Durham E-Theses

Analogical Reasoning and Working Memory

ROBSON, ADAM,JAGO

Abstract

Analogical Reasoning (AR) is the ability to find a relationship between two objects that is not based on featural (attribute-based) similarities. As such, reasoning by analogy is thought to be crucial in learning and scientific discovery.

Analogies have played an important role in the conceptualisation of both IQ (Spearman, 1927) and cognitive development (Piaget, Montangero & Billeter, 1977). Yet very little is understood regarding the component processes which underlie analogical thought. Recently, there has been a resurgent interest in the field: one brought about by modern computational methodologies which purport to model the cognitive architecture of analogical thinking. A prominent feature has been the introduction of capacity based processing constraints claimed to arise in the reasoning processes from limited Working Memory Capacity (WMC) resources (Halford, 1992, 1993, 1998; Hummel & Holyoak, 1997; Morrison, Doumas, & Richland, 2011; Richland, Morrison & Holyoak, 2004, 2006, 2010).

Adopting a Working Memory (WM) perspective (Baddeley & Hitch, 1974; Baddeley 2000) the aim of this research is to investigate whether individual differences in WM mediate AR, as well as critically assessing the current theories of AR in relation to this.

In chapter 1 the research behind AR-WM is reviewed with reference to modern interpretations of what analogy is and how it might be measured.

In chapter 2 (Experiment 1), a flexible new scene-based measure of analogical ability, the Richland Picture Analogies (RPA; Richland, et al, 2004, 2006) is introduced, the data confirming effects of complexity and distraction hypothesized by Richland and her colleagues. Experiment 2 related performance on the RPA with quantitative measures of WM, concluding that IQ was related to relational responding in the RPA over and above that of WMC. Experiment 3 further explored the role of WM, observing an effect of processing/storage (WMC) but not storage (STS).

In chapter 3, the role of WMC was further examined. Experiment 4 using a reaction time (RT) paradigm demonstrated that featural responding was unlikely to be a prepotent response, and instead related to conflict resolution. Experiment 5 adopted a dual-task methodology and attempted to explore the involvement of WMC under load in conditions of complexity and distraction. Unfortunately, the low level of variance proved an insurmountable problem. Experiment 6 examined Executive Functions (EFs) as a potential explanation for both IQ and WMC effects in the RPA. Overall, it is concluded that WM does indeed mediate analogical performance within the RPA, but that effects of relational-complexity, as suggested by Halford (1992, 1993, 1998) are not as evident as might have been supposed.

Instead the data from Experiments 2-6 suggests that individual differences in processing efficiency as well as the ability to divide and control attention in novel circumstances may explain the variance in relation responding

reported by Richland et al. (2004, 2006) and found in Experiment 1. It is

hypothesized that one of the core aspects of AR is task relevance, the

research concluding that other interpretations of how WM affects AR

should be considered beyond the traditional theories.

Working memory and analogical reasoning

One volume

Adam Robson

Submitted as part of the doctoral program (PhD)

University of Durham

Department of Psychology

2012

κα`ὶ μὴν ἀριθμόν, ἔξοχον σοφισμάτων,
ἐξηῦρον αὐτοῖς, γραμμάτων τε συνθέσεις,
μνήμην ἁπάντων, μουσομήτορ᾽ ἐργάνην.

"...and yes, I invented for them numbers, too, the most important science;

and the stringing up of letters, the art of Memory, the mother of the Muses."

(Prometheus: on memory as the source of creative wisdom)

- Aeschylus, Prometheus Bound, Line 459.

"Whether or not [we] talk of discovery or of invention, analogy is inevitable

in human thought because we come to new things in science with what

equipment we have, which is how we have learned to think, and above all

how we have learned to think about the relatedness of things."

-Oppenheimer (1956), Analogy in science, p. 129-130

Table of contents

List of tables

Analogical reasoning

Experiment 1

Experiment 2

Experiment 3

Experiment 4

Experiment 5

Experiment 6

11

List of illustrations

List of Common Abbreviations

AR: Analogical reasoning

CE: Central Executive

EB: Episodic Buffer

EF: Executive Function

F^s**:**   Featural Score (number of featural responses)

PL: Phonological Loop

RT: Reaction Time

R^s**:**   Relational Score (number of relational responses)

SMT: Gentner's Structural Mapping Theory

VSSP: Visuo-Spatial Sketch Pad

WM: Working Memory

WMC: Working Memory Capacity

Dedication

To my parents, Barry and Margaret Robson, for making my greatest life ambition come true, and my dream of studying at the University of Durham a reality. And to my fiancée Chiara who kindly put up with me for so long, whilst I was buried in my books and statistics and whom gave so much support when things looked impossible.

Chapter one

1.0 Introduction

Reasoning by analogy has been described as "pervasive in everyday experience" (Sternberg, 1977, p. 353); a "crucial factor in knowledge acquisition at all ages" (Brown, 1989, p. 370).

> "Most of our ordinary mental work - that is, our commonsense reasoning - is based more on thinking by analogy" (Minsky, 1988, p. 329)

It is generally believed that AR is essential to cognitive growth, facilitating learning by allowing us to understand difficult concepts, solve novel problems, discover new facts and make scientific discoveries (Chen, 1999; DeLoache, Miller & Pierroutsaks, 1998; Donnelly & McDaniel, 1993; Gick & Holyoak, 1980; Glynn, 2007, 1995; Glynn & Takahashi, 1998; Goswami, 1994; Hallyn, 2000; Col, France & Taylor, 2005; Hayes & Tierney, 1982; Holyoak, Junn & Billman, 1984; Idling, 1997; James & Scharmann, 2007; Novick, 1988. For reviews see Gentner, Holyoak & Boicho, 2001; Goswami, 1991; Vosniadou & Ortony, 1989).

1.0.1 Origin of the term analogy. The term *analogy* comes from the ancient Greek αναλογια which is also '*analogia'* in Latin but is often translated as '*proportio'* or '*proportionalitas*', 'an understanding of proportionality' (Ashworth, 2008). The stem for analogy is probably from the root word λόγος [reckoning/relation/explanation/debate/verbal-statement/subject-matter] (Liddell & Scott, 1996).

1.0.2 Defining analogical reasoning. Despite AR's perceived importance in cognition, there is considerable debate as to how the processes underlying AR might be represented, analogy being famously described as both "a notoriously difficult term to define" (Goswami, 1991, p. 1) and "a fuzzy concept that means different things to different people." (Dejong, 1989, p. 346).   Such criticisms are not unusual in psychological theory however, and these statements remain true of other scientific concepts, notably intelligence, which as Dejong (1989) points out is still considered a valid concept.   Consequently, the first problem facing any analogical researcher is identifying aspects of the term 'AR' which most theorists can agree upon.

One such point appears to be that the core of AR is, at least in part, the ability to find a similarity between two or more objects/arguments, a process known as Relational Reasoning (RR) and which may be seen as the more

Standard

Pair A   Pair B
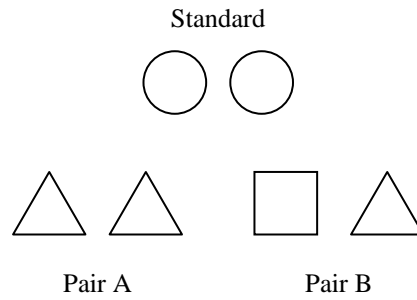
Figure 1. Gentner's representation of analogy (Gentner, 2001). Pair A is the same as the standard, but Pair B is not. The relationship is not based on the standard having the same shape (or 'featural similarities') as Pair A.

general process of identifying (any) relations between objects, regardless of outcome or process[1].

AR appears disparate from RR in that, in an analogical comparison, the relationship that is utilized is supposedly not based upon physical properties[2] (see Figure 1). A relationship based upon observable physical attributes is known as a 'featural' relation, whilst a relationship based upon non-physical attributes is a 'relational' relation.

---

[1] Here the term 'object' implies a conceptual chunking of meaningfully bound information. 'Relations' can be identified as "predicates that link two or more arguments" (Markman & Gentner, 1993).

[2] An idea that shares much in common with Plato's theory of forms (Watt, 1997) and the concept of lower order (a posteriori) as opposed to higher order (conceptualist, or a priori) thinking. Plato described ideas of things (forms) as being the highest form of reasoning possible, whilst observations made about the material world were illusions", made from metaphorical shadows (Watt, 1997), an idea expanded upon by Piaget in his stage theory of analogical development.

From the literature, whether a featural relationship is required for a problem to be 'analogical' remains fuzzy and defining the boundaries of featural similarity is an obvious step in defining analogical thinking. However, as has already been indicated, a number of radically different perspectives of AR exist, which have led to competing definitions of analogy.

1.1 Measurement of analogical reasoning.

Traditionally AR has played a key role in the development of psychological theory, some theorists going so far as to postulate that AR either encompasses "all of" (Spearman, 1923, p. 66), or is a "central component of" (Goswami, 1992, p. 1) human cognition. It is therefore perhaps not surprising to discover that the term has long been understood to be synonymous with the concept of intelligence (Sternberg 1977; Spearman, 1923), with AR tasks making up core components of early measures of general ability or 'g' (Raven, 1938; Miller, 1947).

Today the most frequently encountered measurements of analogical ability typically can be divided into three domains. Those using a variation of the classical analogy (probably the most common form of analogical test), those using variations on the problem analogy format, and a third paradigm, scene based analogies has recently seen a growing interest among researchers. It is this methodology that will be adopted for this thesis.

 1.1.1 The classical analogy. Names most commonly associated with analogy include Plato, Aristotle, Aquinas, Francis Bacon, and John

Stewart Mill as well as numerous others (see Shelly, 2003 for a review), their philosophies forming fundamental concepts which have helped define the direction of analogical research to this day.   However it is from Aristotle and his analysis of analogical structure that we define what is known as the classical analogy which is described by Collins and Burstein (1989) as a 4 figure model using 2 pairs of arguments/objects.

These object-pairs are called 'terms', the individual objects/arguments being called 'figures': hence in the classical format there are four figures (a, b, c and d) and two terms comprising of two figures each (term i: with figures a & b, and term ii: with figures c and d: being written a:b::c:d), the objective being to imply one of the figures through the induction of the other three by correctly identifying the best fitting relation between terms (what Gentner, 1992 calls the similarity-constraint).   The classic analogy becomes a test of reasoning when participants are asked to identify one or more missing term/terms, the missing term (in most cases this is usually only the 4th 'd' term) is known as the 'target term.' Answers are either i) open choice where participants are required to make up their own term without help, or ii) multiple (restricted) choice where participants are given a limited number of options, most commonly 4. The classical analogy is exemplified by the U.S based Miller analogies test (Miller 1947), which uses the multiple choice, 4 term, 4 option verbal analogies.

