategies which always result in the solution, or typically faster rule-of-thumb heuristics which apply operators when certain conditions are met). Such a mental representation would indeed be capable of representing many types of insight problems, not merely verbal ones. However, the idea of riddles being represented as a problem space is not incompatible with the notion of propositions and situational models if we: view start and intermediate states as work-in-progress sets of concepts, propositions, and situational models; view operators as those mental processes which build up the current set of concepts, propositions, and situational models, based on on-line verbal processing; and view the goal state as being a comprehensible representation of the complete verbal input. In such a perspective, the problem space is simply a larger overarching organization of concepts, propositions, and situational models. And it should be noted, that even Ollinger and Knoblich acknowledge that we probably don't represent the problem space in full, but rather in part -a "subjective" problem space (p9).

After we initially construct a riddle as a set of concepts, propositions, and situational models, we then begin the process of evaluating whether those representations make sense. Initially, they will not, unless we have heard of the riddle's solution before. Essential to this comprehensibility evaluation process are the two higher order representations: propositions and situational models (van Dijk & Kintsch, 1983 p.150, 189, 339).

One level at which comprehension in riddles can fail is at the level of local coherence. Broadly, what this means is that while we expect a riddle's component propositions to be interrelated, they are initially difficult to relate (for further treatment of local coherence, see Carroll, 2008 p158; Graesser, Millis, Zwaan 1997; and van Dijk & Kintsch 1983 p149). For example, in the "Twins" riddle, " Marsha and Marjorie were born on the same day of the same month of the same year to the same mother and the same father - yet they are not twins. How is that possible?" we may form the propositions: BORN(MARSHA AND MARJORIE, SAME DAY) and ARE NOT(MARSHA AND MARJORIE, TWINS). We expect these two propositions to be related by virtue of semantically co-referring to the same concept. Initially, this riddle seems puzzling because the dominant concept of TWINS does not form a common referent for the two propositions. However, when a bridging concept to which both propositions can refer to (i.e., "TRIPLET") is inferred, the two propositions are suddenly related, and comprehension occurs. Failure to connect the propositions within the riddle—that is failure to obtain local cohesion—may result in the initial incomprehensibility of the riddle. And people do spend mental resources on creating such local coherence. For example, Haviland and Clark (1973) found that people took time to mentally create bridging inferences to connect propositions when reading. Particularly, they found that proposition pairs that required a bridging inference took participants longer to understand than proposition pairs whose relationships were obvious from having an explicitly shared referent; this suggests that people expend effort to make propositions cohere locally, especially when it is not obvious that they do cohere.

A second level at which riddle comprehension may fail is at the level of global coherence. Broadly, what this means is that the constructed situational model of the riddle fails to identify with some external overall organization (i.e., schema) (van Dijk & Kintsch, 1983, p189) (for more on global coherence, see Carroll, 2008 p160; and Graesser, Millis, and Zwaan, 1997). Schemas can be thought of as mental representations that exist outside of the verbal material (including such things as narrative frameworks and world knowledge), which in turn specify the relations among constituents in the verbal material (Anderson, Reynolds, Schallert, & Goetz, 1977). The specific constituents of a particular schema can be variable. As a result, a schema is a general type (i.e., template), rather than a token (i.e., exemplar). If we were to compare a schema to a situational model, the relationship would be that the situational model is a specific instance of a more general schema (Zwaan & Radvansky, 1998). Global coherence is necessary for comprehension. To experience how comprehension can fail when our situational model of the text doesn't match an external schema, consider the following disjointed passage found in Dooling and Lachman (1971):

> "JOE LOOKED OUTSIDE/ FROM CRAMPED QUARTERS/ NUMEROUS UNKNOWN OBJECTS/ MOVED SWIFTLY BY/IN VAGUE BLACKNESS/ AROUND HIS FIELD/TWO FEARLESS COMPANIONS WORKED ALONG/ MANIPULATING BUTTONS/ WHILE READING COMPLEX PATTERNS/ FLAT FAMILIAR HOMELAND/ NOW ACTUALLY RESEMBLED/ A TINY RUBBER BALL/ EVERYONE HERE AND AT HOME/ KNEW THAT ONLY LIFELESS THINGS/ WOULD BE FOUND/ AMONG HUGE COLD MOUNTAINS/ SURROUNDING DEEP BARREN VALLEYS/ BUT ALL IMPORTANT PAPERS/ ANXIOUSLY AWAITED THEIR ARRIVAL/ FOR NO MAN/ HAD EVER MADE/SUCH BIG NEWS."

After reading this passage, we may come away feeling a bit confused because our situational model of the passage may not immediately match any schema. However, if you, the reader, were to have known that the matching schema for this passage was "The First Space Trip to the Moon", you might have comprehended the passage much more easily, and remembered its contents better (this was what Dooling & Lachman found). Now recall the riddle of "Professor Bumble". In this riddle, comprehension was likely initially stymied because we formed a situational model of Professor Bumble blatantly committing a traffic violation by driving against the direction of traffic while a policeman looked on. This situational model did not match any of our prior schemas concerning traffic law enforcement. The riddle can be solved, however, when we replace the concept DRIVING, with WALKING, which reconstructs the former situational model to that of Professor Bumble walking down a street while a policeman looks on. This updated situational model does match our past schema of how a policeman typically behaves in such a situation. As this example riddle shows, comprehension can be initially stymied because we are unable to achieve global coherence (i.e., match our constructed mental representation of the riddle to an external schema), but comprehension can be obtained once we achieve it.

After initial comprehension of a riddle fails, we turn to processes involved in breaking down and rebuilding the initial set of mental representations of the riddle (i.e., the initial set of concepts, propositions, and situational model of the riddle) so that comprehension can be achieved. These processes may be triggered after a certain amount of time has passed or several attempts have been made where no solution (i.e., comprehension) can be obtained (Seifert, Meyer, Davidson, Patalano, & Yaniv, 1994, p111) with the currently activated concepts, propositions, and situational model. Alternatively, a complete halt in comprehension may not need to occur. Rather, restructuring of the riddle may be initiated when the constructed propositions or situational model, interpreted as they are, fail to add sufficient *additional* meaning to the current state of the riddle, according to some metric (this metric is currently unknown for riddles). This notion comes from the Criterion for Satisfactory Progress Theory view of insight (Ormerod, McGregor, Chronicle, 2002). The aforementioned halted or slowed progress on riddles may both fall under the term "impasse."

The need to break down our initial set of mental representations of the riddle comes from the idea that this initial set (including the riddle's initial concepts, propositions, and situational model) blocks related solution concepts from being retrieved. This is related to the idea of functional fixedness in insight, where the common use of an object may prevent the use of the object in novel ways that actually solve the problem. Functional fixedness can be seen in Duncker's candle problem, where the participant's goal of creating a platform on a wall to hold a candle, with nothing more than matches, candles, and a box of thumbtacks, was impeded by the participant's typical usage of each object -the solution was to not use the box as a thumbtack holder, which was its original function, but rather to pin it to the wall to use as a platform (Duncker, 1945). Similarly, Luchins (1942) showed that when participants were asked to solve a problem of obtaining a certain amount of water using multiple measuring jugs of various sizes, they used a more difficult strategy when that strategy was primed by previous problems, where it worked, rather than a new simpler strategy that was not primed by previous problems. Control participants who did not receive priming of the more difficult strategy used the simpler strategy instead. Both Duncker's and Luchins' studies support the general notion that an initial dominant, but misleading, set of mental representations of a problem can block later appropriate solutions if not dealt with. Recall that in our "Professor Bumble" riddle we may initially form the set of concepts, propositions, and situational models consistent with the idea that Professor Bumble was driving. This initial set of mental representations is sticky and blocks our consideration of alternative concepts of the riddle that offer the solution. Only by breaking down this initial set can we realize the key concept to solving this riddle, that Professor Bumble was not driving, but WALKING.

One potential process for breaking down the initial set of concepts, propositions, and situational model is forgetting. Demonstrating that forgetting contributed to insight solution, Smith and Blankenship (1989) found that participants who were tasked to solve an insight problem (a rebus puzzle) were more likely to solve it if they forgot a misleading cue concept which suggested an inappropriate solution strategy. This finding, applied to our "Professor Bumble" riddle, would suggest that after a period of remaining stumped, we may forget our initial concept of DRIVING (and presumably, the propositions and the situational model constructed with it). Once this initial set is forgotten, we may notice other modes of transportation, such as WALKING, which leads to the solution propositions and situational model. However, it should be noted that we are not claiming that forgetting is a passive process only due to lack of rehearsal. In Smith and Blankenship's study, they found a marginally significant improvement (p=.06) in the rebus problem's performance in participants who listened to music during a break over those who did not listen to music during the break period (though both groups who listened to music during the break and those who didn't listen to music during the break performed better than those who had no break). This finding offers an interpretation of forgetting not simply as lacking rehearsal (though lack of rehearsal may certainly contribute to forgetting) but also as having an inhibiting distractor. A simple, albeit speculative, example of how a successful riddle solver may possibly offer his/her own mental distractor to forget the dominant misleading concept (and presumably, its constructed propositions, and situational model) may be by subvocalizing or reading parts of the riddle not tied to those dominant misleading concepts, such as the part of the "Professor Bumble" riddle that simply says, "street".

A potential process especially suited to breaking down the initial situational model is chunk decomposition. Knoblich, Ohlsson, Haider, and Rhenius (1999) defined chunks as "recurring constellations of features or components" of objects or events and suggested that decomposition of chunks may be necessary for overcoming problem impasses. Knoblich, Ohlsson, and Raney (2001) obtained results that evidenced the role of chunk decomposition in solving insight problems. In particular, they showed that successful solvers of their insight problems (matchstick math problems) attended longer to visual chunks of the problems requiring decomposition than non-solvers. So if chunk decomposition operates in insight, how might it work in riddles? Consider our "Magician" riddle from earlier. Initially, the situational model that we form of this riddle may have involved a scene where a magician threw a ball sideways. However, if we visually decompose (i.e., chunk decompose) this scene into the image of a magician, and the image of a ball, we can more easily reorganize the component scene concepts (because they are no longer tightly bound to an organization). For example, with the above scene decomposed, we may more easily trade the concept of the ball moving horizontally for the concept of the ball moving vertically, which is indeed the solution to the riddle (i.e. the magician threw the ball up and caught it coming down). Considering that the chunk decomposition in Knoblich, Ohlsson, and Raney (2001) occurred with visually grouped chunks, chunk decomposition, if operating in riddles, most likely decomposes situational models, which have a spatial dimension (Zwaan & Radvansky, 1998). It is less clear whether the process of chunk decomposition may operate to directly decompose propositions.

With the initial set of concepts, propositions, and situational model either weakened or decomposed, the task now turns to retrieving the correct non-dominant concepts to rebuild the riddle propositions and situational model, so that they can achieve local and global coherence. One way in which such concepts are retrieved is via the "opportunistic assimilation" account (Seifert, et. al, 1994). According to this account, we initially generate a special memory trace called a “failure index” when we come across an incomprehensible text. The failure indices point back to the concept where comprehension failed. Their role is to reactivate that troublesome concept in the riddle when the riddle solver chances upon cues present in the environment (e.g., words on a page) related to the riddle's solution, though internally generated cues are also speculated to help as well (e.g., daydreaming). This way, the cue can retrieve the correct reinterpretation of that concept. For example, in the "Twins" riddle, when comprehension fails, a failure index may be placed at the inferred concept of TWINS, which is originally troublesome because the riddle specifies that Marsha and Marjorie cannot be twins. When the solver of this riddle encounters a cue for other forms of siblinghood in the environment (such as seeing triplets walk by, or perhaps internally encountering triplets through rumination), then the riddle may be reactivated, and the correct concept associated with the cue (of TRIPLET) may be inserted as the solution concept.

There are really two main notions that need to be justified in claiming the above account of solution concept retrieval: (1) that initial failure to solve the problem contributes to insight and that (2) external cues contribute to insight. In support of the first notion, Seifert et al. (1994, p99) found that participants were more likely to solve verbal problems after surreptitiously receiving solution cues when the same verbal problems had been attempted prior to receiving the solution cues. Supposedly, failure to solve the problem the first time made the solution cue more salient, as would have happened in the opportunistic assimilation account of solution retrieval.

Further evidence for the importance of failure in reaching insight solutions comes from Patrick, Ahmed, Smy, Seeby, and Sambrooks (2015), who turned the recognition of failure to comprehend riddles into a training protocol. In their training protocol, participants were trained to explicitly check their interpretation for each part of the riddle and revise their interpretation if inconsistency arose between it and some specification of the riddle (i.e., comprehension failure). Those who received this failure recognition training performed better (68% correct) on riddles than those who either: received training involving being given the correct riddle solution (39% correct), or received no training at all (30%). Furthermore, the benefit of failure recognition training persisted two days after the initial training period.