Classical analogies of both forced and open choice formats can also be divided into quasi or non quasi analogies (Levinson & Carpenter, 1974). In

quasi analogies (which according to Levinson and Carpenter, are easier), the analog is given to the participant (i.e. a bicycle is steered with handlebars, but a boat is steered with…) however in non quasi analogies (bicycle is to handlebars as rudder is to…) the similarity-constraint needs to deduced independently.

From the classical analogy format four different types of analogy may be commonly derived: that of the verbal analogy (using words for each term), the pictorial analogy (using pictures for each term), pattern based analogies (using shapes and/or proportions for each term.), and numerical analogies (using numbers or formula for each term).   Verbal and pictorial analogies are usually the most commonly encountered out of all classic analogy types.

Some similarity-constraints are apparently easier than others, particularly antonyms and functions (Sternberg & Nigro, 1980; Willner, 1964), however generally the difficulty of the classical analogy framework may be altered by varying which term is the target, increasing the number of terms, increasing the number of target terms, and for multiple choice questions varying the number of possible answers. Multiple choice (forced) answers can be made more difficult by increasing/decreasing the similarity between the options available (Piaget et al., 1977), similar but not 'correct' objects creating what are known as 'distracters', an important term within AR, and which usually have featural similarities to the target but not relational ones.

1.1.2 The problem analogy.   The problem analogy follows

directed, situational, goal-driven reasoning, where appropriate behaviour is

abstracted through the cueing of a 'base' situation to a novel 'target'

situation. Usually through a story in which the initial conditions, legal

operations, and goal states of a problem are explicitly specified in one

domain before being applied in another (Glick & Holyoak, 1980, Holyoak et

al., 1984): i.e. having to listen to a story and then complete a physical tool

based task to replicate the 'lesson'.   An example of this is Duncker's

(1945) famous radiation problem[3], or Glick and Holyoak's (1983) 'Red

Adair' tasks[4].

Typically however, success at problem analogies depends not on whether

the participant can come up with a solution, but whether they can

successfully transfer the (experimenter) intended solution from the base to

the target problem, therefore it is not a strict measure of creativity but of

interpreting experimental demand.

---

[3]Where a cancer must be removed via targeted radiation rays, but the rays themselves will
be fatal unless applied in smaller doses. The solution is to administer several smaller rays,
rather than the one big one. This solution is then applied to new situations, such as Glick &
Holyoak's (1983) 'Red Adair' tasks.
[4] I.e. 'The General' problem (where a number of armies must attack a fortress down several
mined roads in order to be successful), 'The Commander' problem (where an island must
be attacked over several bridges at once), 'The Fire Chief' problem (where water must be
applied to save a house all at once rather than one bucket at a time) and the 'The Red Adair'
problem itself (where a fire could only be extinguished if enough foam was used through
lots of smaller hoses as no large hose was available).
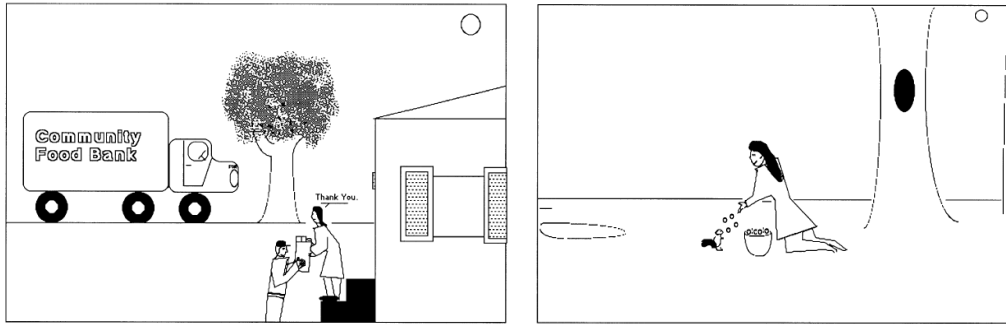
Figure 2. The cross mapping task (Markman & Gentner 1993).

Participants are asked to cross map an object from the base scene (left) onto

the target scene (right). In this case the woman in the base can be mapped

onto the woman in the target (a featural answer) or the squirrel (a relational

answer) as they are both receiving food.

1.1.3 Scene based analogies (cross mapping paradigm). The

'scene based analogy' task was originally developed by Markman and

Gentner (1993) and expanded upon by other researchers such as Waltz et al.

(2000) and Richland et al. (2006). This form of analogy uses a visual cross

mapping paradigm with pictorial scenes as both the base and target

problems (Figure 2). Participants are shown a 'target object' in the base

scene and asked to identify the 'best fitting relation' from a list of objects in

the target. In scene based analogies, mappings (representations of relations)

can be increased or decreased by adding interacting objects to pictures.

In Markman and Gentner's paradigm there are both featural answers

(answers that physically resemble the target object) and relational answers

(answers that rely on attributes other than appearance). These can both be

represented simultaneously within the same scene if desired (Morrison et

al., 2010; Richland et al., 2006), with featural answers taking on the role of 'distracters'.

Scene based analogies have the advantage over classical analogies in that they better define the relational interactions intended by the experimenter and isolate featural distracters within a relevant context, featural objects being more obvious.

Research using this paradigm has shown that younger children are less likely to choose relational answers (as opposed to featural ones) than their older siblings (Richland et al., 2004; 2006; 2010), and that if participants are encouraged to build an integrated representation (Waltz et al., 2000) of the scene, by being asked to map multiple objects from the base onto the target (called multiple mapping as opposed to "one shot" mapping where the participant is asked to map just one object), then more relational answers will also be chosen.

Table 1. An example of Piaget's classical analogies (Piaget et al., 1977).

Bicycle: Handlebars: Ship [Rudder/Sail/Bird]

If the participant chooses sail as the d term, is it because they are associating

the term with a ship, or is it because they understand that sails can also be

manipulated in order to change the direction?

1.1.4 Fundamental problems within analogical testing.

(i) The associative dilemma. According to Willner (1964), most tests of

classical analogy do not follow a pre-determined set of rules for dealing

with correct answers not anticipated by the researcher (a fact apparently true

even today), thus may not be measuring analogical thinking.   Instead it is

common for participants to use what Gentile (1977) calls an 'associative

mechanism': what Goldman and colleagues (Goldman, Pellegrino,

Parseghian & Sallis, 1982) describe as a simpler associative understanding

of the classical analogy.

This works when we analyse just the c and d figures in an analogy and

produce an answer based on them alone. As Willner (1964) points out, it is

entirely possible for well over half the number of correct answers in

classical analogies to be derived via purely associative responses from the c

and d terms, without even considering a and b.   According to Wilner it is

possible to create a classical analogy using a rigorous methodology

accounting for the a:b terms, however most researchers do not seem to have adopted this.

(ii) The validity dilemma. In order for a relationship to be analogical we must dismiss non relevant attribute based information, this means that such forms of response must be present. However, what makes this information irrelevant has been the source of much debate, authors such as Goswami (1992, 1991) suggesting that it is in fact the question, not any innate analogical skill, which determines what we are looking for within a relational problem.

This goes against the traditional understanding of analogical thinking which assumes featural responses to be 'undesirable' in analogical tasks. But this may not always be correct, and sometimes the most basic answer may be the best one. As Dedre Gentner (one of the main proponents of featural/relation separation) points out: "If something looks like a tiger, it probably is one". (Gentner, Ratterman & Forbus, 1993, p. 567).

What this means is that if the question is open ended (i.e. "how are these two the same?", a method applied in many early forms of analogical research, most notably Piaget) and no goal state has been decided upon prior to the task, then theoretically any relational answer is correct so long as we ourselves can justify the response given. On many occasions an answer may not have been accounted for by the experimenter yet still fulfil the criteria for higher order thought (Table 1)…and so is marked 'wrong'.

There is no real way of identifying such responses quantitatively; rather we must rely on more traditional qualitative based methods of research, such as asking why a particular answer is chosen. Yet this is sadly a practice adopted by few researchers.

One possible conclusion from these conceptual issues is that analogical 'reasoning' differs from analogical 'thinking'; the former essentially being relational thought that arrives at a specific pre-defined goal-state, which the latter lacks.

1.2 Definitions of analogy

The main theories of AR may be broadly divided into three different domains: Classical Structuralism, Modern Structuralism and Domain Knowledge theories.

    1.2.1 Classical structuralism. For most of last century the idea of structuralism in analogy (i.e. that there was more than one form of reasoning involved in analogies) has dominated analogical research.

Although Piaget can be seen as the definitive structuralist, the idea of structuralism - that relational thought represents a higher form of reasoning than lower featural forms - is an old one that dates back to Plato and Aristotle and the theory of forms.
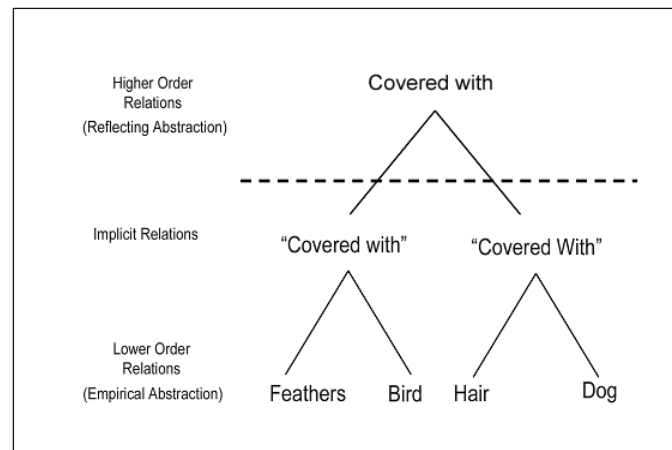
Figure 3. Piaget's hierarchical view of analogical reasoning. Taken from Goswami (1992) and adapted to show Piaget's overlying theory of abstraction.

Here true meaning is not found in actual physical objects, but the idea of objects, comprehension of which is assumed to be extremely difficult. "Truth", as Plato puts it, is "literally nothing but the shadows of the images" (Plato, 1999, p. 68). The structuralist view simply extends this one step further and applies it to reasoning.

Classical structuralism is centred around Piaget et al.'s (1977) use of the classical analogy format (see page 21). According to the structuralists (Inhelder & Piaget, 1958; Lunzer, 1965; Piaget et al., 1977; Piaget & Inhelder, 1969) the process of empirically abstracting a relationship between figures a:b and c:d represents 'lower-order' thought (see Figure 3), requiring less complicated (and early developmental) processes then the 'higher order' establishment of further connections between terms. If the participant

can understand what the similarity-constraint underlying the terms is- then they have achieved true 'higher order' thinking.