In support of the second notion critical to the opportunistic assimilation account of solution retrieval, that external cues facilitate insight, are findings by Rowe, Hirsh, and Anderson (2007). Specifically, their study demonstrated that broad visual attention (which would allow for greater detection of external solution cues in the visual field) was associated with better insight performance, by showing that performance on an insight problem known as the Remote Associates Task (RAT) was positively correlated with degree of distractibility by items in a person's visual periphery, as measured by the Eriksen flanker task.

Also consistent with the role of external cueing in insight is a study by Kounios et al. (2008), which found resting state electroencephalogram (EEG -a measure of brain activity based on electrical signals detected at the surface of the scalp) differences between a group of high-insight participants and low-insight participants in the occipital lobe, using an anagram insight task. High-insight participants showed less alpha wave activity in the occipital lobe, which the authors interpreted as reflecting less visual gating prior to seeing the anagram. Like broad visual attention, this would potentially allow more external insight-related cues to stay activated during the insight problem. Low-insight participants, on the other hand, showed more beta wave activity in the occipital lobe than high-insight participants, which the authors interpreted as reflecting more focused visual attention. This would have potentially excluded those external cues in the visual periphery from being available in the subsequent anagram solving.

Alternatively, the search for alternative solution concepts may occur through outward spreading activation (Collins & Loftus, 1975) originating from the features of the dominant but incorrect concepts, or originating from other concepts within the riddle. However, unconstrained, spreading activation occurring in this manner may explode the search space uncontrollably, producing more leads than can possibly be evaluated. Therefore, the notion of a constrained spread has been proposed (Ollinger & Muller, 2017).

One way in which spreading activation may be constrained in riddles is through an abstraction heuristic. The meaning of abstraction here is essentially the same as used by van Dijk and Kintsch (1983) for propositions. The idea here is that the solver can abstract from a problematic concept to a higher, more general conceptual node which is more directly connected to the solution concept (assuming concepts are organized in a hierarchical fashion). For example, consider the following "Ancient Invention" riddle, which goes: "There is an ancient invention that is still used in many parts of the world and that allows people to see through the walls. What is it?" (solution: a window). In this riddle, we may initially be constrained by the concept of an ANCIENT INVENTION. It may cause us to imagine some class of old and esoteric inventions, not everyday inventions, which windows are probably categorized as in most people's minds. However, if we generalize from the concept of ANCIENT INVENTION to the more general concept of INVENTION, we move closer to the class of everyday inventions and thus to the concept of WINDOW. Support for the abstraction view of solution concept retrieval comes from Forster, Friedman, and Liberman (2004), who demonstrated in a series of experiments that thinking about the distant future engaged abstraction processes; and that, presumably by virtue of engaging abstraction processes, distant future thinking also facilitated generation of a greater number of insight problem solutions.

Once the correct solution concepts are retrieved, they replace the misleading initial concepts to build up new propositions and a new situational model of the riddle. This time, however, comprehension via local coherence and global coherence succeeds with the newly constructed propositions and situational models, and the riddle is solved. To summarize, riddles may be viewed as an instance of verbal comprehension where the conditions for comprehension (local and global coherence) may be viewed as conditional rules that, when violated, suppress the activation of the current riddle concepts, propositions, and situational model, and engages processes that retrieve alternative concepts -among which lies the key to the riddle's solution.

Before leaving discussion of our cognitive model of insight, it is worth noting that some insight researchers characterize insight as occurring in two stages: 1) obtaining the correct constituents of the solution, which, in riddles, involves *retrieving* the relevant concepts pertaining to the riddle; and 2) *constructing* the correct constituents into their solution structure, which, in riddles, involves the construction of the appropriate propositions and situational model, using the appropriate concepts, as well as *evaluating* those constructed propositions and situational models. These two stages are also termed "approach recognition" and "approach execution," respectively (Sternberg, 1995, p580-582). However, only the first stage is actually hindered in short verbal riddles of the type we will be using in this study. We can see this is the case because incremental progress on the problem solution, with many states in-between the initial set of constituents and the final solution state, would be a feature of a prolonged second stage. However, as we saw in Metcalfe's (1986) study (also see Jausovec & Bakracevic 1995, and Smith & Kounios 1996), correct riddle solutions were obtained suddenly with no apparent in-between progress. This indicates that for riddles, the second stage occurs quickly and automatically. This is not to say that construction and evaluation of the correct constituents (concepts) of a riddle, once retrieved, does not occur (they do, because the retrieved solution concepts must still be constructed into propositions and/or a situational model). But what this does imply is that the critical work in comprehension of riddles lies in obtaining the right solution constituents (concepts). Once the correct set of constituent concepts are active in memory, they are automatically constructed into propositions and a situational model, which are then evaluated for local coherence and global coherence.

Insight Partly Involves Common Processes

In the past, some researchers have viewed critical insight processes as being common and non-exclusive to insight (e.g., Chronicle, MacGregor, & Ormerod, 2004; Ball & Stevens, 2009), while others have viewed critical insight processes as being unique to insight (e.g., Knoblich et al., 1999). However, here we view insight as possessing both types of component processes, common and unique, with common processes no less capable of affecting performance on an insight task than unique processes. Some processes, such as spreading activation of concepts, building propositions and situational models, and evaluating statements for coherence, may occur during any kind of verbal processing task, including riddles. Other processes, such as chunk decomposition, may be unique to insight. Consistent with the common process notion, various components of general intelligence (such as working memory capacity, and fluid reasoning) were found to correlate with performance on insight problems (Chuderski & Jastrzebski, 2017). And consistent with the unique process notion, working memory capacity and fluid reasoning, while able to account for much of the variance in insight performance (around 2/3 of it), could not account for all the variance in insight performance (Chuderski, 2014).

Other Insight Tasks

Several tasks have been used in the neuroimaging literature to characterize insight. However, as Dietrich and Kanso (2010) have pointed out, robust neuroimaging findings (i.e., consistent brain locations associated with a particular cognitive process) are hard to come by, even when studies are grouped under the banner of "insight". But if we look to the field of deductive reasoning research, we can see that the practice of carefully noting the subtle cognitive process differences between different tasks, and regrouping results based on such a detailed analysis can provide a more consistent picture of the locations associated with each cognitive process (Prado, Chadha, Booth, 2011). Therefore, while acknowledging that differences in location of neuroimaging results can arise due to other reasons (e.g., when one utilizes functional magnetic resonance imaging [fMRI] which detects brain blood flow versus [EEG]), at least one key to explaining the location of one's insight neuroimaging results may be distinguishing between the subtle cognitive differences that exist between one's own insight task and others' insight tasks. Core among the processes underlying all forms of insight is the notion of mental representation restructuring (Weisberg, 1995). Yet, in different insight tasks, the notion of *what* mental representations are restructured, as well as *how* they are restructured differs. Here, we will attempt to identify the *what* and *how* of restructuring in two other tasks, and contrast them to our own riddle task in an attempt to use the *what* and *how* of restructuring as potential dividing lines to distinguish our insight task from others.

The Remote Associates Task (RAT) (Mednick, 1962) was originally conceived, not as an insight task, but as a measure of general creativity defined by the ability to make distant semantic associations while also satisfying multiple constraints. In service of these goals, the task involves asking the solver to find a single word which is distantly connected (with "distantly connected" possibly being operationalized as infrequently co-occurring in writing) to each of three words (e.g., being given the words "rat", "blue", and "cottage", and being tasked to find a word that binds them, such as "cheese"). However, it has since been used as an insight task in neuroimaging research, after its initial conception, due to the sudden obtaining of the solution that can occur in the task (e.g., Jung-Beeman et al. 2004), which suggests a cognitive gap between the initial problem representation and the solution. *What* appears to be restructured in the RAT are just lexical concepts (such as "blue") and the concepts associated with them (such as "sky"), as opposed to the higher order propositions and situational models which also require restructuring in riddles. *How* restructuring occurs may include weakening the associational relationship between the cue words (such as "blue") and their dominantly associated concepts (such as "sky"), such that non-dominantly associated concepts to "blue" (such as "cheese") can be retrieved and evaluated for coherent associations with other cue words. Once a lexical concept (tied to a word) is retrieved, that is associated with the lexical concept of all three cue words, that word becomes the solution. While in the RAT, solution representation construction links concepts more so by association, in riddles solution representation construction links concepts more so using rule-based relationships between concepts to build higher order representations that become the solution. For example, in the proposition OBSERVED(POLICEMAN, SCENE), from the "Professor Bumble" riddle, the concepts POLICEMAN, OBSERVED, and SCENE enter into an actor-action-object relationship, where there is a specific relationship between each concept rather than a nebulous associational binding between words based on frequency of word co-occurrence.

Another class of insight task which has been used in neuroimaging includes mathematical hidden rule-discovery tasks. For example, Lang et al. (2006) used a task called the Number Reduction Task (NRT), which required participants to generate a string of 5 numbers from a given string of 6 numbers. The explicit rules which were given allowed participants to generate the number string in a straightforward but tedious manner. However, there was a hidden underlying rule that the latter 2 generated numbers would always mirror the first 2 generated numbers. Once the hidden rule was discovered, solution time per trial would drop. As in the RAT, the "solution" to the NRT is also acquired suddenly, which is reflected in the sudden drop in solution response times for those participants who discover the hidden rule. On the surface, *what* appears to be restructured in mathematical hidden rule discovery tasks are the transformations applied to numbers in the problem. This stands in contrast to riddles, where the transformations applied to concepts, which allow them to form propositions and situational models, remain the same from the riddle's initial formulation to its solution state. *How* the initial transformation in mathematical hidden rule discovery tasks is restructured may be through noticing of "change invariances", such as noticing the way that certain moves often have predictable results (Kaplan & Simon, 1990), thus hinting at a more efficient transformation.

Development of Insight Performance

When it comes to the question of how insight develops between adolescence and adulthood, we can refer to a relevant study by Kleibeuker, De Dreu, and Crone (2013). In their study, they looked at visual insight performance (i.e., Gestalt Completion Test & Snowy Picture Test), and verbal insight (i.e., RAT) performance across adolescence and into adulthood.

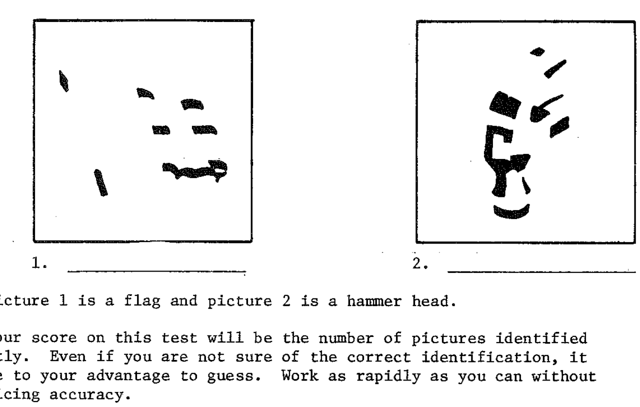
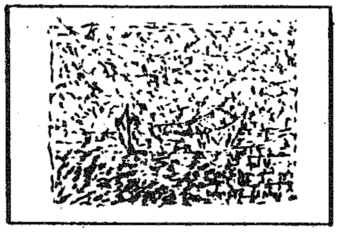


Figure 3. Example Gestalt Completion Test item (Ekstrom, Dermen, & Harman, 1976)



2. ____________________

The picture shows a small boat sitting in the water. Boat, rowboat, or other similar words would be correct answers.

Your score on this test will be the number of objects that you name correctly. Work as quickly as you can without sacrificing accuracy. If some pictures are difficult, skip them and return to them later if you have time.

Figure 4. Example Snowy Picture Test item (Ekstrom et al., 1976)

The authors found that the two older age groups (18-19, and 25-30 years old) performed more accurately on both visual and verbal insight tasks than the two younger groups (12-13, and 15-16 years old). Furthermore, the two visual insight measures were better fitted (had more variance explained) by a post-hoc ANOVA contrast (Tukey's HSD) between the two older and the two younger groups, which perhaps suggests a more sudden change in visual insight performance with aging; while the verbal insight measure was better fitted by a continuous curvilinear model where verbal insight was regressed against log-transformed age, which perhaps suggests a more gradual improvement in performance for verbal insight with aging.

It is possible that the difference in insight performance between the older and younger groups was a result of cognitive processing differences in the way they solved insight problems. Such cognitive differences could be due to the gradual cognitive maturation and brain maturation that occurs over the course of adolescence.