To Piaget, analogical thinking is therefore establishing "relations of relations, but without the equality of cross products." (Piaget et al., 1977). A process which Piaget calls 'reflecting abstraction'

 Young children are apparently unable to grasp higher order functioning as they are not cognitively developed enough to be able to comprehend the similarity-constraint (Piaget, 2001; Sternberg & Nigro, 1980). In terms of Piagetian theory this may be because young children are naturally egocentric and cannot expand their viewpoint of an object to include a relationship with which they are not familiar, or which they consider less immediate.

1.2.2 Piagetian structuralist theory. Although Piaget is often quoted as being the definitive structuralist his work is presented here as a subset. This is because Piaget makes a number of assertions about analogical thinking that may be considered particular to his developmental theories. Although structuralism originated through Piaget's work, it has since been applied across scientific and academic disciplines and has gone beyond the claims originally made by Piaget, particularly regarding his stage theory.

Piaget is accredited with the formation of the formal theory of constructivism which proposes that humans create meaning from operative processes by experiencing rules and testing them rather than through observation or passive repetition. To Piaget, rather than being a symbolic system made up of static representations of real world structures, knowledge is a hierarchical and 'active' process. Campbell (2000) interpreted this perspective as knowing what to do with something under certain possible conditions.

Piaget emphasises that alone mere figural representations ('figural' being described as declarative representations which cannot be generalized: Piaget uses perception and language as examples) are not sufficient to explain an organism's world knowledge unless they are intrinsically 'embedded' with other transformational/operative data. This equates to figurative knowledge having very little influence on a child's cognitive development. Instead meaningful 'high-level' knowledge is determined by both application and interaction. [5]Crucial to Piaget's structuralist perspective is the belief that key developmental changes occur in our reasoning ability throughout childhood. Piaget famously categorizing cognitive development into stages (Piaget & Inhelder, 1969). Analogy is a critical component of this developmental framework, with its own developmental timetable.

---

[5]Piaget describes his methodology through his cognitive-structural theory of 'assimilation' and 'accommodation' wherein an individual comprehending the world assimilates it into schemes that must then be accommodated for application with non identical objects or situations (for more information see Campbell, 2000). The key point in Piaget's work is that knowledge is never viewed in isolation, rather it is judged with reference to the individual comprehending the object and how they have accommodated its position in the world.

Piaget and his colleagues used a pictorial version of the classical analogy (see p. 21) in which children of different ages were asked to put together pictures that "seemed to go well together." (Piaget et al., 1977). After first establishing that children were already familiar with all the pictorial terms presented, Piaget's task required the participant to bring together pairs of pictures and then use them again to create an analogical answer that comprised of four pictures in total, mimicking what we now know as the classical analogy. If the participant did not immediately identify an analogy then the problem was metamorphosed into a quasi-analogy with the experimenter verbally specifying three of the terms and querying the participant regarding the fourth (thus prompting an easier solution). If this still did not result in the analogy being successfully solved, then a forced choice solution was given with three possible pictures being used as answers for the 4th unidentified term. When the participant correctly answered, the experimenter then proposed a number of counter-examples (i.e. "Would a bell or a bicycle go as well as the tiller" in the handlebars:bike:tiller:boat analogy).

Piaget noted that by using this methodology three stages of analogical development within AR could be identified, all of which fitted sequentially within his developmental stage theory and which predicted a gradual increase in AR ability over time. Piaget concluded that this was due to the child's growing ability to project qualitative information onto a higher level process where it could be reorganized and analysed for meaning, a process

he termed abstraction réfléchissante (reflecting abstraction): a "higher level" ability not accessible or fully developed at a younger age. According to Goswami (1992) Piaget made three prevalent claims about analogy from this research.

(i) Stage based reasoning. Analogy is a developmentally sophisticated skill. Young children find it difficult to reason analogically before they have passed the required developmental milestone, which for Piaget was the formal operational period of development. This means that a fundamental change occurs around the ages of 11-12 years, wherein children fully understand analogical thinking. Importantly, prior to this, children are usually unable to reason in this way. There is good evidence for this change (Gallagher & Wright, 1977; Piaget et al, 1977) and it is well documented that success in analogy is associated with age (Holyoak et al, 1984; Gentner, 1988).

(ii) Counter suggestions. Piaget noted that generally, older children could dismiss counter suggestions (alternative solutions to the analogy that did not represent what Piaget saw as a correct form of reasoning) far more consistently than younger children. As a direct result, the Piagetian theory of analogy purports that the ability to dismiss counter suggestions is a key developmental stage that must be reached in order for successful AR to develop. If this stage is not reached then analogies can still be solved but without a correct understanding as to why the solution has been reached.

Table 2. *Piaget's stages of analogical development.*

STAGE I:

With the exception of just a few "advanced cases" Stage I (occurring during the preoperational development stage) constituted children aged 5-6 who failed to construct any analogies, the majority of children using either egocentric or associative behaviour to define their answers. Despite this, Piaget was optimistic that children were beginning to be able to reason about relations, suggesting that the second half of this stage (stage IB) was defined from its counterpart by what he called more "stable" relations at the elementary level (i.e. at the 2 picture comparisons, not the 4 pictures).

STAGE II:

Stage II constituted children aged 7-8 and 10-11. Roughly coinciding with Piaget's stage of concrete operations, this stage in analogical reasoning is highlighted by children's willingness to accept counter suggestions as alternate answers. While the children were able to solve the analogies through "groping attempts" (trial and error) they frequently showed an inability to grasp the overall analogy, offering alternate explanations for why pictures go together instead of the one intended by Piaget. In the latter half of Stage II, Stage IIB children appeared to be more willing to reject alternate answers, and also were more likely to show an understanding of the overriding analogy combining terms. However they were still unable to construct them without feedback from the experimenter.

STAGE III:

It was only at Stage III (11-12 years) at the formal operational stage of development that children were able to do this on a more regular basis, with more rational solutions and a more frequent willingness to refute counter suggestions.

(iii) Proportional reasoning. A third claim of Piaget's is that analogies involve proportional reasoning. Piaget implies that all proportions are preceded by an "understanding of the corresponding analogy" (Piaget, 2001) implying that analogies are just "logical proportions". However, the claims appear to have been fitted around his earlier work on stages (where the understanding of proportions plays a large role in defining the formal operational period of development), and appears to have been made in order to cement AR in a wider developmental context. As a result little was written about this particular issue and Piaget's thoughts on proportions and analogies remain largely unsupported.

1.2.3 Modern structuralism. Models of analogy such as those formed by Piaget and his structuralist colleagues have enjoyed moderate success in their field since their original inception (Inhelder & Piaget, 1958; Lunzer, 1965; Piaget & Inhelder, 1969; Piaget et al, 1977). However, towards the end of the last century there was mounting pressure from a cognitive, neuropsychological, as well as computational standpoint, to show that a workable definition of the representations involved in analogical thinking could be achieved through mathematically structured hierarchical systems of predicates that could be modelled either computationally or neurologically.

The ensuing theories are the post-modern structuralists, defined by the work of Dedre Gentner and her colleagues in a direct response to classical structuralism.
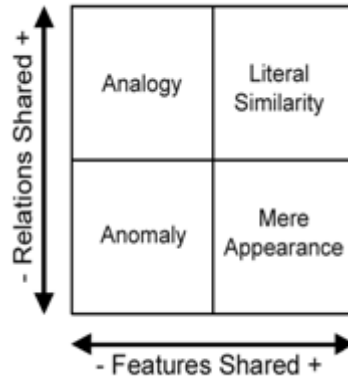
Figure 4. A simplified version of Gentner's Similarity Space (Gentner, 1989). The axis represent whether the entity shares relational or featural similarities, not the number of relations/features.

These are based around two core principles, structural mapping and the relational shift/systematicity principle, both of which are for practical purposes inseparable.

Structural mapping theory. For Gentner, knowledge is defined as a propositional network of nodes and predicates, which interact to form concepts about objects, systems and the world in general. Predicates are descriptive arguments about a concept, the number of predicates defining its complexity (see Gentner's taxonomy of relations, p. 38).

This understanding forms the basis of what is known as Gentner's 'Structural Mapping Theory' of analogical reasoning (SMT,), 'mapping' being defined here as a series of "one-to-one correspondences", the aim of which is obtain the best match possible (Gentner, 1983).

37

"The basic intuition of structural mapping theory is that an analogy

is a mapping of knowledge from one domain (the base) into another

(the target) which conveys that a system of relations that holds

among the base objects also holds true among the target objects"

(Gentner, 1988, p. 43).

Gentner's taxonomy of relations.   The structural aspect of SMT comes

from the Piagetian concept that certain forms of analogical response

represent higher order forms of thought over others. For Gentner, this is

also true, describing what she calls attributional and relational relationships

(see below), and which she defines by the number of predicates associated

with each concept. Yet she clearly refutes the traditional structuralist

assertion that featural responses are cognitively inferior to other forms of

relations. Instead she claims that superficial properties are as useful as their

'higher order' counterparts (Gentner et al., 1993, p. 567), a statement which

has obvious evolutionary validity given how important, immediate and

readily-accessible relational information might be (as Goswami points out

when she highlights the advantage of quickly recognizing a tiger).

Attributional relations are unitary predicates used to describe properties or

entities such as LARGE, RED, SQUARE or FURRY. *Relational* relations

are used to describe events, comparisons, or states applying to two or more

entities or arguments (binary or greater) such as HIT, INSIDE, FASTER-

THAN, or LARGER-THAN (Gentner, 2001).

For example, take the above relational predicate 'LARGER-THAN' when discussing the relationship between a badger and a skunk. In order to understand this we would need i) the predicate i.e. 'x is LARGER-THAN y' ii) the first argument, i.e. 'BADGER' iii) the second argument i.e. 'SKUNK' and iv) the binding of the similarity-constraint to these, i.e. BADGER is larger than SKUNK). This is a binary 'lower order' relational relationship (lower-order relations take objects as arguments, for example HIT [ball, table] and INSIDE [ball, table]) but it could also be reduced to a unitary argument such as 'BLACK and WHITE' which requires only the knowledge of one relational object. Higher order relations such as IMPLIES or CAUSE take other predicates as their argument. For example CAUSE, when discussing pocketing a ball in snooker requires the two predicates HIT [cue, stick, ball] and ENTER [ball, pocket].