Development of Linguistic Abilities in Adolescence

During adolescence, language abilities continue to develop. Of potential importance to riddles are developments in syntax, semantics, and pragmatics that occur during adolescence (see Nippold, 1993 and Nippold, 2000 for an overview). Nippold, Hesketh, Duthie, & Mansfiled (2005a) showed that produced syntactic complexity as measured in mean length, in words, of each T-unit (an independent clause plus any dependent clauses), as well as the number of T-units with dependent clauses, rose from childhood through adulthood (a range of 7-29 years), and remained steady from adulthood through to middle age (up to 49 years). Many riddles are presented in complex sentences with dependent clauses. For example, in the "Professor Bumble" riddle, "*A policeman observed the entire scene* but did nothing about it " is a complex sentence. It has both an independent and dependent clause (independent clause italicized). Likewise, for the "Magician Riddle," "*A magician claimed* to be able to throw a ping pong ball so that it would go a short distance, come to a dead stop, and then reverse itself " the first sentence has an independent clause (italicized) and multiple dependent clauses. Therefore, the increased ability to syntactically process these riddles as adolescents age may facilitate overall comprehension of riddles.

Furthermore, Nippold, Hegel, Sohlberg, and Schwarz (1999) showed that the ability to identify key features associated with a word, and abstract from the word's lexical concept, as measured by an abstract word-defining task, improved from adolescence to adulthood. This is relevant to riddles because the process of identifying the initially misleading concept in the riddle, and abstraction from that concept to a higher conceptual node where other concepts can be more easily retrieved, are potentially crucial processes in retrieving the solution concepts in riddles. Specifically, in their study, participants were asked to give definitions of words in the form "an X is a Y that Z" where X is the given word, Y is its superordinate category term, and Z is a characteristic feature of X. An example answer that would have been acceptable to the prompt, "peace", was, according to the authors, "Peace is a condition in which people are free of strife and disorder." Two points were given for having both an appropriate superordinate category (i.e., meaning the participant was able to abstract) and key feature (i.e., meaning the participant was able to identify critical concepts related to the lexical concept of the target word). One point was given for having either the superordinate category or feature. No points were given for having neither. Critically, young adults scored higher on this task than children and two adolescent groups, showing that feature identification and abstraction abilities were improving with age.

There is also evidence that older adolescents understand figurative language, such as proverbs, better than younger adolescents (Nippold, 2000). Figurative language is akin to insight in that comprehension of both requires access to non-dominant concepts. The ability to understand proverbs positively correlated with the ability to provide more "accurate, clearer, and more complete" definitions to keywords that were used in the proverbs (Nippold & Allen, 1998, cited in Nippold, 2000). What this might suggest is that the ability to retrieve concepts from a richer and more specifically related semantic network may underlie improved proverb comprehension. This is relevant for riddle solving, because the riddle solver may benefit from being able to access a richer conceptual network with well-specified relationships, which may help the riddle solver to see alternative non-dominant solution concepts.

Pragmatics development over adolescence may also be relevant to riddle solving. In particular, Larson and McKinley (1998) found that the number of abrupt topic shifts in videotaped conversations with adults and familiar peers decreased with age over adolescence. What this may indicate is that, with aging, adolescents may maintain attention to certain trains of thought better. This may be of importance for reaching impasse for riddles, as one may need to maintain attention to a certain interpretation of a riddle long enough to realize that it is untenable, which triggers restructuring of the riddle's concepts, propositions, and situational model.

Pragmatics, in the form of the ability to acknowledge other viewpoints during persuasive essay writing, also becomes more common in adult writers than older adolescent writers, and in older adolescent writers than young adolescent writers (Nippold, Ward-Lonergan, & Fanning, 2005b). This ability to adopt alternative perspectives may be related to the ability to retrieve alternative concepts while solving riddles. Together, improving syntactic, semantic, and pragmatic processing abilities over adolescence, and into adulthood, leads us to expect that riddle insight performance will improve with age from adolescence to adulthood.

Neural Structural Development with Aging

The general picture painted by macroscopic neurodevelopmental trajectories seems to suggest that the brain regions related to higher cognitive functions do not complete maturation before adolescence. Maturation occurs well into adolescence and beyond, leading us to expect cognitive developments in this time period. This is evidenced by cortical gray matter (GM) volume development patterns showing later peaking in higher order associational regions such as the dorsolateral prefrontal cortex, inferior parietal gyrus, and superior temporal gyrus, than primary sensorimotor regions (Giedd, 2008, Gogtay, 2004). Gray matter refers to the neuronal cell bodies in the central nervous system (Bear, Connors, and Paradiso, 2007, p802). Because the general trend for GM volume development seems to conform to an inverted U trajectory from childhood to late adolescence, we can at least see here that these higher order associational areas develop later than other parts of the brain. Furthermore, work by Taki et al. (2013) in mapping the development of GM volume at the sub-lobular level confirms that cortical GM volume increases well into adolescence, and shows that GM volume peaks in several areas after the age of 17. And while later peaking suggests later neural and cognitive maturation, it does not imply the end of such maturation. This is because, while no hard conclusions can be drawn about exactly what happens at the microstructural level (cellular and synaptic), one possible account for this general pattern is that of pre-adolescent GM volume increase due to synaptogenesis, followed by post-adolescent decrease due to reduction in synapses per neuron and increases in myelination (Taki et al. 2013). By this account, the increase in GM volume leading up to its peak is maturation by synaptogenesis (i.e., connecting brain regions that need to work together for a cognitive task), and the decrease in GM volume post-peak is maturation by synaptic pruning and more efficient inter-neuron communication (i.e., making the connected brain regions communicate more efficiently). In particular, the notion that age-related synaptic pruning occurs in adolescence is supported by histological data from Huttenlocher (1979), which showed decreases in synaptic density in the middle frontal gyrus into mid-adolescence.

White matter (WM), which refers to central nervous system tissue classified as neuronal axons (Bear, Connors, and Paradiso, 2007, p816), development paints a similar picture of protracted neural development. During development, we see total WM volume increasing with no peaks during childhood and adolescence (Giedd 2008), suggesting prolonged neural development in white matter structures into adulthood. The general pattern in white matter development, based on autopsy, seems to progress from inferior to superior and from posterior to anterior, with the brain stem and cerebellum myelinating before the cerebral hemispheres, and with the frontal lobes myelinating last within the cerebral hemispheres (Sowell, Thompson, Holmes, Jernigan, & Toga, 1999). This suggests that WM supporting higher cognitive functions may mature later than WM supporting sensorimotor functions. Also supporting late maturation of higher cognition-related white matter structures, a recent review (Lebel, Treit, & Beaulieu, 2017) showed that Fractional Anisotropy (FA) values often continue to increase into the 20's for frontal-temporal areas, cingulum, uncinate fasciculus, superior longitudinal fasciculus, and superior fronto-occipital fasciculus. FA is a measure of inter-neuron communication efficiency based on the bias in water diffusion within WM tracts. On the other hand, FA of the corpus callosum and tracts in the occipital lobe typically stabilize before the 20's. Higher cognitive functions are likely to mature concomitant to white matter development, as FA values have also been shown to positively correlate with measures of cognitive performance such as working memory and reading ability, with FA values also showing a positive correlation with age (Nagy, Westerberg, Klingberg, 2004).

Neuronal Events Underlying Macroscopic Volume Changes

Studies in the last two decades have found relationships between performance on certain tasks and regional GM brain volume. Often, these studies find that larger regional GM volumes are associated with improved performance (or more experience) on some task. For example, an early study found that the GM volume of the posterior hippocampus was positively correlated with months of experience in London taxi cab drivers (Maguire et al., 2000). Another study found that three months of juggling practice increased the GM volume of bilateral mid-temporal areas and left intra-parietal sulcus over participants who didn't practice juggling (Draganski et al., 2004). Furthermore, Seyed-Allaei, Avanaki, Bahrami, and Shallice (2017) found that participants who solved a mathematical rule discovery problem (called the Nim problem) with insightful solutions had larger GM volumes in the anterior medial, ventrolateral prefrontal, and parietal cortices than those participants who did not find the insightful solutions. Such a "bigger is better" phenomenon is often explained in terms of synaptogenesis and dendritic arborization leading to improved inter-neuronal communication (Maguire et al., 2000; Draganski et al., 2004; Kanai & Rees, 2011).

However, smaller regional GM volumes have also been associated with improved performance on certain tasks as well. For example, Bendetowicz, Urbanski, Aichelburg, Levy, and Volle (2017) found a negative correlation between insight performance and GM volume in the left rostrolateral prefrontal and left posterior parietal regions, with participants who solved more RAT-type problems showing smaller GM volumes in these brain regions. Another study by Jung et al. (2010) found a negative correlation between cortical GM thickness of a region within the lingual gyrus and scores on a creativity survey (the Creativity Achievement Questionnaire). One explanation for this "smaller is better" phenomenon was proposed by Kanai and Rees (2011), who suggested that smaller GM volume may benefit a cognitive function by reducing the distance over which neurons need to communicate. This is because, over longer distances, the thickness of dendrites needs to increase disproportionately more, compared to increases in length, in order to maintain the same communicative efficacy. Therefore, a large biological burden is placed on dendrites with increases in distance over which they need to communicate. Pruning of unneeded synapses may reduce the overall GM volume, reducing the distance over which remaining synapses need to communicate.

An additional explanation for negative correlations between GM volume and task performance, also based in synaptic pruning, is that the process of synaptic pruning also strengthens the remaining synaptic connections, resulting in their more efficient cognitive functioning (Blakemore & Choudhury, 2006). Either or both of these two accounts may lead to more efficient retrieval of concepts used in restructuring an insight task's initial representation, and explain why "smaller is better".

Admittedly, though, these explanations of what goes on at the neuronal level when we observe either macroscopic increases or decreases in GM volume being associated with better performance on a task are, at best, post-hoc explanations. Current research is simply unable to exactly account for why we observe GM volume increases with better cognitive performance in some situations, but GM volume decreases with better cognitive performance in other situations.

Purpose of Study

In this study, we tried to address the following questions:

1. Is there a cognitive difference, involving common processes, between adults and adolescents when it comes to how they solve insight problems (i.e., riddles)? Kleibeuker et al. (2013) have demonstrated that performance on insight problems improves between adolescence and adulthood. During roughly the same developmental period, past research has demonstrated that linguistic abilities (a set of common processes), and brain structures continue to develop (Nippold, 1993; Larson & McKinley, 1998; Nippold et al., 1999; Nippold, 2000; Nippold et al., 2005a, 2005b). Past research has also demonstrated the partial role of common processes in insight (Chronicle et al., 2004; Ball & Stevens, 2009; Chuderski, 2014; Chuderski & Jastrzebski, 2016). Taken together, these suggest that such a cognitive difference may indeed exist. We used behavioral performance data to verify this.

2. Does an insight-volume relationship exist for riddle problems? Past research has recently demonstrated that an insight-volume relationship exists for two types of insight tasks: the RAT task (Bendetowicz et al., 2017), and the Nim problem (Seyed-Allaei et al., 2017). We used Voxel Based Morphometry (VBM) analysis to see whether this relationship extends to riddle-type problems. Furthermore, we tested whether this insight-volume relationship: 1) is revealed in GM regions exclusive to adults *or* adolescents, 2) is revealed in GM regions common to both adults and adolescents. Bendetowicz et al. (2017) and Seyed-Allaei et al. (2017) have examined the insight-

volume relationship in adults, but not in adolescents. Due to the changing balance of neuronal events that occur as people age (such as the balance of synaptogenesis versus pruning), when we analyzed adults and adolescents separately, we speculated that we might see insight-volume relationships manifest, which might otherwise have been hidden if we were to have examined only adults.

3. Does an insight-WM tract integrity relationship exist for riddle problems? It is reasonable to expect associations between insight problem solving and WM tract integrity, given that measures of divergent thinking (another type of creativity) performance were previously found to be correlated with measured WM tract integrity (FA) within the frontal cortex, as well as between the frontal cortex and multiple other regions (Jung, Mead, Carrasco, & Flores, 2013; Takeuchi et al., 2010). We tested whether this relationship exists by analyzing the relationship between riddle performance and brain WM tract integrity, as measured by FA. Furthermore, we tested whether certain insight-WM tract integrity relationships: 1) exist exclusively in adults or adolescents, 2) were revealed in WM tracts common to both adults and adolescents. As with the VBM analysis, we surmised that developmental factors might give rise to unique FA findings when analyzing adults and adolescents.

Methods

Participants

The current study applied factorial and parametric analyses to behavioral data, to archival T1 weighted structural Magnetic Resonance Images (MRI), and to diffusion-weighted MRI. A total of 84 participants' data were analyzed (39 females; 45 males). Their ages ranged from 11.16 to 16.92 (*Mean = 14.77; SD=1.47*) for adolescents (22 females; 21 males), and 20 to 24.33 (*Mean=21.86; SD=1.26*) years for adults (17 females, 24 males). All participants were adolescents or older, defined as having began puberty according to the Self Rating Scale for Pubertal Development, which classifies pubertal status based on self-reported secondary sex characteristics (Petersen, Crockett, Richards, & Boxer, 1988). This was done because our focus is on the relationship between insight and brain structures in the developmental span occurring after childhood (i.e., adolescence or early adulthood). The data were collected as part of a larger neuroimaging study in Tenerife, Spain, where participants performed multiple tasks measuring creativity, reasoning, and risky decision making.