This method of categorization has proved to be highly successful computationally, providing theorists with a practical concept of how analogical relations may be represented (Halford, 1992, 1993; 1998; Hummel & Holyoak, 1997, see 'complexity-constraint theory'; p. 89). In SMT Gentner further compartmentalizes analogy, breaking relational reasoning down into several overlapping areas of what she calls 'similarity space' (Figure 4), a hypothetical meta-region dominated by four areas of similarity: analogical, literal, mere appearance, and anomaly. Thus we are able to separate different forms of analogy (relational and featural), but also define the type of dimensions that defines analogy as opposed to similarity

in the real world. 'Analogical' relations are when objects share a majority

of relational but not featural attributes (i.e. a ball and a ship [floats]). Mere

appearance relations are the opposite of analogies when objects share

featural but not relational attributes (i.e. a golf ball and a football

[spherical]). Literal similarities are when objects share both featural and

relational attributes, and so are highly similar in different ways (i.e. a ship's

wheel and a steering wheel [round] & [controls direction]). Anomalies are

so called because they do not share any form of relation and therefore do not

encompass the definition of similarity[6]. Hierarchical presentation of

relationships in this manner has given rise to what may be called

'complexity theory' (the word 'complexity' implying that more processing

power may be required in decoding 'complex' (relational) over 'simple'

(featural) information). I will discuss in detail how complexity may

constrain analogical thinking later.

Structural mapping: the relational shift. Gentner (1983) argues that there

is sufficient support in the literature to show that young children are

inherently limited in their reasoning abilities compared to their older peers.

Using evidence from a series of studies based on the types of answer given,

---

6 According to Gentner (2001) some definitions do not always fit the perspective, namely 'impure similarities' or similarities that cross the boundaries within similarity space. Probably the most frequently associated impure term with analogy is that of the metaphor. Figures of speech such as metaphors or similes are terms frequently (often mistakenly) interposed with analogy, particularly within education circles. Depending on what is being reasoned, figures of speech are forms of relational reasoning that overlap into different categories of similarity but are not ones that should automatically be classed as analogy unless they fulfil certain criteria. This is because figures of speech can be used to describe attributes such as "a football is like a golf ball" [spherical] as well as relational concepts such as "a football is like a boat" [floats].

Gentner showed that younger children (usually aged between 3 and 7 years) are more likely to respond featurally than relationally if given the choice, despite relational responses being preferable overall (Gentner, 1998; Sternberg & Downing, 1982; Sternberg & Nigro, 1980). Gentner refers to the gradual change in preference as "relational shift" but calls the preferential process of selecting relational information as 'systematicity'

"The systematicity principle states that a base predicate that belongs to a mappable system of mutually interconnecting relations is more likely to be imported into the target than is an isolated predicate." (Gentner & Toupin, 1986, p. 280).

Within the mapping process systematicity is crucial, offering a potential resolution to many long-standing questions, particularly as to why analogical ability emerges at very different times in different individuals and environments (Gentner, 1998).

Like the knowledge based accounts (see p.60) Gentner predicts that young children can successfully reason analogically only if the similarity-constraint (i.e. the relationship sought) being used is within a familiar domain (Gentner, 1977a, 1977b, Ratterman & Gentner, 1998), the kind of similarity children can perceive being "determined by the nature of their domain representation, and, in particular, by the amount and kind of relational knowledge they possess in the domain" (Ratterman & Gentner, 1998, p. 455). As children get older their pool of domain-experience

broadens, allowing them to compare objects on levels other than those which are immediately obvious (systematicity). Yet Gentner claims that maturation has nothing to do with this experiential gain, instead drawing a line between general maturation and the experiential accretion of relational knowledge, relevant to the AR problems domain. If they do not have relevant knowledge, they must rely on featural similarity to solve analogies.

1.2.4 The knowledge based account. Driven by the theories of Usha Goswami and Anne Brown (Brown 1989; Brown, & Kane, 1988; Crisafi & Brown, 1986; Goswami 1992; Goswami & Brown 1989), the knowledge based account of AR stems from the idea that the classical structural models of analogy fail to account for analogical success in children who have not yet reached the formal operational stage of development.

Although Piaget himself had stated that children could, and frequently did, solve analogies outside of the later developmental stages (Piaget is often misquoted in this respect) he firmly believed they were doing so mainly by chance, and were often not showing higher order thinking.

Goswami and her colleagues noted that a large body of children are consistently able to reason analogically outside of these boundaries, noting that authors have historically shown that children as young as 4 years and in some cases even as young as 2 and 3 (Crisafi & Brown, 1986) can complete

analogies[7] so long as the analogies are within the domain of their experience and task understanding (Alexander, Willson, White & Fuqua 1987; Gentner, 1989; Goswami & Brown, 1989; Vosniadou, 1989). By giving children relational tasks they can understand and therefore complete (such as analogies based on physical causalities), Goswami and Brown demonstrated that analogical failure in young children results not from a lack of "cognitive competence" (Goswami & Brown, 1989) but a lack of knowledge.

The knowledge based theory therefore proposes that the ability to give analogical responses is available from birth and that structuralists may be incorrect in assuming that you require complex levels of thought in order to solve an analogy. Global age related improvements are not necessary, as domain specific knowledge of the specific relations can account almost entirely for analogical success (Goswami 1992). This is fundamentally intuitive: the first stage in any AR task should be an understanding of what the basic figures and terms are. If we are unfamiliar with any of them, then we will not succeed, or at the very least be able to give an adequate explanation of the correct answer (thereby being unable to show true analogical understanding in the eyes of the experimenter).

---

[7] This depends on whether you class much of the research presented in this section as true analogical studies or not. It is extremely hard to get valid explanations from a notoriously unreliable year group. Brown and Kane (1988) in particular frequently talk about analogical "transfer", the ability to "extract and apply rules over examples"

This can be seen in Piaget's famous handlebars:bike::rudder:boat: analogy. In order to understand this analogy, you will have to have experienced the terms before; but a rudder is often partially below the water line on a boat, and therefore out of sight. In this instance, it would not be unreasonable to suspect that a child who lives near the coast, or who is otherwise familiar with boats, would be better able to solve the analogy than a child who is not. This does not mean that such a child has better analogical skill… just more relational experience.

Since the knowledge based approach has been adopted by post-classical theorists such as Gentner, it is unfair to assume that all non-knowledge based theorists reject the assumptions made above.   Goswami (1992) however, has adopted a self-titled "extreme approach" to the knowledge based theories, one that rejects structuralism entirely.

1.2.5 Relational primacy perspective. Despite being instrumental in forming the modern interpretation of the knowledge perspective, Goswami is keen to distance herself from the mainstream interpretation of relational primacy, offering instead what she calls a more "extreme version of the knowledge based view." (Goswami, 1992).

Like her colleague Anne Brown, Goswami rejects Piaget's claims that only older children can reason analogically, but goes further by suggesting that everything within analogy can be explained by either i) performance factors (i.e. task understanding and meta knowledge of the task), or ii) domain

44

knowledge. Goswami argues that if children do not know or comprehend the analogical problem presented, then their analogical ability (i.e. their knowledge of the relations being used) is hidden, but is not absent.

In order to account for success or failure outside of performance, Goswami proposes what she calls the relational difficulty hypothesis, which states that:

> "…the constraint on the development of AR is the recognition or discovery of relationships in the developing knowledge base, rather than the recognition and use of relational similarity itself."
> (Goswami, 1992, p. 13-14).

Importantly, in relational primacy this process is automatic, the entire process being reliant on whether the child does or does not have the information required to solve the task. To this end Goswami makes an important prediction:

> "If the relations in an analogy are already part of conceptual knowledge, then recognizing their similarity should not constitute an extra cognitive load." (Goswami, 1992, p. 13-14).

Any idea of hierarchical relationships as proposed by the structuralists is therefore meaningless. Designating some relationships as 'higher' or 'lower' order is a purely arbitrary process defined by the experimenter, a

classification that is made only after the solution has been obtained. "Higher order analogies" in the Piagetian sense are always exactly the same as the lower order, any attempt to differentiate them syntactically (or proportionally) being doomed to failure because the relationship can always be reduced to a single argument.

Age based changes in the relational primacy theory. A central feature of relational primacy theory is that it does not predict global age related shifts in analogical ability (i.e. the 'relational shift' hypothesis), which Goswami argues is an artefact of the analogical test.   Goswami is keen to point out that despite this claim, age related changes in ability can and do take place but the change is entirely an ideographic one, centred round whether or not the child has experience of the specific analogy.   The fact that other authors have detected such relational changes is due to either the child's full or partial understanding of the relations involved within the particular analogy.

In cases that are ambiguous, Goswami proposes that children will use other forms of reasoning, falling back on self-descriptive observables (featural relations) or other associations that they have acquired. This accounts for the distraction effect observed by Piaget (1977). Crucially, what an object looks like (i.e. if it is featural or relational) should not affect analogical outcome.

Performance factor account. The performance factor account is Goswami's concession to information-processing and individual

46

differences. She proposes that "meta knowledge" of the analogical task can account for some or all relational thinking. It is based on the work of Goldman and her colleagues (Chen & Daehler, 1989; Goldman et al., 1982) who suggested that a simpler mechanism for understanding analogy existed. This may be by association, but may also include the level of instruction or experience with the specific task.

Analogy is often a term which can only be applied post-hoc after the analogical 'problem' has been solved (Goswami, 1992), implying that most reasoners either do not implicitly understand what is required of them or that participants have not specifically been asked to reason relationally. Performance factors are therefore factors which aid task understanding and the formation of appropriate schemas or rule-sets. What Goswami terms the "facilitation gradient" of "performance factors" (Goswami, 1992, p. 74). For example, being able to interpret an ambiguous goal, interpret experimenter demand, comprehend instruction, hold a goal in mind, or reject alternative responses known to be incorrect.

Brown and Kane (1989) suggest that in many cases analogies are really measuring "meta knowledge". By frequently observing the type of responses available they learn to answer relationally. Of course obtaining this meta knowledge is not in itself an easy task. In many cases analogical problems use confusing, contradictory or vague language (arguably brought about by the confusion surrounding the definitions of the term analogy)

citing words such "patterns[8]" (Richland et al. 2004, 2006), "hints[9]" (Gick & Holyoak, 1980) or "correspondences." (Waltz et al., 2000), which may be confusing to younger participants who need to decipher their instructions before they can adapt to the task.

Similarities and differences between SMT and relational primacy. Despite their apparent differences, both SMT and relational primacy share a number of similar theoretical stances. Both state that domain-specific knowledge is fundamental in AR tasks (without which we cannot fully comprehend what is required of us), and both suggest that featural matching during this process may (in some circumstances) be a fall back strategy, brought about by an absence of prerequisite information. Indeed both theories directly mirror Piaget's concept of experiential learning being crucial to the developmental process; if you have no personal experience of something: you are unable to accomplish the task.

The biggest different between the two theories surrounds processing. For Gentner the focus is on the later active one-to-one mapping of the descriptive information in the base onto candidate targets; the relationships being held in mind and processed for meaning. For Goswami, it is on early processing, specifically the ability to recognize and conceptualize the appropriate similarity-constraint in the base object. Ratterman and Gentner

---

[8] "A certain pattern exists in both the top picture and the bottom picture, and the child's job is to find this pattern." (Richland et al, 2004)

[9] "Subjects in the experimental conditions were told that the first story might give them some hints for solving the test problem" (Gick & Holyoak, 1980)

(1998) coined the term 'relational primacy' to describe the differences between these approaches.

It can be argued that the construct of a schema/rule set and the descriptive complexity of a relational object do in fact represent different approaches to the same concept of strategy management. A rule-set simply describing the overall appropriateness of an object while complexity describes the number of ways in which they are similar.

The main differences between the two theories are that Goswami predicts that the similarity-constraint can be understood by all ages if the child is able to comprehend what is being asked of him/her. Secondly, she predicts that relational-complexity does not load mapping. In the case of the latter, relational primacy predicts that once a similarity-constraint is understood, it is a part of knowledge and the child can find the correct solution regardless of how complex it is. However, recognizing the appropriate similarity-constraint it not always easy and may require a more detailed schema (i.e. further processing[10]) if the target is ambiguous.

---

[10] Because Goswami's focus is on the presence of knowledge and sees mapping as influenced by the performance factors described above, it has received a degree of negative attention from cognitive psychologists who have criticised the Goswami theory for not predicting processing differences in problems where the similarity-constraint is already known (Richland et al, 2004; Ratterman & Gentner, 1998). However it may be more beneficial for researchers to consider that both theories are more similar, viewing the same mapping process from different standpoints with a greater emphasis on different stages of the process.

How ambiguity might be managed is not discussed in Goswami's (1992) theory however, it is suggested that ambiguity is not demonstrated by the number of attributes relational objects may share (one being more appropriate than another in a task and therefore being the correct solution), but the overall goal.

A further point of contention is that of Gentner's relational shift. Goswami argued that such a change could be a function of knowledge, a child progressing from no-knowledge about a similarity-constraint to knowledge. However, Gentner (1998) dismisses such a claim. For Gentner, whilst the relational-shift does involve a change in knowledge, featural-matching is more than an error suggested by Goswami. Instead it is a prerequisite for relational thought and a product of a developing system, the sheer number of featural errors made by younger year groups suggesting a fundamental change in the way relational-reasoning is conducted. As Gentner states, featural responses are "a crucial and necessary step in the progression from comparisons based on overall similarity to comparisons based on relational similarity" (Gentner, 1998, p. 456).

"The relational shift hypothesis predicts a shift from interpretations based on object commonalities to interpretations based on relational commonalities as domain experience increases. The relational primacy view predicts that children should either process relational comparisons correctly from the beginning or, if errors are made before the child possesses

50

adequate domain knowledge, there should be a variety of errors, with no one type predominating." (Gentner, 1998, p. 457).

1.3 Working memory

In this section we will focus on the contribution of working memory (WM) to reasoning and cognitive development. WM refers to the processes in short term memory beyond that of passive storage which includes the cognitive ability to manipulate the information being stored. It is believed to "underpin our capacity for complex thought." (Baddeley, 2007, p. 1) and as such has been implicated in important areas such as child development, learning (for a review see Alloway, 2006) and language comprehension (Gathercole & Baddeley, 2003), diagnosing learning difficulties (Alloway, Gathercole, Willis & Adams, 2004; Swanson, 1994), reasoning and thinking (Logie & Gilhooly, 1998), conscious thought (Baars & Franklin, 2003), and even moral decision making (Moore, Clarke & Kane, 2008), childhood poverty and stress (Evans & Schamberg, 2009) … amongst many others.

One of most widely accepted models of WM is the Baddeley & Hitch model. Developed in 1974 the original tripartite model is a modality specific, resource limited, modular system containing three slave systems: the phonological loop (PL) the visuo-spatial sketch pad (VSSP) and the central executive (CE).
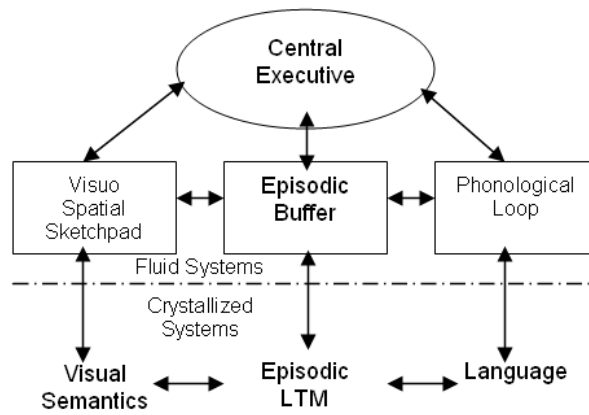
Figure 5. The multiple resource model of WM (Baddeley, 2000). The top section represents the multiple resource model itself, which is based around fluid processing systems. The bottom represents crystallized areas of LTM that are accessed by WM.

Within the model the central executive fulfils the most important role (Gathercole & Baddeley, 2003), its main functions being to both process and store information, as well as regulate the information flow between the slave systems and long term memory. The other systems, namely the PL and VSSP, are maintenance and rehearsal modules which deal with specialised domain specific information, namely verbal and visual/spatial. The WM model is limited by what is referred to as working memory capacity (WMC).

Although much debate surrounds the definition (Engle, 2002) WMC refers to storage and attentional limitations within STM which vary according to the domain and content of the information stored. Baddeley has recently

indicated that these constraints are attentional, however because the WM construct is modular (with evidence for visual and verbal systems) these resources may be further categorized into what will be termed global 'executive' and domain-specific STS (Short Term Store) capacities.

In 2000 a multiple resource model (Figure 5) of WM was introduced (Baddeley, 2000) when a fourth module, the Episodic Buffer (EB) was added to the WM system. The EB is a limited capacity store, linking the CE to LTM, its main purpose being to integrate information into coherent episodes, binding different dimensions of information from the other systems, as well as any form of perceptual input into chunks (such as colour: shape: type etc).

As with most cognitive models Baddeley's model of WM is intended as a framework for research, and as such is not meant to be a detailed representation of neurological networking. As an experimental framework the model has undoubtedly been highly lucrative and has successfully been applied as an investigational tool in a wide range of roles. Part of WM's success has been due to its simplicity; Baddeley himself stating that upon its creation, he expected the model to change, more components to be added[11] and existing modules to be fractionated[12].

---

[11] The exact quotation is: "I assume that the working memory system will ultimately prove to comprise considerably more than two or three subcomponents" Baddeley (1995) p. 114.
[12] Baddeley is quoted as saying "I assume that the central executive can in common with the other components of working memory, be fractionated into subcomponents." Baddeley (2007) p. 119

To an extent this has already happened with other authors adding new insights to the existing structure; such as Logie's visual spatial model (Logie & Pearson, 1997) or Baddeley's own inclusion of the Norman and Shallice (1986) supervisory attention model.

1.3.1 Dual task methodology. Undoubtedly the reason for the success of the WM model has been its ability to make specific predictions based on its subcomponents within a limited-capacity system through the creation of tests designed to overload global or domain specific resources. This method of overloading has become known as the 'dual task methodology'.

Within the dual task methodology participants perform a primary task concurrently with a secondary task, the idea being that if the primary task uses the same cognitive resources as the secondary task then the level of performance at either will significantly drop. Secondary tasks usually involve concurrent processing in specific domains such as continually articulating a word (known as articulatory suppression) or remembering a series of letters (DeStefano & LeFevre, 2003). As such the dual task methodology provides "the most compelling evidence on the specific processes involved in the task of interest because they can be used to isolate the roles of the different working memory components." (DeStefano & LeFevre, 2003, p. 3)
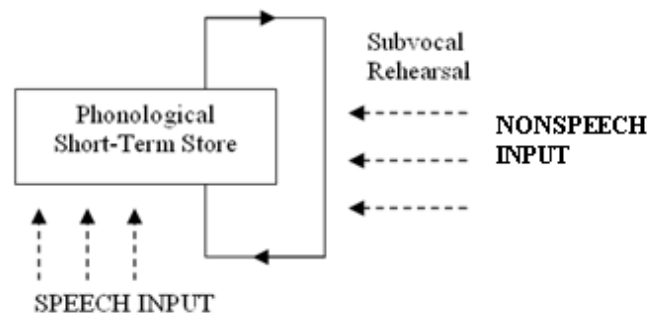
Figure 6. The Phonological Loop (Baddeley, 1986).

1.3.2 The phonological loop. The PL is a slave system specialized in the storage of verbal material. It is a two component process containing a short term phonological store and the sub vocal rehearsal system (Figure 6). As its name implies, the phonological store is responsible for the storage of phonological code (both acoustic and non-acoustic). This code decays over time (Baddeley, 1974) but may be refreshed by the rehearsal system. Critically, auditory information (such as speech) has direct access to the phonological store, whilst non-verbal information (such as words), must be recoded through the rehearsal system before it is stored.

Experimental methodology of the phonological loop. Evidence for the architecture of the phonological loop comes from a number of experimental phenomena associated with phonological STM, such as the phonological similarity effect, articulatory suppression and the word length effect.

In the case of articulatory suppression the recall of printed words is affected by the suppression effect of repeatedly vocalizing words such as "the, the, the." This disrupts the rehearsal process resulting in reduced phonological coding and refreshing, meaning that the phonological code decays without replacement (Gathercole & Baddeley, 2003).

Further evidence for the process of rehearsal was found in what became known as the word length effect (Baddeley, Thomson & Buchanan, 1975) which states that longer words are harder to recall than shorter words (within both verbal and visual stimuli). This alone implied a maximum capacity limit, but when the size of the words being recalled was examined the findings suggested that the observed word length effect was "due to the articulatory duration of the items, and not simply the number of syllables they contained" (Gathercole & Baddeley, 2003, p. 9). This suggests that rehearsal is a real time process and that capacity within this process is not down to the number of spaces available during processing for phonological code but rather to the length of time information is held within the system. This temporal rather than volumetric based capacity has a drastic effect on recall: since words with longer spans take longer to encode, the loop is less able to refresh the short term store, meaning that other items (of any span) decay more rapidly. What is more, since longer words take more time to rehearse, more short words may be maintained compared to longer ones.