Tasks

For the behavioral insight data to be analyzed in this study, participants solved short verbal riddles in their native language, Spanish. There were ten verbal riddles in total. Participants were given 90 seconds to solve each riddle. Any answer which fit the description of the riddle was judged as correct. An accuracy score for the total percentage of correctly solved riddles was recorded. The riddles (originally in English, but translated to Spanish for this study) are provided in the appendix.

Furthermore, Intelligence Quotient (IQ) scores were obtained for each participant using either the Wechsler Adult Intelligence Scale (WAIS) or Wechsler Intelligence Scale for Children (WISC) (Wechsler, 2004; Wechsler 2014). The full scale IQ score is a composite of subtests from four categories (verbal comprehension, perceptual reasoning, working memory, and processing speed), and is thought to reflect general intelligence. Thus, it will be used as a measure of common cognitive processes (i.e., cognitive processes which are utilized in typical non-insight behaviors). According to Tulsky, Zhu, and Prifitera (2000) (cited in Goldstein & Herson, 2000), the full scale IQ score (which will be referred to simply as IQ score from here on) is highly correlated between the WISC and WAIS ($r = 0.88$). Thus, IQ scores from the WAIS and WISC are treated as comparable within this study.

MRI Data Acquisition

For the VBM data, sagittal T1 weighted images were scanned on a 3 Tesla GE-Medical System MRI scanner. This was acquired using a 3-dimensional fast spoiled-gradient-recalled pulse sequence with a slice time of 8.7ms, with slices 1mm thick. Scans were acquired within 90 days of behavioral testing when possible, or else participants were rescanned. Age during behavioral testing and age at the time of last scanning were highly correlated ($r=0.998$). Therefore, all ages used in this study were based on age at the time of last scanning.

For each participant's diffusion weighted imaging data, two acquisitions of diffusion weighted MRI scans were combined. Each acquisition occurred in opposite phase encoding directions, and consisted of 6 non-diffusion-weighted images, as well as 55 diffusion-weighted directional images, with each directional image consisting of 57 slices, each 2.4 mm thick.

Voxel Based Morphometry

The first set of analyses involved Voxel Based Morphometry (VBM). According to Mechelli, Price, Friston, and Ashburner (2005), VBM is a whole-brain unbiased technique for detecting local differences in brain tissue composition after aligning major anatomical structures. VBM is appropriate for our purposes because it allows us to look at normal anatomical variation rather than anatomical changes due to trauma or lesions, since we are examining normally developing participants who are not expected to have any lesions that may alter the major brain structures but may have smaller local tissue differences due to learning and development. It is also appropriate because it allows us to look across the whole brain for structure-function relationships. VBM utilizes a mass-univariate statistical approach to detect differences in tissue volume or density at the voxel level. In our case, we are interested in analyzing differences in GM volume. The steps that were performed for our VBM analysis involved (see Appendix: Overall VBM Workflow): parsing each participant's MRI image into separate tissue types, normalizing GM images onto a common stereotactic space (the Montreal Neurological Institute brain -i.e. the MNI152 brain) to create a template while preserving variance related to age and gender via the Template-o-Matic toolbox (Wilke, Holland, Altaye, & Gaser, 2008), parsing the images again into separate tissues, aligning the images again onto a common stereotactic space (MNI152), smoothing adjacent voxels using a 6mm full width at half maximum (FWHM) smoothing kernel, modulating each voxel by the warping factor used during normalization so that each voxel's value reflects their participant's GM volume in native space, and finally performing statistical analyses using participants' GM volume images. This particular workflow notably parses structural images a second time based on the template generated after the first parsing to increase tissue parsing accuracy. The software used for this workflow included Statistical Parametric Mapping 12 (SPM 12) (Wellcome Department of Cognitive Neurology, London, UK) running in Matlab (Math Works, Natick, MA), as well as the Template-o-Matic toolbox (Wilke et al., 2008) running on SPM 12. Locations of VBM findings were obtained from the Anatomy toolbox running in SPM 12 (Eickhoff et al. 2005). Power analyses were not feasible here, as computation of power for our VBM analyses would require statistical contrasts generated from an independent set of GM images, such as those from a pilot study or those from another study running analyses using similar behavioral data, neither of which were available.

Five types of analyses were conducted using VBM. First, GM volume was linearly regressed with riddle score in adults and adolescents combined (i.e., the GM All Regression), while controlling for age in years, total intracranial volume, IQ, and sex as covariates of noninterest. Total intracranial volume was controlled because we are interested in local GM volume differences that will allow us to identify brain regions whose structures are related to insight. Sex was controlled because, while interesting on its own, sex differences are outside the focus of this study. This analysis was performed once to test for positive regression betas, and once to test for negative regression betas, because SPM does not regress automatically in both directions. The motivation for this analysis was to reveal insight-volume relationships that might manifest when we try to increase detectability by maximizing the number of participants in the analysis.

Second, GM volume was linearly regressed with riddle score in adults only (i.e., the GM Adult Regression), while controlling for age in years, total intracranial volume, IQ, and sex as covariates of noninterest. This analysis was performed once to test for positive regression betas, and once to test for negative regression betas. The motivation for this analysis was to reveal any insight-volume relationship that might be more detectable in adults, given their stage of neural development.

Third, GM volume was linearly regressed with riddle score in adolescents only (i.e., the GM Adolescent Regression), while controlling for age in years, total intracranial volume, IQ, and sex as covariates of noninterest. This analysis was performed once to test for positive regression betas, and once to test for negative regression betas. The motivation for this analysis was to reveal any insight-volume relationship that might be more detectable in adolescents, given their stage of neural development.

Fourth, using a factorial model, the coefficient of the regression beta between GM volume and riddle score of adults was contrasted with that of adolescents (i.e., the GM Interaction Contrast) while controlling for total intracranial volume, IQ, and sex. This analysis was performed in both directions (i.e., to test for a greater regression term in adults, and a greater regression term in adolescents) because SPM does not automatically test for such a difference in both directions. Conceptually, this analysis was done to reveal any differences in the slope of GM volume and riddle score, between adults and adolescents. The motivation for this analysis was to see if any insight-volume relationships might be unique to either adults only or adolescents only.

This would be the case if the analysis were significant, and showing a significantly non-zero regression beta for one age group, while showing a zero regression beta for the other age group.

Fifth, a conjunction analysis (Friston, Penny, and Glaser, 2005) between the regression of volume and insight of adults and adolescents was performed (i.e., the GM Conjunction). This analysis used the same model as the GM Regression Contrast. This analysis was performed once to look for conjunctions between positive regression betas in adults and adolescents, and once to look for conjunctions between negative regression betas in adults and adolescents -SPM does not search for conjunctions in both positive and negative regression betas automatically. Conceptually, this analysis was done to reveal any brain regions where volume and insight were consistently related in *both* adults and adolescents. The motivation for this analysis was to see if any insight-volume relationship persisted across adolescence to adulthood.

These five types of analyses were repeated without controlling for IQ as a covariate of noninterest. The motivation for this variation on the above analyses was due to prior literature suggesting that insight, in part, consists of common intelligence-related processes. Therefore, it is of interest to see both the insight-volume relationships which includes the components of insight that are related to common intelligence-related processes, as well as those relationships which are due to insight without the effects of some common intelligence-related processes.

Fractional Anisotropy

The second method that was used to analyze our MRI data was Fractional Anisotropy (FA), which is a form of Diffusion Tensor Imaging (DTI) analysis. DTI involves the analysis of magnetic resonance brain images processed to reflect the flow of water in brain tissue. FA specifically analyzes the robustness of WM tracts by detecting the bias of water in white matter voxels to flow in one direction rather than ballooning out in all directions. Well-developed WM tracts should show a greater amount of unidirectional diffusion (i.e. reflecting a non-leaky axon), while lesser developed WM tracts should show more multi-directional diffusion. Higher anisotropy values are thought to reflect white matter fiber integrity, density, and myelination, and thus reflect more efficient neuronal communication (Kanai & Rees, 2011).

To perform FA, images were first preprocessed in FSL software (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012) (see Appendix: Overall TBSS FA Workflow). This involved: estimating distortion maps for susceptibility induced distortions (i.e., image distortions due to tissue type borders); stripping the skull and other non-brain tissues; applying correction for susceptibility induced distortions, estimating and correcting for eddy currents (i.e. distortion due to changes in gradient of the magnetic current), and estimating and correcting for participant movement; estimating FA tensors and generating an FA image for each participant; normalizing each FA image to a standard space (i.e. MNI152); extracting a standard WM skeleton from all FA images; and projecting each participant's FA image onto the standard skeleton. Finally, statistical analyses were performed on the skeletonized images in SPM 12. This method is consistent with the Tract-Based Spatial Statistics (TBSS) FA approach described by Soares, Marques, Alves, and Sousa (2013). For the FA analyses, data from a 12.75 year-old female were excluded due to severe image artifacts. Locations of FA results were obtained from the JHU ICBM-DTI-81 atlas (Mori, Wakana, Van Zijl, & Nagae-Poetscher, 2005) in FSLeyes (an FSL tool). For the same reasons as our VBM analyses, it was not feasible to calculate power for our FA analyses.

The same five basic types of analyses which were run for the VBM analyses were also run for the FA analyses, replacing images of GM volume with images of skeletonized FA, and using FA values in place of GM volume (i.e., the analyses performed were: All FA Regression, FA Adult Regression, FA Adolescent Regression, FA Interaction Contrast, and FA Conjunction Contrast). The motivation for each of these FA analyses are analogous to the corresponding GM analyses. These five types of analyses were also run with and without controlling for IQ as a covariate of noninterest.

Results

Behavioral Results

Behavioral performance data for adults and adolescents were statistically analyzed in RStudio (Ver. 0.98.1103), utilizing the car package as well as the default statistical package; additionally, an online statistical tool, "Calculation for the test of the difference between two independent correlation coefficients," was used (Preacher, 2002). All standard deviations were sample standard deviations calculated using the n-1 method. Statistical power was obtained using the pwr package in RStudio, or by reference to a power table (Cohen, 1988).

Adult IQ scores had a mean of 99.98 ($SD = 14.16$). Adolescent IQ scores had a mean of 92.09 ($SD = 15.28$). A summary of these means and standard deviations are shown in Table 1. A two-tailed pooled Student's t-test between adult and adolescent IQ scores was significant ($t(82) = 2.45, p < 0.001, d = 0.53$). *This indicates that adults had higher overall IQ scores than adolescents.* The power for this t-test is 0.95, assuming a large effect size ($d = 0.8$) and an alpha of 0.05. Power becomes 0.62 when considering a medium effect size ($d = 0.5$).

The assumptions for the above Student's t-test were met as follows. For adolescents, no IQ score exceeded 2.56 standard deviations of the mean. Shapiro Wilk's test of adolescent IQ scores was non-significant ($p = 0.715$), indicating normality. For adults, no IQ score exceeded 2.61 standard deviations of the mean. Shapiro Wilk's test of adult IQ scores was non-significant ($p = 0.177$), indicating normality. Homogeneity of variance in IQ between adults and adolescents was supported by a non-significant Levene's test ($p = 0.432$).

Due to the significant confounding difference in IQ scores between adults and adolescents, pairs of participants from each age group were formed by matching adolescents and adults who were within 3 IQ points of each other. This resulted in 27 pairs of participants (54 total), who were retained for the remaining behavioral analyses. The new set of adolescents ranged from 11.16 to 16.92 (*Mean = 15.02, SD = 1.58)*, and consisted of 14 females. The new set of adolescent IQ scores had a mean of 98.59 (*SD = 14.66*). The new set of adults ranged from 20.00 to 24.33 (*Mean = 21.85, SD = 1.26), and consisted of 12 females.* The new set of adult IQ scores had a mean of 99.19 (*SD = 14.57*). A summary of these IQ means and standard deviations are shown in Table 1. In these participants, IQ did not differ significantly between age groups ($t(52) = 0.149$, $p = 0.882$, $d = 0.04$). The power for this t-test is 0.82, assuming a large effect size ($d = 0.8$) and an alpha of 0.05. Power becomes 0.44 when considering a medium effect size ($d = 0.5$).

The assumptions for the above Student's t-test were met as follows. For adolescents, no IQ score exceeded 2.43 standard deviations of the mean. Shapiro Wilk's test of adolescent IQ scores was non-significant ($p = 0.811$), indicating normality. For adults, no IQ score exceeded 2.48 standard deviations of the mean. Shapiro Wilk's test of adult IQ scores was non-significant ($p = 0.340$), indicating normality. Homogeneity of variance in IQ between adults and adolescents was supported by a non-significant Levene's test ($p = 0.812$).

Adolescent riddle scores had a mean of 0.28 (*SD = 0.19*). Adult riddle scores had a mean of 0.44 (*SD = 0.13*). A summary of these means and standard deviations are shown in Table 1. A two-tailed pooled Student's t-test between adult and adolescent riddle scores was significant (*t(52) = 3.166, p = 0.003, d = 0.86*). *This indicates that adults had higher overall riddle performance than adolescents.* The power for this t-test is 0.82, assuming a large effect size (*d = 0.8)* and an alpha of 0.05. Power becomes 0.44 when considering a medium effect size (*d = 0.5)*.

The assumptions for the above Student's t-test were met as follows. For adolescents, no score exceeded 1.94 standard deviations of the mean. Shapiro Wilk's test of adolescent riddle scores was non-significant (*p = 0.251*), indicating normality. For adults, no score exceeded 2.10 standard deviations of the mean. Shapiro Wilk's test of adult riddle scores was non-significant (*p = 0.163*), indicating normality. Homogeneity of variance in riddle scores between adults and adolescents was supported by a non-significant Levene's test (*p = 0.953*).

For adolescents, riddle scores were not significantly correlated with IQ scores (WISC) (*r(25) = 0.023, p = 0.910*). For adults, riddle scores trendingly correlated with IQ scores (WAIS) (*Pearson's r(25) = 0.37, p = 0.059*). A summary of these correlations are shown in Table 2. These two correlations were not significantly different (*z = 1.258, p = 0.209*) according to an online test which utilized Fisher's r to z transformation to test for the difference between two independent correlation coefficients (Preacher, 2002). *Based on these results, we cannot conclude a difference in the relationship between riddle performance and IQ (i.e. common processes) between adults and adolescents.* The power for the riddle score to IQ score correlation within each age group was 0.78, assuming a large effect (*r = 0.5)* and an alpha of 0.05. These powers become 0.34 when considering a medium effect (*r = 0.3).* The power to detect a difference in correlation of riddle score to IQ between adults and adolescents was between 0.8 and

0.9, assuming a large effect size (*r-r = 0.5)* and an alpha of 0.05. This power falls between 0.25 and 0.5 when considering a medium effect (*r-r = 0.3).*

The assumption of homoscedasticity for the above Pearson correlation was met by a non-significant Breush-Pagan test for adolescents ($p = 0.718$) and adults ($p = 0.633$), indicating a homoscedastic relationship between IQ and riddle scores in both adolescents and adults.

Table 1: Riddle and IQ Scores for Non-IQ-Matched and IQ-Matched Participants

	Mean	SD
Non-IQ-Matched		
Adolescent IQ Scores	92.09	15.28
Adult IQ Scores	99.98	14.16
Adolescent Riddle Scores (% Correct)	0.27	0.19
Adult Riddle Scores (% Correct)	0.43	0.20
IQ-Matched		
Adolescent IQ Scores	98.59	14.66
Adult IQ Scores	99.19	14.57
Adolescent Riddle Scores (% Correct)	0.28	0.19
Adult Riddle Scores (% Correct)	0.44	0.13

Table 2: Riddle and IQ Correlations for IQ-Matched Participants

	Riddle-IQ correlation Pearson's R	P-Value
IQ-Matched		
Adolescents	0.023	0.91
Adults	0.37	0.06

Voxel Based Morphometry Results

For the following VBM results, which are summarized in Table 3 and 4, significance was defined as a p-value < 0.05, and trending was defined as a p-value less than 0.1, but greater than or equal to 0.05. These results were family-wise error rate (FWE) corrected in SPM, which is SPM's statistical correction for multiple comparisons based on Random Field Theory (Flandin & Friston, 2016). It reflects the chance of finding a false positive in each analysis. Such a correction takes into account information about volume searched and smoothness of the data, such that the number of comparisons corrected for is fewer than the number of voxels, which would be the case in a pure Bonferroni correction.

In addition to the main VBM analyses previously described using all 84 non-IQ matched participants, another set of VBM analyses was conducted using the 54 IQ-matched participants analyzed previously in the behavioral analyses. This was done because our adolescent and adult groups were found to differ on IQ scores. The set of analyses with non-IQ matched participants has more participants, and therefore provides greater power to detect and map general insight-structure in the relationships in the brain. However, the analyses with IQ-matched participants provides a more legitimate basis for between-age comparisons of insight-structure relationships because it controls for the confound of IQ-differences between groups. Both are provided since we are interested in both the detection of general insight-structure relationships regardless of the effect of age or IQ, as well as the comparison (i.e. commonalities and differences) of insight-structure relationships between age groups. Also note that analyzing IQ-matched groups serves a different purpose than running the analyses with

IQ as a covariate of non-interest. In the first case, we are trying to improve comparisons between age groups by controlling for IQ differences between groups. In the second case, we are trying to observe the component of insight that is relatively distinct from common processes, and how it relates to brain structure.

Table 3: Main VBM Results for Non-IQ-Matched Participants

Area	BA	Peak MNI coordinates	Peak t score	Cluster size	FWE corrected p-value	Partial r^2 of riddle score on peak voxel volume
All Negative Regression						
L/FOC	11	-10 27 -24	4.06	471	cluster p > 0.1; voxel p = 0.101	0.204
R/ITG	20	57 -27 -22	4.78	1873	cluster p = 0.023; voxel p > 0.1	0.167
Adult Negative Regression						
R/ITG	37	54 -45 -14	5.65	2958	cluster p = 0.001; voxel p = 0.045	0.076
Adolescent Negative Regression						
R/Thal	NA	10 -12 20	5.59	504	cluster p > 0.1; voxel p = 0.041	0.196
L/FOC	11	-10 27 -24	5.26	1330	cluster p = 0.098; voxel p = 0.095	0.217
All Negative Regression -IQ controlled						
L/FOC	11	-10 26 -24	5.03	447	cluster p > 0.1; voxel p = 0.046	0.220
Adolescent Negative Regression -IQ controlled						
R/Thal	NA	10 -12 20	5.59	499	cluster p > 0.1; voxel p = 0.045	0.197
L/FOC	11	-10 27 -24	5.19	1313	cluster p = 0.104; voxel p > 0.1	0.216

BA refers to Brodmann's Area. L/FOC refers to left frontal orbital cortex. R/ITG refers to right inferior temporal gyrus. R/Thal refers to right thalamus (with probable neuronal connections to the temporal cortex as labeled by Anatomy Toolbox). Cluster size is given in number of voxels. P-values are family-wise error corrected.

For our main VBM analyses using all 84 participants (see Table 3 and Figure 5), the All Negative Regression revealed a trending ($p = 0.101$) voxel in the left frontal orbital cortex (L/FOC). This voxel became significant ($p = 0.046$) when IQ was controlled for. The same regression also revealed a significant ($p = 0.023$) voxel in the right inferior temporal gyrus (R/ITG). This voxel became less significant ($p > 0.1$) when

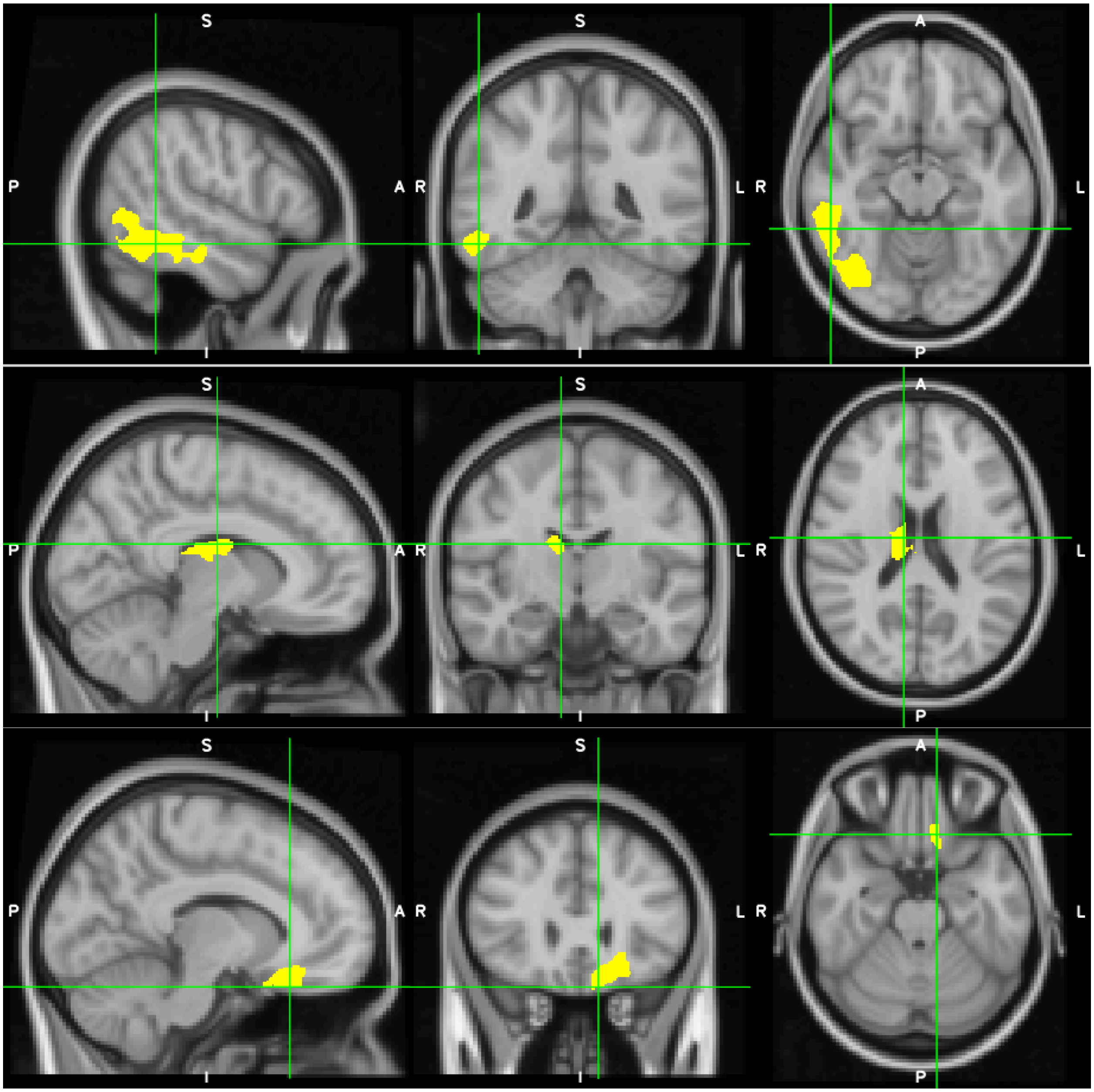


Figure 5: Main VBM Results for Non-IQ-Matched Participants: (Top) A cluster from the Main GM Adult Negative Regression (not IQ controlled) with peak voxel in the right inferior temporal gyrus (Middle) A cluster from the Main GM Adolescent Negative Regression (not IQ controlled) with peak voxel in the right thalamus. (Bottom) A cluster from the Main GM Adolescent Negative Regression (not IQ controlled) with peak voxel in the left frontal orbital cortex. Images are thresholded at p = 0.01 uncorrected. Refer to Table 3 for family-wise error corrected p-values of specific clusters and peak voxels. Peak voxels are indicated by crosshairs.

IQ was controlled for. *The results of this analysis confirms that an insight-volume relationship does exist for riddle-based insight.*

The main VBM Adult Negative Regression revealed a significant ($p = 0.001$) cluster, with peak activation in the R/ITG and extending to the right fusiform gyrus (R/FG) when IQ was not controlled for. This regression became less significant ($p > 0.1$) when IQ was controlled for. *The result of this analysis is consistent with the All Regression analysis in confirming an insight-volume relationship.*

The main VBM Adolescent Negative Regression revealed a significant ($p = 0.041$) voxel in the right thalamus (R/Thal) when IQ was not controlled for. The Anatomy Toolbox in SPM labeled this voxel as most likely having neuronal connections to the temporal cortex. The same regression also revealed a trending cluster ($p = 0.098$) in the L/FOC which extends to the left middle orbital gyrus (L/MOG) and left superior orbital gyrus (L/SOG). When IQ was controlled for, the voxel in the R/Thal and the cluster in the L/FOC became slightly less significant (*$p = 0.045$ and $p = 0.104$ respectively*). *The results of this analysis are consistent with the All Regression Analyses in confirming an insight-volume relationship, and further revealed another region, the R/Thal, where there was an insight-volume relationship.*

There were no significant or trending results for the main VBM Interaction Contrasts. There were also no significant or trending results for the main VBM Conjunction Contrasts.

Table 4: VBM Results for IQ-Matched Participants

Area	BA	Peak MNI coordinates	Peak t score	Cluster size	FWE corrected p-value	partial r^2 of riddle score on peak voxel volume
All Negative Regression						
L/SOG	11	-14 26 -18	5.69	711	voxel p = 0.015; cluster p > 0.1	0.370
Adult Negative Regression						
R/MTG	37	51 -60 2	4.78	1676	voxel p > 0.1; cluster p = 0.016	0.48
All Negative Regression IQ controlled						
L/SOG	11	-14 24 -20	5.59	710	voxel p = 0.021; cluster p > 0.1	0.366

BA refers to Brodmann's Area. L/SOG refers to left superior orbital gyrus. R/MTG refers to right middle temporal gyrus. Cluster size is given in number of voxels. P-values are family-wise error corrected.