Since the word length effect and articulatory suppression both appear to affect the rehearsal process (which as mentioned is limited by temporal-

space) it was hypothesized that the advantage short span words have over

longer span words would be abolished under conditions of suppression, i.e.

when rehearsal is prevented or made more difficult. This was found to be

the case in a further experiment (Baddeley, 1995); wherein both verbal and

visual recalls were affected[13]. "The results were clear." writes Baddeley:

"The standard word length effect was present under control conditions but

was abolished under articulatory suppression" (Baddeley, 1995, p. 81).

This therefore provided good evidence for both the sub vocal rehearsal

process and the existence of the PL.

Another effect that gave rise to the structure of the PL was the phonological

similarity effect[14], which, as with most of the evidence for the PL, was a

well known phenomenon prior to the development of WM. Simply put, this

effect occurs when participants make acoustic errors on visually presented

stimuli in a recall task, phonologically similar letters being recalled less

accurately than dissimilar ones (Conrad, 1964; Conrad & Hull, 1964).

Salame & Baddeley (1982) suggest that this effect results from the

degradation of an item that is phonetically similar to another already being

held in the STS. In such cases, when the number of discriminating features

between the two is small, the effect is particularly catastrophic as the loss of

potentially definitive/unique features will render it identical to the other

maintained item(s).   In these experiments, when the recall stimuli were

---

[13]Although with visual presentation the effect was similar to verbal presentation, the direct
access phonological code has to the STS means that with phonological presentation "It may
be necessary to continue articulatory suppression during recall as well as during list
presentation" (Gathercole & Baddeley, 2003; p. 11)

[14] Conrad (1964) used the term 'Acoustic Similarity'.

presented visually the phonological similarity effect seemed to disappear under articulatory suppression, however it remains within auditory stimuli interference (Gathercole & Baddeley, 2003). This discrimination is accounted for because the effect takes place within the STS, not the articulatory loop. Articulatory suppression blocking the rehearsal process and interfering with the access visual information has to the STS, whilst verbal material, which has direct access, is not affected.

1.3.3 The visuo-spatial sketch pad. The VSSP is a slave system involved in the generation of images and in the retention of information with visual or spatial dimensions. It is responsible for processing verbal material that has been encoded in the form of imagery, integrating visuo-spatial information from multiple sources (such as visual, tactile, and kinaesthetic; as well as episodic and semantic LTM)[15] (Baddeley, 1995; Gathercole & Baddeley, 2003; Baddeley, 2007). Practical applications for such a system include mental calculation, mental mapping (i.e. route learning) and mnemonic strategies, skills that are critical in occupations such as architecture or engineering (Baddeley, 2007).

Experimental methodology of the visuo-spatial sketch pad. The idea that the visuo-spatial memory are distinct processes from the phonological system comes from a number of sources (See Baddeley, 2007; Baddeley,

---

[15] The WM model assumes that the VSSP is a storage system for integrating visual and spatial information. This visuo-spatial code is acquired, not just from vision, but also touch, language and LTM, so the term "domain specific" when used for the VSSP is a little ubiquitous.

1995 for a review), the most influential of which has been neuropsychological evidence citing double-disassociations between patients with impaired visuo-spatial ability and intact verbal ability, and vice versa (DeRenzi & Nichelli, 1975; Hanley, Young & Pearson, 1991).

The WM model proposes that such a visual-spatial system may be further divided into two separate but otherwise complimentary visual and spatial systems, and over the past few decades supporting evidence also has emerged to suggest that visual-memory systems might be separate from spatial-memory along these axis, such research stemming from further double-disassociations in FLHI patients (Darling, Sala, Logie & Cantagallo, 2006) and interference tasks which reported an observed difference between location and appearance memory (Darling, Sergio & Logie; 2009) . Admittedly much confusion exists around how 'visual' and 'spatial' memory may be defined, however very broadly speaking 'spatial' may be defined as the location of an object (where it is), whilst 'visual' memory refers to the detail within an object (what it looks like). (Darling et al., 2006).

The joint visual and spatial WM hypothesis arose from a number of task based experiments (Baddeley, 1995) designed to tap one system but not the other (such as sound cuing in a darkened room, or the recall of colours and shapes), using modality specific interference via the dual task technique. Two of the most frequently cited visual and spatial tasks, which are often used as a psychometric measure of visual and spatial abilities respectively,

are the Wilson, Scott and Power (1987) matrix task (participants must remember the order of shapes) and the Corsi (1971) block tapping task (participants must physically tap out the correct sequence of answers on different blocks). More recently, passive and active aspects of VSWM have been highlighted, which allow the subdivision of storage and attentional components (Hamilton, Coates, & Heffernan. 2003). As a consequence of such evidence, two components of visuo-spatial working memory have been proposed: a visual cache thought to store the appearance of a stimulus and an "inner scribe", believed to store crucial spatial information such as locations (Logie, 2003).

Logie's elaborative model. In contrast to the PL and the CE, Baddeley's accounts of the VSSP are far less developed (Pearson 2001), the VSSP being described as a "younger sibling to the theoretically more mature phonological loop" (Logie & Pearson, 1997, p. 241). As a result, unlike the PL, there is currently no officially accepted (sub)model within Baddeley's WM for the layout of what the components of the VSSP may look like. There are a number of reasons for this, the most critical being that visual spatial working memory is a notoriously hard concept to study: there is no analog for the word length effect, most visual tasks seem to involve at least some elements of spatial memory, and all spatial tasks (developed thus far) are moderated by extraneous factors such as motor skills/muscular control.

Although Baddeley (1995) does not describe the inner architecture of the VSSP Logie and colleagues (1995) provide us with a more descriptive

model of what the components sub-processes may look like. Logie differentiates the VSSP's functions into three distinct processes (all of which may be roughly analogous to the visual spatial equivalent of the PL model created by Baddeley et al.).

i)   Visual cache, a short term store: a passive process not involved in rehearsal.

ii)   The inner scribe, an active rehearsal loop for both visual and spatial information: critically, the scribe is also responsible for the manipulations/transformations of the information within it).

iii)   And, the central executive: a control process for the other two modules.

Although fundamentally in agreement with the overall perspective of the VSSP[16] Baddeley and Logie differ in terms of emphasis.   The key difference between the two perspectives is that Baddeley holds the Sketch Pad to be a single system for the processing of visual and spatial information, whilst Logie suspects that the processes are separate (Mohr & Linden, 2005). To Logie, Baddeley's VSSP is a "mental workspace" (Baddeley, 2007, p. 92), an area where visual-spatial manipulations are made on data already stored in LTM memory. Conversely, Baddeley's stance on the VSSP is based around the short term storage of visual information relying on the direct processing of sensory code from

---

[16] Baddeley notes that the lack of evidence is still "sparse" regarding the VSSP and does not yet justify totally "abandoning Logie's concept" (Baddeley, 2007, p. 92) of the rehearsal process despite his concerns, detailed above.

perceptual input rather than a unitary inner scribe which holds, and then processes, all forms of information in LTM as well.

In the quadripartite model the VSSP does not manipulate information; instead this role is fulfilled by the CE/E. Although Baddeley's tripartite concept did account for transference of information from LTM, it did not qualify how this might take place until the quadripartite model was introduced.

Logie's elaborative model has been called indirect because of its nature of handling data from LTM rather than directly from sensory input (Baddeley, 2007). As such it should be noted that Logie's model fits the original tripartite model better than the quadripartite, as the manipulation of information from LTM in the tripartite model is now performed up by the EB in the quadripartite rather than the CE. This effectively means that Logie's use of the CE as a workspace is now redundant.

1.3.4 The central executive. Described as the most important component of working memory (Baddeley, 2007) the CE fulfils a number of critical supervisory roles: coordinating strategy selection and planning, and controlling the transmission and (in the tripartite model) retrieval of information from LTM to the other subsystems. As previous described, WM is defined as being a 'capacity' limited, system and one of its primary roles of the CE is to govern the flow of attentional through a number of

executive processes[17] (see p. 84). Like the VSSP the CE started life as one of the least studied aspects of working memory[18], however by the mid 1990s the lack of evidence had become something of a concern (Baddeley, 1996). As with the VSSP there was no cognitive model, and little could be done to hide the ambiguity surrounding the critical (directorial) decisions the CE was being asked to make; a point noted by Merlin Donald when he infamously criticised the CE as being little more than an inscrutable homunculus[19] (Donald, 1991).

This initial critique of the model was acknowledged by Baddeley and his colleagues, who adopted the Norman and Shallice (1986) supervisory attention system (SAS) as a possible alternative to the CE. This move effectively abandoned the idea that the CE was in any way involved in storage (i.e. was a mental workspace) and postulated instead that attention was the processing resource famously implied by limited capacity (Baddeley, 1995). Although frequently misquoted as being a unitary construct (Baddeley, 2007, p. 118), the CE is now assumed to have at least four particularly important executive component processes that are believed to globally effect WM including the capacity to focus attention, divide attention, switch attention, and to provide a link to LTM (Baddeley, 1996).

---

[17] Processes should not themselves be seen as being limited capacity; rather these processes define what we understand as a limited resource system; constraining attentional control in a number of different ways.

[18] In Baddeley's 1995 book concerning the tripartite model represented, work on the CE covered 11% (one chapter) of the overall written content. Comparatively an entire book (Gathercole & Baddeley, 2003) has been written on the articulatory loop.

[19] The exact quote is: "The C.E. is a hypothetical entity that sits atop the mountain of working memory and attention like some gigantic Buddha, an inscrutable, immaterial, omnipresent homunculus, at whose busy desk the buck stops every time memory and attention theorists run out of alternatives". (Donald, 1991, p. 327)
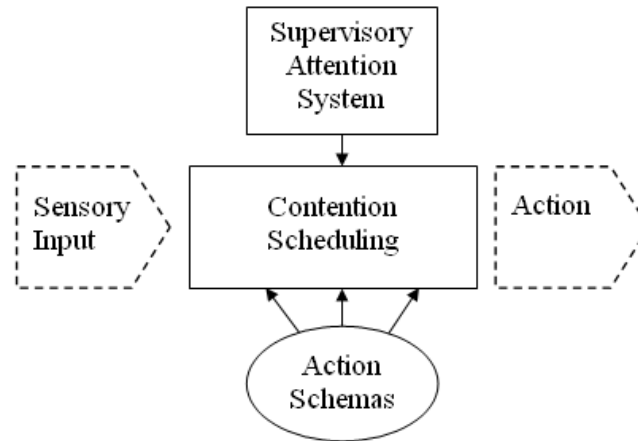
Figure 7. The Mk II Supervisory System Model. (Adapted from Shallice &
Burgess, 1996)

The supervisory attention system (SAS). According to Norman and
Shallice (1986) behavioural output (action) is either automatic (via a
contention scheduling system) or controlled by executive processes (via a
supervisory attention system: Figure 7). Automatic actions are stored as
schemas which are chosen by their relevancy and which are used until the
system either runs out of attentional resources, or another schema overrules
the first. In the latter case, which schema is preferentially chosen is
decided by a process of "contention scheduling", which uses inhibitory and
excitatory processes to decide upon the action relevant to the situation. For
automatic processes (such as driving a car) the relevant action is chosen if
there are appropriate environmental stimuli present (such as being in a car).
In the cases where automatic tasks cannot be carried out, or where new
stimuli are involved, the SAS process becomes involved, increasing or
decreasing excitatory variables until a single schema is selected.