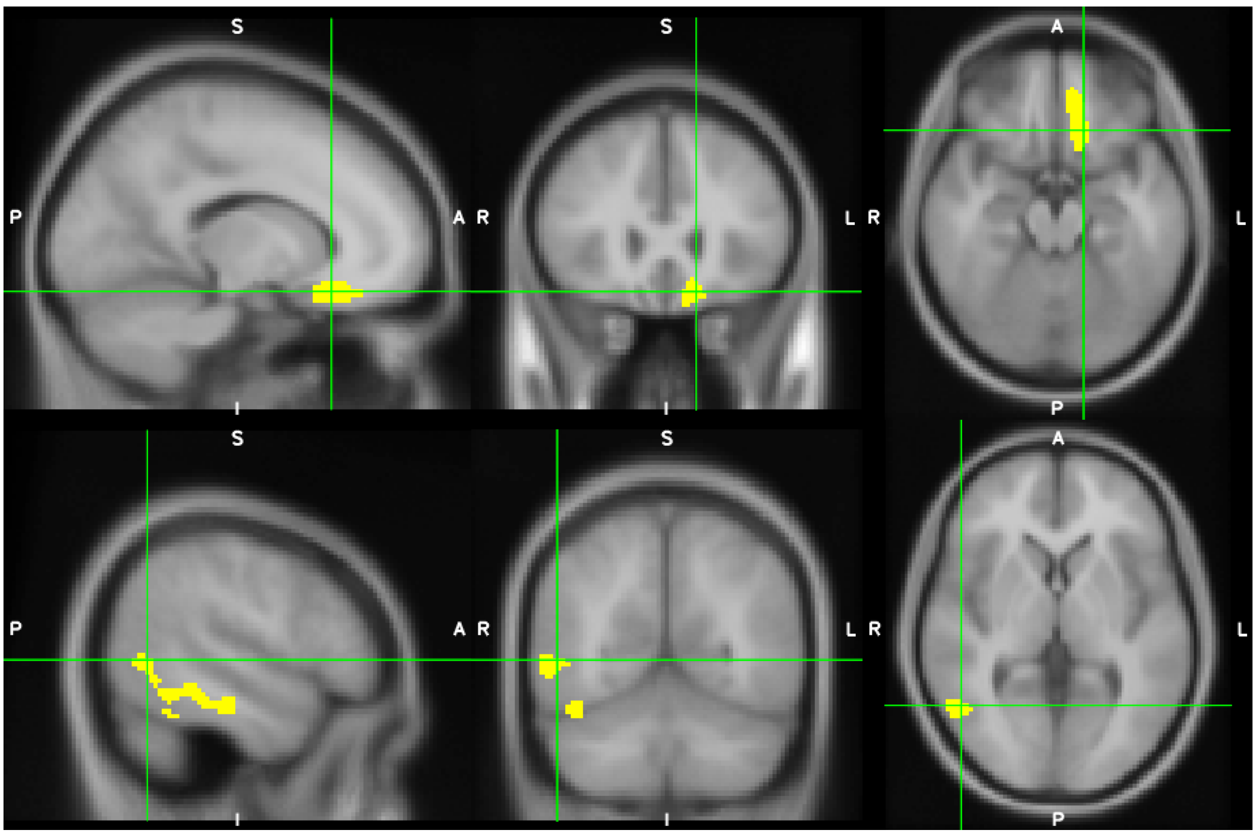


Figure 6 VBM Results for IQ-Matched Participants. (Top) A cluster from the IQ-Matched GM All Negative Regression (not IQ controlled) with peak voxel in the left superior orbital gyrus. (Bottom) A cluster from the IQ-Matched GM Adult Negative Regression (not IQ controlled) with peak voxel in the right middle temporal gyrus. (Bottom). Images are thresholded at p = 0.01 uncorrected. Refer to Table 4 for family-wise error corrected p-values of specific clusters and peak voxels. Peak voxels are indicated by crosshairs.

In general, fewer results were significant when we analyzed the 54 IQ-matched adolescents and adults (see Table 4 and Figure 6). The IQ-matched All Negative Regression revealed a significant voxel in the left superior orbital gyrus (L/SOG) when IQ was not controlled for. The peak voxel in this regression became slightly less significant when IQ was controlled for ($p = 0.021$).

The IQ-matched Adult Negative Regression revealed a significant cluster in the right middle temporal gyrus (R/MTG) when IQ was not controlled for. The cluster in this regression became non-significant ($p > 0.1$) when IQ was controlled for.

No other VBM analyses using IQ-matched adolescents and adults were significant, including the IQ-matched Interaction Contrasts and IQ-matched Conjunction Analysis. *Therefore, we cannot conclude whether any insight-volume relationships are unique to either age group, or common to both age groups.*

Fractional Anisotropy Results

The only significant results (see Table 5 and Figure 7), occurred when IQ was controlled for in the FA Adult Positive Regression. Here, three clusters, with peak voxels in the body of the corpus callosum, left anterior limb of the internal capsule, and right anterior corona radiata reached significance. *This demonstrates that an insight-WM integrity relationship exists.*

None of the FA analyses using IQ-matched adolescents and adults were significant. This includes both the IQ-matched Interaction Contrast, and the IQ-matched Conjunction Analysis. *Therefore, we cannot conclude whether any insight-FA relationship are unique to either age group, or common to both age groups.*

Table 5: Main FA Results for Non-IQ-Matched Participants.

Area	Peak MNI coordinates	Peak t score	Cluster size	FWE corrected p-value	Partial r^2 of riddlescore on peak FA
Adult Positive Regression					
Body of CC	-12 17 25	4.59	187	cluster p = 0.041; voxel p > 0.1	0.369
L Anter Limb of IC	-19 9 10	4.51	202	cluster p = 0.025; voxel p > 0.1	0.361
R Anter CR	24 18 11	4.27	216	cluster p = 0.017; voxel p > 0.1	0.336

Body of CC refers to body of corpus callosum. L Anter Limb of IC refers to left anterior limb of internal capsule. R Anter CR refers to right anterior corona radiata. Cluster size is given in number of voxels. P-values are family-wise error corrected.

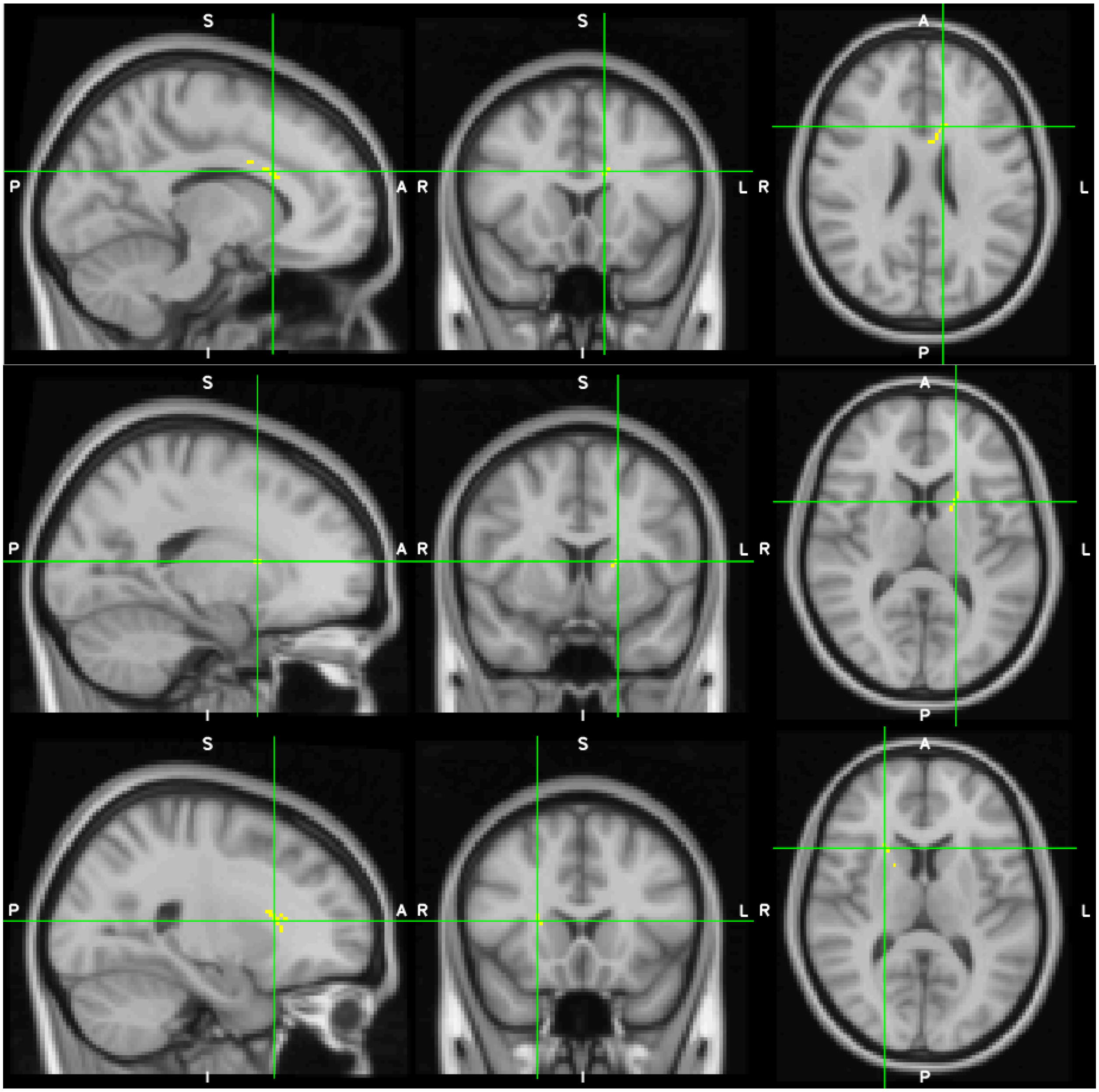


Figure 7: Main FA Results for Non-IQ-Matched Participants: Significant clusters from the FA Adult Positive Regression (IQ controlled), including (Top) a cluster with peak voxel in the body of the corpus callosum, (Middle) a cluster with peak voxel in the left anterior limb of internal capsule, (Bottom) a cluster with peak voxel in the right anterior corona radiata. Images are thresholded at p=0.01 uncorrected. Refer to Table 4 for family-wise error corrected p-values of specific clusters and peak voxels. Peak voxels are indicated by crosshairs.

Discussion

Addressing the Main Purposes of the Study

Past research has shown that insight ability improves from adolescence to adulthood (Kleibeuker, De Dreu, and Crone, 2013). Corroborating this idea, we found that riddle solving was significantly more accurate in adults than adolescents. Past research has suggested that insight critically consists, in part, of common processes (Chronicle et al., 2004; Ball & Stevens, 2009; Chuderski, 2014; Chuderski & Jastrzebski, 2016). Past research has also indicated that linguistic competence (a common process) continues to mature over the course of adolescence (Nippold, 1993; Larson & McKinley, 1998; Nippold et al., 1999; Nippold, 2000; Nippold et al., 2005a, 2005b). Past research has also shown that brain structures which potentially support common processes continue to mature during adolescence (Sowell et al., 1999; Gogtay, 2004; Giedd, 2008; Taki et al., 2013; Lebel et al., 2017). Based on these past findings, we posited that adults have more matured common cognitive processes, than adolescents, which they can apply towards solving insight problems. However, our behavioral analyses do not support this conclusion, as insight was not found to be differently correlated with IQ (common processes) in adults versus adolescents.

Before discussing the implications of our MRI data, we will point out that for conclusions regarding overall insight-structure relationships we will defer to the results of our main analyses due to the greater statistical power it provides us by way of more participants. However, for conclusions regarding age-specific insight-structure relationships, we will defer to the results of our IQ-Matched analyses, which mitigates the confound of differing IQ between age groups, and are therefore more valid in comparing (i.e. detecting commonalities and differences in) insight-structure relationships across our two age groups.

Based on our main FA results, we have confirmed that insight performance is related to macroscopic brain structure in terms of white matter tract integrity. This means that more efficient communication between multiple brain regions is related to insight performance. Specifically, our finding that increased FA in the corpus callosum is related to improved riddle solving performance suggests that inter-hemispheric communication is potentially critical to insight. One reason why both the left and right hemispheres are critical for insight may be that insight starts off with processing dominant concepts, but then shifts to processing non-dominant concepts. The left hemisphere may be specialized for processing dominant concepts, while the right hemisphere may be specialized for processing both dominant and non-dominant concepts. Support for this view of hemispheric specialization comes from a study by Faust and Chiarello (1998), which found that when participants were presented with sentences ending with a word that had multiple meanings (e.g., "He could not wait for even one second."), that sentence-final word primed target words (i.e., reduced their recognition times) related to both its dominant and non-dominant meanings (e.g., "time" and "number") when the target word was presented in the left visual field. This suggested that the right hemisphere (which processes the left visual field) maintained concepts associated with both the dominant and non-dominant meanings of the sentence-final word. In contrast, target words, presented to the right visual field primed only target words related to the dominant meaning of the sentence-final word, suggesting that the left hemisphere maintains only concepts associated with the dominant meaning of the sentence-final word.