Although switching the CE with the SAS effectively silenced the critics and circumvented the infamous 'Homunculus' problem, new problems were created in the process. One big side effect of adopting the SAS was that the system had no storage component: it worked in parallel to a STS. Since the main criterion for WM is that it should be able to manipulate the material being stored "incorporating information indirectly, either from LTM or by allowing cross modal encoding" (Baddeley, 2007, p. 85), the new model had effectively removed cross modal semantic manipulation within information stored in LTM. Therefore a new component was required, one that restored the tripartite CE's buffering facility. This new component was the EB.

Experimental methodology of the central executive. The process of investigating the CE has been made easier by the recent advances in (and availability of) neuroimaging techniques- which, since the creation of the quadripartite model, have now convincingly isolated the frontal lobes as being involved in the executive processes (for review of this see Henson 2001; Kane & Engle, 2002; Roberts, Robbins & Weiskrantz, 1998; Smith & Jonides, 1997; Stuss & Knight, 2002).

Prior to this, the investigational methodology of the CE centred on neuropsychological techniques investigating inhibition and inappropriate behaviour in frontal lobe damaged patients (frontal lobe head injury or 'FLHI' patients). Since WM operates on the assumption that the selection and processing of information is a critical attribute of the CE, conceivable

central executive deficits would impair processing ability, which would in turn explain well known side effects of such FLHIs, such as poor planning and organisation skills (Hartman, Pickering & Wilson, 1992). This assumption was supported by a series of behavioural experiments in normal participants, each of which looked at the four areas of WM indicated by Baddeley (1996) as being involved in executive control (for full a review of these see Baddeley 2007; Baddeley 1996). As mentioned previously, these were

i)   The capacity to focus attention.

ii)   Divide attention.

iii)   Switch attention.

iv)   To provide a link to LTM.

Three of these executive processes are attentionally based- the fourth, the link to LTM, came about when the SAS model was introduced to the WM framework and the CE was stripped of its storage capacity: therefore the EB makes up the fourth executive function of the model.

Because executive functions share much common ground (specifically inhibition), identifying separable aspects of each is paramount, but understandably difficult. Yet recent evidence has arisen that suggests that this task is far from impossible, with growing evidence appearing over the last ten years to suggest that three functions in particular (separate interpretations of Baddeley's big four) may be isolated (Miyake, Friedman,

Emerson, Witzki, Howerter, & Wager; 2000), shifting, inhibition and inhibition and updating.

The following four sections will detail each of Baddeley's functions in order, briefly describing the methodology and reasoning behind isolating each function.

(i) Attentional focus. Focusing attention is one of the most crucial features of the central executive (Baddeley, 2007). It is now readily accepted that attentional capacity is limited but that different factors such as task familiarity[20] (Schneider & Shiffrin, 1977) or number of tasks being held in mind at any one time (Baddeley, 2007) can reduce attentional demands. Historically two of the most frequently encountered experimental methodologies used to investigate attentional focus have been random verbal generation and the generation of items from semantic categories (RVG). Semantic categorization (such as naming as many words or animals as possible that begin with certain letters, without repetition) is both attentionally demanding (Vallar, & Baddeley 1984) and known to be vulnerable to frontal lobe damage (Milner, 1964). Random verbal generation (generating supposedly random numbers or letters out of sequential order, such as 1, 2 or A, B and in the case of letters, without commonly encountered acronyms) is known to share common

---

[20] The most likely reason for this is because (as modelled in the SAS) "habit-based automatic processes become more efficient" (Baddeley, 2007, p. 126).

processes/resources with decision making whilst showing improvement across the developmental course (Baddeley, 1966; Holding 1989).

Interestingly, RNG has been shown to be comparatively less sensitive to the mediatory effect of rule comprehension although it has been suggested that this may be because RNG is representative of general high level functioning (Towse & Mclachlan, 1999) rather than a specific attentional function.

Methodologies such as RNG or RVG may be highly-useful to WM psychologists as they exemplify the different facets of attentional constraint, tapping the same executive systems, specifically inhibition as other more complicated tasks. Thus, when conducted simultaneously, such tasks as RNG or RVG are understood to affect the concurrent task (Vallar, & Baddeley, 1984).

Separating 'active' executive functioning from 'passive' storage processing is one of the major tasks of the WM executive and it should be noted that a counter argument to the attentional claim is the position that the findings from studies using just verbal generation or categorization tasks do not automatically imply a distinct executive system comparable to the CE. This is the case of the Holding chess study, where Holding (1989) demonstrated that counting backwards in threes interfered with the retention of chess pieces within carefully structured chess problems, but rather than assuming this was interfering with an executive resource, Holding attributed his findings to the limited capacity phonological loop. This assumption has

since been refuted. In a number of experiments centred on similar chess problems, Baddeley (2007) investigated whether or not the CE could be implicated instead in any of these claims. Using dual tasks designed to disrupt the three primary aspects of WM (PL, VSSP and CE namely articulatory suppression, spatial tapping and backwards counting accordingly). Baddeley and his colleagues found that articulatory suppression had little or no effect on the retention of the chess positions; spatial tapping had a moderate disruption, whilst backwards counting had a severe effect on performance, implying that backwards counting was an executive demanding process rather than a purely phonological one. Critically in a follow on study (using chess problems restrained by a time limit) Baddeley also found that "performance was not impaired by a demanding concurrent load" (Baddeley, 2007, p. 125), suggesting that task difficulty was unrelated to attentional demand.

(ii) Division of attention. If one of the features of the CE is to allocate resources to "heavy" cognitive loads it is reasonable to assume that the ability to schedule 2 or more cognitive tasks is either a separate process from other executive functions, or a subtask of the overall ability to focus attention. Baddeley (2007) assumes the latter, stating that it is "potentially dissociable" (Baddeley, 2007, p. 136) but also states that although work is "appears promising" it is by no means "firmly established" (Baddeley, 2007, p. 138). Unsurprisingly therefore, the executive ability to divide attention enjoys the least empirical support out of all four abilities currently

associated with the CE, with much of the work being either pragmatically or intuitively based.

As mentioned previously the WM model assumes that some processes are more automatic than others, but it is highly evident that we are able to conduct more the one cognitive task simultaneously, and that some of these tasks interfere with one another (such as driving, reading, or playing the piano). What experimental evidence there is comes mainly from Alzheimer's disease patients, whom it is argued have trouble performing more than one skill concurrently, but who have a level of performance on the individual (single as opposed to dual) tasks that remains steady throughout the majority of the disease (Baddeley, Bressi, Della Sala, Logie & Spinnler, 1991), implying that Alzheimer's disease is disrupting an executive functioning facility that is able to divide attention. Other more general studies support this association, such as Hartman, Pickering and Wilson (1992) who found that FLHI patients found it more difficult to divide their attention over two tasks[21].

(iii) Task switching. Task switching is perhaps one of the better known control processes thought to be associated with the CE, with research citing its importance dating back to the early quarter of last century (see Baddeley, 2007, for a review).

---

[21] Studies by both Baddeley (2007) and Hartman (1992) found that the act of conversation had a particularly big negative effect on dual tasks: this in turns suggests that social activity has a high executive loading.

As with other processes implied by Baddeley as being core to CE function, Task Switching has been shown to be linked to FLHI patients (Milner, 1965). Historically, psychometric measurements of task switching include tasks such as the Wisconsin Card Sorting Task[22] (WCST; Wells printing and digital services, 2000) or the Trail Making Task[23] (TMT; Reitan, 1958). Other commonly encountered tasks such as the Stroop are also thought to involve similar core processes.

Although difficult to define with any precision (Rogers & Monsell, 1995) a 'task' may be described as a subset of cognitively available representations, which may contain a number of different processes, learned material, and/or perceptual inputs. Switching tasks involves arranging these processes, either consciously or subconsciously to achieve a goal (Monsell, 2005).

In order to switch routinely between parallel tasks using the same input we must a) maintain two or more task-sets active and accessible whilst b) simultaneously being able to disregard (reconfigure) the appropriate set as required (Rogers & Monsell, 1995). Additionally, on top of these, we must concurrently hold the initial instructions as to which set is needed at which point and be able to recall our current position in the problem (i.e. which methodology is required next); earlier studies having already shown that

---

[22] Participants are shown a number of cards and told to match a stimulus card with them. Participants are not told the rule governing how the cards, and at set periods of time, this rule changes. The errors are taken as the score.

[23] The TMT comes in two sections. In the first, participants join sequential numbers on a sheet of paper using a pen/pencil (i.e. 1,2,3 etc.), in the second, they must alternate between letters and numbers (i.e. A1,B2,C3 etc).

task difficulty in task-switching is increased when the stimulus provides a clue as to which is next (i.e. is not identical) (Spector & Biederman, 1976).

Most commonly we measure the cognitive costs of switching between tasks (switch cost) either by comparing a standard number of correct/incorrect responses in a task where switching is not required to a task using identical stimuli where switching is required (error cost), or by determining how long the participant takes to complete either (time cost) (Rogers & Monsell, 1995).

Using a verbal version of the TMT (involving articulated days of the weeks and months rather than letters and numbers) and articulatory suppression (involving repeating days of the week, then months articulated continuously) as dual tasks, Baddeley asked participants to add or subtract (depending on condition) digits from a column depending on a stated or unstated rule. It was found that without signs stating the rule participants were highly impaired on the primary task (judged by higher reaction times), especially under conditions where the verbal TMT was being conducted concurrently.   From this Baddeley hypothesised that the verbal trail making test was using similar attentional resources, increasing what Baddeley called the "switch cost". Attentional demand may not be related to switching itself however, as "switching under certain circumstances, appears to enhance performance" (Baddeley, 2007, p. 133).