In the main FA results, two projection fibers' FA were also found to be related to riddle solving. This means that efficient communication between cortical and subcortical structures is related to insight performance. One possibility is that subcortical structures may be acting as a gateway between cortical structures that serve insight processes.

Based on our significant main VBM results, we have confirmed that insight performance is related to macroscopic brain structure, as measured by GM volume, in three cortical areas. Speculatively, we take the negative direction of this relationship to be caused by synaptic pruning, and consistent with the notions that: smaller GM volume may reduce demands on dendrite integrity, as well as imply fewer but more efficient synapses. This finding also supports the contention of Bendetowicz et al. (2017), and Seyed-Allaei et al. (2017) that inter-individual differences in insight ability manifest themselves in GM volume differences. This finding also joins other studies in the last twenty years that have found relationships between GM volume and cognitive functioning in various domains (Maguire et al., 2000; Draganski et al., 2004; Kanai & Rees, 2011).

Based on our IQ-matched analyses, we were unable to conclude that certain insight-volume or insight-FA relationships are unique to either age group, due to the interaction contrasts being non-significant. Similarly, we were unable to conclude that certain insight-volume or insight-FA relationships are consistent across age groups, due to the conjunction contrasts being non-significant. However, both remain open possibilities, and might be revealed given greater statistical power (e.g., via more participants, or a longitudinal design).

Possible Function of the Right Inferior Temporal Gyrus

A potential role of the R/ITG (as identified in the main set of our VBM analyses) in insight might be the maintenance of non-dominant concepts. In our cognitive model of riddle solving, after the incorrect dominant representation (existing as propositions, and a situational model), which was evaluated as incomprehensible by failing criteria of local and global coherence, is suppressed, semantic retrieval of non-dominant verbal concepts becomes necessary for rebuilding the mental representation of the riddle. The role of the R/ITG could be to maintain the activation of both dominant and non-dominant concepts (either denoted, connoted, or inferred from the riddle) such that when the dominant representation is weakened, the non-dominant concepts remain activated, and are more easily retrieved.

Possible Function of the Left Frontal Orbital Cortex

A potential role of the L/FOC (as identified in the main set of our VBM analyses) in insight might be detecting coherence in the current riddle representation. As specified in our cognitive model of riddle solving, detection of coherence, or lack thereof, is a key part of reaching impasse, and of evaluating whether a solution has been reached with the current set of propositions and situational model. Consistent with this role of the L/FOC are results from Maguire, Frith, and Morris (1999), who analyzed positron emission tomography (PET) images from participants while they read coherent and incoherent stories. They found that a subtraction of coherent minus combined-incoherent story conditions elicited PET activation in the ventromedial orbitofrontal cortex (Brodmann Area [BA]11), and the left temporal pole (BA38). This implicated the orbitofrontal region in some *aspect* of comprehension. However, since this area did not appear in the contrast implicating comprehension, that of an incoherent story (made coherent by a picture cue) condition minus incoherent story (not made coherent by a picture cue) condition, the ventromedial orbitofrontal cortex was not implicated in coherence building per se. Rather, the authors speculated that the orbitofrontal cortex may be involved in *rewarding* increased coherence. By this explanation, the L/FOC would not be active in producing coherence, but rather in making it more salient when it is achieved (detection rather than production). Also consistent with this role of the L/FOC, Ferstl, Rinck, and von-Cramon (2005) found that fMRI activation from stories with temporal inconsistencies minus stories which were temporally consistent revealed activation in the bilateral inferior frontal gyri (orbitalis) close to the L/FOC cluster in this study.

Possible Function of the Right Thalamus

A potential role of the R/Thal (as identified in the main set of our VBM analyses) in insight might be to utilize feedback from failed coherence detection in the frontal lobe to redistribute activation originating in the temporal lobe away from incorrect dominant concepts towards the correct concepts. As specified in our cognitive model, riddle solving involves both retrieval of appropriate solution concepts as well as evaluation for their coherence within larger representations (either as propositions or a situational model). It would be intuitively plausible if there were to be a location where these two processes could interface (e.g., in the R/Thal). Perhaps messages from the frontal lobe signaling failed coherence (and impasse) might signal the R/Thal to initiate processes that redistribute activation away from the dominant concepts and towards the non-dominant solution concepts, the latter of which may be maintained in the right temporal lobe.

Consistent with this function of the R/Thal, are our findings of two WM projection tracts whose FA values were related to riddle solving. Also, consistent with this function of the R/Thal was the finding that the significant GM region within the R/Thal was labeled as probably having neuronal connections to the temporal lobe by the Anatomy Toolbox in SPM12.

Also consistent with this function of the R/Thal, is a study by Behrens et al. (2003), which used DTI to map connections between the medial-dorsal thalamus to both the temporal lobe and prefrontal cortex. The anterior limb of the internal capsule was found to form a probable pathway between the medial-dorsal thalamus (close to our thalamus location) and the prefrontal cortex.

However, an alternative explanation for the function of the R/Thal, is that the R/Thal may simply pass on inhibitory signals from the frontal cortex to suppress dominant incorrect interpretations in the temporal cortex. But whether the R/Thal passes on excitatory or inhibitory messages, the position of the thalamus within the brain makes it likely that it functions as a bridge between other regions involved in riddle solving.

Comparing Location of Results Between Structural Insight Studies

The riddles used in our study, the variant of the RAT utilized by Bendetowicz et al. (2017), and the "Nim" problem utilized by Seyed-Allaei et al. (2017) resulted in both similarities and differences in the location of their structural correlates. Speculatively, these locational similarities and differences may reflect cognitive similarities and differences between our tasks (though other factors, such as variations in the image preprocessing pipeline, may also affect the location). What is cognitively similar in all our tasks is that they all involve the restructuring of the initial representation of the insight problem -essentially a mental set shift. Relevant to this, Dosenbach (2007) identified a frontal-parietal control network which is said to be involved in domain-general set initiation or shifting. Consistent with the implication of this frontal-parietal network, both Bendetowicz et al. (2017) and Seyed-Allaei et al. (2017) found structural correlates of insight in frontal and parietal regions (with both studies identifying Brodmann Area 40 in the parietal lobe). In our study we saw that frontal lobe volume was related to insight; and though parietal lobe involvement was not seen in this study, another study by Goel et al. (2015) reported facilitation of riddle performance via transcranial direct current stimulation of the temporo-parietal junction (also Brodmann area 40). Therefore, the implication of a domain-general frontal-parietal control network could be a common thread to our studies.

Admittedly, one weakness in the preceding claim that the insight tasks in this study, Bendetowicz et al. (2017), and Seyed-Allaei et al. (2017) involve the fronto-parietal control network is that the frontal locations found by these studies do not exactly match the locations implicated by Dosenbach colleagues (2017). But it should be kept in mind that at this time we do not know exactly how fMRI findings translate into VBM findings, so a difference in the location of VBM versus fMRI results does not necessarily preclude the roles we have suggested for these locations revealed by VBM, as long as we are not making a strong claim that these locations *must be* the fronto-parietal control network and nothing else.

What is cognitively different between our studies is that our tasks invoked potentially different representations and different mechanisms of restructuring them (see section "Other Insight Tasks"). That is, the RAT appears to process mainly conceptual level representations in the verbal domain, restructuring associational relationships between them; and the Nim problem appears to deal mainly with transformations in the mathematical/spatial domain, restructuring them using change invariance. Either cognitive difference, representation type or restructuring mechanism, could potentially cause us to see the involvement (as identified in the main set of our VBM analyses) of the R/Thal and R/ITG, which was not seen in these other two studies. For example, one possibility might be that the R/ITG is invoked uniquely in riddles because it maintains semantic concepts destined for combination into higher level verbal representations such as *propositions or situational models*. It might also be the case that the R/Thal is uniquely invoked by riddles because it is involved in restructuring *rule-based* relationships among concepts rather than associational relationships among concepts.

Conclusion

In conclusion, we demonstrated that insight ability, as measured by riddle solving performance, was related to GM volume and FA -two macroscopic measures of brain structure. However, we were unable to conclude whether any of these insight-structure relationships were unique to each age group, or common to both age groups, though they remain open possibilities. Furthermore, given the results of our behavioral analyses, we were unable to conclude that there are any cognitive differences in the way adults and adolescents solve insight problems.

Since the structural neuroimaging techniques utilized in this study are more appropriately utilized for studying relatively stable behavioral measures such as overall performance on insight, future research using event-related functional neuroimaging techniques such as fMRI may better clarify the specific moment-to-moment cognitive differences between adult and adolescent insight.

Bibliography

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. (1977). Frameworks for comprehending discourse. American educational research journal, 14(4), 367-381.

Ball, L., & Stevens, A. (2009, January). Evidence for a verbally-based analytic component to insight problem solving. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 31, No. 31).

Bear, M. F., Connors, B. W., & Paradiso, M. A. (Eds.). (2007). *Neuroscience* (Vol. 2). Lippincott Williams & Wilkins.

Bendetowicz, D., Urbanski, M., Aichelburg, C., Levy, R., & Volle, E. (2017). Brain morphometry predicts individual creative potential and the ability to combine remote ideas. Cortex, 86, 216-229.

Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A. M., Boulby, P. A., ... & Thompson, A. J. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature neuroscience, 6(7), 750.

Blakemore, S. J., & Choudhury, S. (2006). Brain development during puberty: state of the science. Developmental science, 9(1), 11-14.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. Handbook of psycholinguistics, ed. by Morton Ann Gernsbacher, 945-84.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. Cognitive psychology, 3(2), 193-209.

Carroll, D. (2008) Psychology of Language 5th Ed. Wadsworth, Cengage Learning.

Chronicle, E. P., MacGregor, J. N., & Ormerod, T. C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. Journal of Experimental Psychology: Learning, memory, and cognition, 30(1), 14.

Chuderski, A. (2014). How well can storage capacity, executive control, and fluid reasoning explain insight problem solving. Intelligence, 46, 258-270.

Chuderski, A., & Jastrzębski, J. (2017). Working memory facilitates insight instead of hindering it: Comment on DeCaro, Van Stockum, and Wieth (2016). Journal of Experimental Psychology: Learning, Memory, and Cognition, 43(12), 1993.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological review, 82(6), 407.

Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. Psychological bulletin, 136(5), 822.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., ... & Schlaggar, B. L. (2007). Distinct brain networks for adaptive and stable task control in humans. Proceedings of the National Academy of Sciences, 104(26), 11073-11078.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: changes in grey matter induced by training. Nature, 427(6972), 311.

Duncker, K., & Lees, L. S. (1945). On problem-solving. Psychological monographs, 58(5), i.

Ekstrom, R. B., Dermen, D., & Harman, H. H. (1976). Manual for kit of factor-referenced cognitive tests (Vol. 102). Princeton, NJ: Educational Testing Service.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage, 25(4), 1325-1335.

Faust, M., & Chiarello, C. (1998). Sentence context and lexical ambiguity resolution by the two hemispheres. Neuropsychologia, 36(9), 827-835.

Flandin, G., & Friston, K. J. (2016). Analysis of family-wise error rates in statistical parametric mapping using random field theory. Human brain mapping.

Fletcher, C. R. (1994). Levels of representation in memory for discourse.

Förster, J., Friedman, R. S., & Liberman, N. (2004). Temporal construal effects on abstract and concrete thinking: consequences for insight and creative cognition. Journal of personality and social psychology, 87(2), 177.

Friston, K. J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. Neuroimage, 25(3), 661-667.

Giedd, J. N. (2008). The teen brain: insights from neuroimaging. Journal of Adolescent Health, 42(4), 335-343.

Goel, V., Eimontaite, I., Goel, A., & Schindler, I. (2015). Differential modulation of performance in insight and divergent thinking tasks with tDCS. The Journal of Problem Solving, 8(1), 2.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., ... & Rapoport, J. L. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. Proceedings of the National academy of Sciences of the United States of America, 101(21), 8174-8179.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse processes. Annual Review of Psychology, 48, 163-189.

Goldstein, G., & Hersen, M. (3rd Ed.). (2000). Handbook of psychological assessment 3rd Ed. Elsevier.

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex-developmental changes and effects of aging. Brain Res, 163(2), 195-205.

Jausovec, N., & Bakracevic, K. (1995). What can heart rate tell us about the creative process?. Creativity Research Journal, 8(1), 11-24.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. Neuroimage, 62(2), 782-790.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., ... & Kounios, J. (2004). Neural activity when people solve verbal problems with insight. PLoS biology, 2(4), e97.

Jung, R. E., Mead, B. S., Carrasco, J., & Flores, R. A. (2013). The structure of creative cognition in the human brain. Frontiers in human neuroscience, 7, 330.

Jung, R. E., Segall, J. M., Jeremy Bockholt, H., Flores, R. A., Smith, S. M., Chavez, R. S., & Haier, R. J. (2010). Neuroanatomy of creativity. Human brain mapping, 31(3), 398-409.

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. Nature Reviews Neuroscience, 12(4), 231-242.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. Cognitive psychology, 22(3), 374-419.

Kintsch, W., & Keenan, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive psychology, 5(3), 257-274.

Kleibeuker, S. W., De Dreu, C. K., & Crone, E. A. (2013). The development of creative cognition across adolescence: distinct trajectories for insight and divergent thinking. Developmental science, 16(1), 2-12.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. Journal of Experimental Psychology: Learning, memory, and cognition, 25(6), 1534.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. Memory & cognition, 29(7), 1000-1009.

Kounios, J., Fleck, J. I., Green, D. L., Payne, L., Stevenson, J. L., Bowden, E. M., & Jung-Beeman, M. (2008). The origins of insight in resting-state brain activity. Neuropsychologia, 46(1), 281-291.

Kounios, J., & Beeman, M. (2014). The cognitive neuroscience of insight. Psychology, 65(1), 71.

Larson, V. L., & McKinley, N. L. (1998). Characteristics of adolescents' conversations: A longitudinal study. Clinical linguistics & phonetics, 12(3), 183-203.

Lebel, C., Treit, S., & Beaulieu, C. (2017). A review of diffusion MRI of typical white matter development from early childhood to young adulthood. NMR in Biomedicine.

Lockhart, R. S., Lamon, M., & Gick, M. L. (1988). Conceptual transfer insimple insight problems. Memory & Cognition, 16(1), 36-44.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. Psychological monographs, 54(6), i.

Luo, J., & Niki, K. (2003). Function of hippocampus in “insight” of problem solving. Hippocampus, 13(3), 316-323.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences, 97(8), 4398-4403.

Mai, X. Q., Luo, J., Wu, J. H., & Luo, Y. J. (2004). “Aha!” effects in a guessing riddle task: An event-related potential study. Human brain mapping, 22(4), 261-270.

Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access.

Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel-based morphometry of the human brain: methods and applications. Current medical imaging reviews, 1(2), 105-113.

Medin, D. L., & Rips, L. J. (2005). Concepts and categories: Memory, meaning, and metaphysics. The Cambridge handbook of thinking and reasoning, 37-72.

Mednick, S. (1962). The associative basis of the creative process. Psychological review, 69(3), 220.

Metcalfe, J. (1986). Premonitions of insight predict impending error. Journal of Experimental Psychology: Learning, Memory, and Cognition, 12(4), 623.

Mori, S., Wakana, S., Van Zijl, P. C., & Nagae-Poetscher, L. M. (2005). MRI atlas of human white matter. Elsevier.

Nippold, M. A. (1993). Developmental markers in adolescent language: Syntax, semantics, and pragmatics. Language, Speech, and Hearing Services in Schools, 24(1), 21-28.

Nippold, M. A., Hegel, S. L., Sohlberg, M. M., & Schwarz, I. E. (1999). Defining abstract entities: Development in pre-adolescents, adolescents, and young adults. Journal of Speech, Language, and Hearing Research, 42(2), 473-481.

Nippold, M. A. (2000). Language development during the adolescent years: Aspects of pragmatics, syntax, and semantics. Topics in Language Disorders, 20(2), 15-28.

Nippold, M. A., Hesketh, L. J., Duthie, J. K., & Mansfield, T. C. (2005a). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. Journal of Speech, Language, and Hearing Research, 48(5), 1048-1064.

Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005b). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. Language, Speech, and Hearing Services in Schools, 36(2), 125-138.

Ollinger, M. & Knoblich, G. (2009). Psychological research on insight problem solving. In Recasting reality (pp. 275-300). Springer Berlin Heidelberg.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(4), 791.

Patrick, J., Ahmed, A., Smy, V., Seeby, H., & Sambrooks, K. (2015). A cognitive procedure for representation change in verbal insight problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 41(3), 746.

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. Journal of Youth and Adolescence, 17(2), 117-133.

Preacher, K. J. (2002, May). Calculation for the test of the difference between two independent correlation coefficients [Computer software]. Available from http://quantpsy.org.

Qiu, J., Li, H., Yang, D., Luo, Y., Li, Y., Wu, Z., & Zhang, Q. (2008a). The neural basis of insight problem solving: An event-related potential study. Brain and cognition, 68(1), 100-106.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. Journal of verbal learning and verbal behavior, 17(4), 403-417.

Redish, A. D. (2001). The hippocampal debate: are we asking the right questions?. Behavioural brain research, 127(1), 81-98.

Rinck, M., Hähnel, A., Bower, G. H., & Glowalla, U. (1997). The metrics of spatial situation models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(3), 622.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1994). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind hypothesis.

Seyed-Allaei, S., Avanaki, Z. N., Bahrami, B., & Shallice, T. (2017). Major thought restructuring: The roles of different prefrontal cortical regions. Journal of cognitive neuroscience, 29(7), 1147-1161.

Smith, S. M., & Blankenship, S. E. (1989). Incubation effects. Bulletin of the Psychonomic Society, 27(4), 311-314.

Smith, R. W., & Kounios, J. (1996). Sudden insight: All-or-none processing revealed by speed–accuracy decomposition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1443.

Soares, J., Marques, P., Alves, V., & Sousa, N. (2013). A hitchhiker's guide to diffusion tensor imaging. Frontiers in neuroscience, 7, 31.

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. Nature neuroscience, 2(10), 859.

Sternberg, R. J., & Davidson, J. E. (1995). The nature of insight. The MIT Press.

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., & Kawashima, R. (2010). White matter structures associated with creativity: evidence from diffusion tensor imaging. Neuroimage, 51(1), 11-18.

Taki, Y., Hashizume, H., Thyreau, B., Sassa, Y., Takeuchi, H., Wu, K., ... & Fukuda, H. (2013). Linear and curvilinear correlations of brain gray matter volume and density with age using voxel-based morphometry with the Akaike information criterion in 291 healthy children. Human brain mapping, 34(8), 1857-1871.

Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.

Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems.

Wilke, M., Holland, S. K., Altaye, M., & Gaser, C. (2008). Template-O-Matic: a toolbox for creating customized pediatric templates. Neuroimage, 41(3), 903-913.

Wechsler, D. (2014). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). San Antonio, Texas: Psychological Corporation.

Wechsler, D. (2004). WISC-IV: Wechsler Intelligence Scale for Children: Technical and Interpretive Manual. Psychological Corporation.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological bulletin, 123(2), 162.

Appendix

Riddle Set (Originally in English; Translated to Spanish for the study)

1. María y Milagros nacieron el mismo día del mismo mes del mismo año, de la misma madre y el mismo padre, pero no son gemelas ni mellizas. ¿Como es posible?

Marsha and Marjorie were born on the same day of the same month of the same year to the same mother and the same father - yet they are not twins. How is that possible?

Solution: They are triplets

2. Tres mujeres- Juana, Daniela y Silvia- tienen entre ellas tres hijos – Samuel, Teresa, y David. A Samuel le gusta jugar con el hijo de Daniela. Silvia de vez en cuando cuida a los niños de Juana. ¿Quién es la madre de Teresa?

Three women - Joan, Dana, and Sandy - have among them three children - Sam, Traci, and David. Sam likes to play with Dana's son. Sandy occasionally baby-sits for Joan's children. Who is Traci's mother?

Solution: Joan

3. Un niño jugando en la playa tiene 6 montones de arena en una zona y tres en otra. Si los pone todos juntos, ¿cuántos montones tendría?

A child playing on the beach has 6 piles of sand in one area and three in another. If you put them all together, how many heaps would you have?

Solution: One

4. Un arqueólogo informó del hallazgo de una moneda romana con la imagen de Julio César en ella, con fecha del 21 antes de Cristo. Otro arqueólogo correctamente afirmó que el hallazgo era un fraude. ¿Por qué?

An archaeologist reported the discovery of a Roman coin with the image of Julius Caesar in it, dated 21 BC. Another archaeologist correctly stated that the finding was a fraud. Why?

Solution: Prior to the start of B.C. coins were not minted with B.C. – it wasn't a known abbreviation

5. ¿Es legal para un hombre casarse con la mujer de su viuda? ¿Por que?

Is it legal for a man to marry his widow's wife? Why?

Solution: No, if his wife is a widow then he is dead and a dead man can not legally marry.

6. Un agricultor posee un hermoso peral. El suministra la fruta a un supermercado cercano. El dueño ha llamado al agricultor para ver la cantidad de fruta está disponible para ser comprada. El agricultor sabe que el tronco principal tiene 24 ramas. Cada rama tiene exactamente 6 ramitas. Dado que cada ramita tiene una pieza de fruta, ¿cuantas ciruelas será capaz de entregar el agricultor?

A farmer has a beautiful pear tree. He supplies the fruit to a nearby supermarket. The owner has called the farmer to see how much fruit is available to be purchased. The farmer knows that the main trunk has 24 branches. Each branch has exactly 6 twigs. Since each twig has a piece of fruit, how many plums will the farmer be able to deliver?

Solution: None, it was a pear tree.

7. Hay un antiguo invento que aún se usa en muchas partes del mundo y que permite a la gente ver a través de las paredes. ¿Que es?

There is an ancient invention that is still used in many parts of the world and that allows people to see through the walls. What is it?

Solution: A window

8. Un mago decía ser capaz de lanzar una pelota de ping pong y hacer que esta avanzara un tramo, se parara y volviera sobre si misma. Además, añadió que no iba a botar la pelota contra ningún objeto ni atar la pelota a nada. ¿Cómo podía realizar esta hazaña?

A magician claimed to be able to throw a ping pong ball and make it move a stretch, stop and come back on itself. In addition, he added that he was not going to throw the ball against any object or tie the ball to anything. How could he perform this feat?

Solution: He threw it up in the air

9. El profesor Benitez se está volviendo más despistado con la edad. Un día, de camino a una clase, se saltó un semáforo en rojo y giró en una calle de un sentido en dirección prohibida. Un policía observó la escena completa pero no hizo nada. ¿Por que el policía no hizo nada?

Professor Bumble, who is getting on in years is growing absent minded. On the way to a lecture one day he went through a red light and turned down a one way street in the wrong direction. A policeman observed the entire scene but did nothing about it. How could Professor Bumble get away with such behavior?

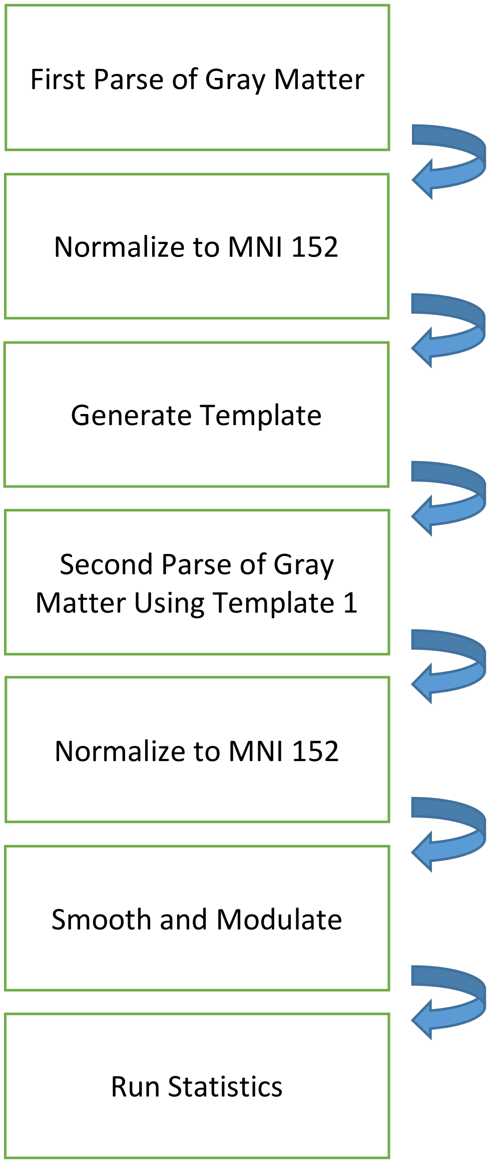
Solution: He was walking

10. Si un barco, durante la marea baja, tiene 6 de sus 12 peldaños de la escalera en el agua. ¿Cuántos peldaños de la escalera estarán en el agua con la marea alta?

If a boat, during low tide, has 6 of its 12 rungs of the ladder in the water, how many rungs on the ladder will be in the water at high tide?

Solution: 6 - the boat floats.

Overall VBM Workflow



Overall TBSS FA Workflow