Intriguingly however, there is some controversy surrounding whether or not these measures represent a specific cognitive function of the CE or a matter of (cognitively simpler) retrieval of learned associations which may or may not be attentionally demanding and have attentional costs associated with them (Monsell, 2005; Baddeley, 2007).   Baddeley himself somewhat sidesteps the issue, stating that task switching "is not a general function, but a process whose costs or benefits are likely to vary depending on the precise situation and the strategy adopted by the subject to deal with it" (Baddeley, 2007, p. 133), despite this, the implication that task switching is an important cognitive ability is still prominent in the field of psychometrics; its connection with memory taken somewhat for granted.

Miyake's three primary functions.   As has hopefully been indicated above, executive functions are highlighted by psychometric tasks targeting specific mental abilities; but a lot of the time such abilities are highly abstract, using terminology such as "planning" or "strategy" to describe success at particular problems.

It makes good theoretical sense that such general functionalities are of course made up of component processes, some of which may be pinpointed to specific functions. In 2000 Miyake and his colleagues used structural equation modelling (SEM) alongside a number of frequently used tasks purported to assess executive functioning. Modelling of the data revealed three clearly separable low-level functions which are assumed to underlie most executive tasks.

i) Shifting. The ability to switch back and forth between operations: the failure of which has been associated with perseveration and attention-maintenance errors in FLHI patients. Tasks used to measure shifting ability appear to be based on temporal metrics, designed to assess how quickly earlier modes of thinking can be disengaged in favour of new measures, such as the Local Global Task (Navon, 1977) which purports to measures what was termed the "switch-cost" (Miyake et al., 2000) of changing between rules which might be applied to otherwise identical problems.

ii) Inhibition. As described above inhibition is perhaps considered to be the most important aspect of executive functioning, and is consistent with Baddeley's 2000 model. However, according to Miyake, inhibition in this sense may be defined as being solely the inhibition of prepotent responses. Common measurements of inhibition are the Stroop task (Stroop, 1935) or Stop-Go tasks (e.g. Logan, 1994), which require a dominant or learnt response.

iii) *Inhibition and updating*. The ability to hold information in mind and dismiss non-relevant 'old' information for relevant 'new' information: thus updating it. Because of its dynamic nature updating is highly associable with the concept of storage and processing in the WM model requiring information to be currently held in mind whilst also selected for meaning. Since the original WM model no longer represents a CE based capacity, such balancing would have to take place in Baddeley's fourth executive-

function-cum-module the episodic buffer, which deals with binding

meaning to maintained information.

1.3.5 The episodic buffer. The adoption of a priming mechanism

such as the SAS had negated the possible role of the central executive in

setting up, maintaining and retrieving temporary representations in long-

term memory (Baddeley, 1996). As a result the CE was no longer seen as

being capable of being a back-up-store for STM, effectively meaning that a

good portion of the mental workspace aspect of WM had been taken out of

the model. This change necessitated that a new aspect of WM be

introduced, one that was capable of storing information that had been

retrieved from LTM and that could also be chunked into meaningful

episodes (an area which had not been fully explored in the tripartite model).

This new component process was the EB, and formed the core component

of the new quadripartite model. Whether it is a unique system or a core

process of the CE is open to interpretation, however recent studies have

suggested that an executive process is more suitable, as the actions of the

EB largely depend on the dynamic capabilities of modality specific modules

such as the VSSP or AL (Baddeley, Hitch & Allen, 2009).

The EB is named because it integrates cross modal information into

coherent episodes (Baddeley, 2007); it also is a limited capacity store[24]

(Cowan, 2005), capacity being defined by 'chunks': "a package of

---

[24] The term 'resource' within the EB is slightly misleading as the EB's capacity is
determined by the limitations of the PL, VSSP and CE; however in this case capacity refers
to the number of chunks maintained.

information bound by strong associative links within chunks and relatively weak links between chunks" (Baddeley, 2007, p. 148). The main role of the EB is therefore to create these internal links, binding information into chunks; the manipulation of these chunks for meaning being performed by static and dynamic binding processes. Static bindings are learned associations (such as 'fire trucks' are 'red'), whilst dynamic bindings are novel combinations of items (Baddeley uses the example of a 'red banana' in 'blue porridge') that can be combined in a number of different ways.

Experimental methodology of the episodic buffer (binding). Up until the time of writing, experiments by Baddeley and his colleagues have mainly centred on the most critical aspect of the EB: that of binding. In order to investigate this phenomenon, two popular modality specific methodologies have arisen (to reflect the multi-dimensionality of the EB), that of colour-shape testing (VSSP) and that of syntax/prose testing (PL).

Syntax/prose testing. Because of the scope of the English language, verbal testing is more varied in nature than the colour-shape paradigm. It is defined by the use of phonological information (either verbal or written) and the investigation of the semantic information bound to it. It originates from the well known effect of phonological capacity (see earlier section on the phonological loop), i.e. that participants are likely to be able to recall six or seven words on their own but when these are put into meaningful sentences they can remember 16+ (Baddeley, 2000). This suggestion is supported by neuropsychological evidence, such as the case of patient P.V. who had a

severely limited digit span of just 1 but a sentence span of 5 (Vallar & Baddeley, 1984), a phenomenon traditionally explained by Miller's process of 'chunking' whereby different types of information are combined to make it more memorable.

Of course one of the main aspects of chunking is that the more meaningful information (bound semantic information) a target phonological item has attached to it, the more likely it is to be recalled, a suggestion which might be interpreted to suggest that a binding-cost is associated with the association of previously unconnected objects. Indeed, as Jefferies and her colleagues showed (Jefferies, Lambon & Baddeley, 2004) as progressive learning increased so did attention-demands. But, like other WM theorists, Jefferies did not automatically prescribe this to the active binding process. Instead it has been suggested that this effect is down to the construction of individual strategies (what Jefferies and her colleagues call multi-word chunking), meaning that the task is more open to executive interference the more external demands (outside of the task) are placed on it. Additionally, in the same vein, Jefferies also showed that the difference in "binding cost" of multiple as opposed to single words is minimal as the timed-recall of unrelated words was the same when compared to meaningful sentences. Instead what seemed to alter outcome the most was the span of the task itself (i.e. task requirement), a heavier load equating to less items recalled and longer processing times. Instead of implicating the EB as being attentionally demanding, this therefore places all demands associated with

phonological encoding firmly at the door of the PL and its individual

capacity (Jefferies, Lambon & Baddeley, 2004).

This conclusion was reinforced recently by Baddeley's laboratory

(Baddeley et al, 2009) who also indicated that although the phonological

binding-capability of the EB may indeed be effortful (i.e. capacity based)

this is only so if the associative semantic-bindings require further attentional

resources beyond the span of the PL. Here Baddeley specifically suggests

that newly developed task-specific arbitrary combinations (i.e. not backed

up by learned facts from LTM require more attentional resources than

relatively automatic semantically based chunks (i.e. sentences that have

meaning beyond the task), but it remains to be seen whether this effect may

be comprehensively demonstrated in the Binding process, or whether, as

Baddeley himself suggests, the binding-demand comes from the number of

chunks capable of being simultaneously maintained.

The EB's role in chunking/binding is further highlighted by previous

phonologically based evidence that individual differences in WM affect

reading comprehension and development (see Gathercole & Baddeley,

1993, for a summary), as well as research from authors such as Poulton

(1958) who showed that increased reading speed leads to decreased

comprehension of the text, suggesting within the confines of the EB that the

reading rate was either interfering with the amount of information able to be

held/updated, or the chunking of semantic meaning to the words (Baddeley,

2007). Baddeley and his colleagues studied these effects in a series of

experiments based on the influences of speed on recall and accuracy within reasoning tasks (Baddeley & Hitch 1974) and sentence comprehension (Baddeley, 2007). Despite intuitive claims to the contrary, executive influence was found to be small, with minimal attentional demands being necessary for binding, executive interference being found to increase verification latency but not the number of errors[25] (Baddeley, 2007). Attentional influence, it seems, is reserved for heavy cognitive loads, with a minimum amount needed for the process of binding, which appears to be fairly automatic.

Colour/shape testing. Similar findings have been found in the colour-shape paradigm. Within visual-centric investigations into the EB participants are normally shown either patches of colour, coloured shapes, or shapes themselves, and after a short delay are given a question-probe and asked whether or not it had just been presented.

A series of these experiments (Allen, Baddeley & Hitch, 2006; Baddeley, 2007) showed that participants were as equally good at recalling coloured shapes as they were recalling individual shapes or colours, and that disruption by secondary tasks (either random-number-generation or a concurrent digit span task) had equal effects on colour-shape combinations as opposed to individual features, implying that (under normal circumstances) the EB either has enough attentional resources to manage the

---

[25] Baddeley initially suggesting that this may be because the PL capacity is more then enough to handle the phonological information used in the individual tasks methodologies, implying that a considerable load is needed to see an effect.

normal binding process, or that binding does not discriminate between the number of dimensions being processed. It is worth noting that even this latter suggestion does not necessarily mean that visual binding is entirely void of attentional demands, just that passive binding is more likely to be a peripheral, pre attentive process. Following on from this suggestion the role of visual attention in the binding processes has instead been suggested by Allen and his colleagues (Allen et al., 2006) as being sequentially based: in other words it is centred upon maintaining bonds in the face of conflicting stimuli (relating the sequential activation of propositions to the ability to inhibit the activation of non-relevant associations whilst allowing/facilitating the synchronous firing of those that are relevant (Morrison 2010)). It has been suggested by Allen et al. that this ability to visually maintain information affects multiple bindings, implying fragile rather than resource demanding connections.

1.3.6 Distinguishing WM and intelligence tasks.   Intelligence and WM are both core constructs that are central to our understanding of human cognition. Broadly speaking, an intelligence quotient (IQ) is measured by variety of tasks which abstract general "high level" functions and processes (Kyllonen & Christal, 1990, p. 426) in areas such as mathematical, verbal and/or written skills (Spearman, 1923, 1927) from which the normal developmental course has been standardized. WM is measured by tasks which rely upon span (capacity) testing.

Despite this, the dividing line between IQ and WMC (which places processing above storage) has not been so clear. While WM measures have been established as measures of general ability, it has also been argued that the two terms are frequently interchangeable, if not indivisible (Colom, Rebollo, Palacios, Juan-Espinosa & Kyllonen, 2004; Colom, Jung & Haier; 2007), with underlying factors such as speed of processing (see p. 83) associated with both (Kyllonen & Christal, 1990).

It has been argued (experimentally) that components of WM may be uniquely discriminated against IQ (Ackerman, Beier & Boyle, 2005; Colom et al., 2007; Oberauer, Schulze, Wilhelm & Süß, 2005) or even that that WM is actually a better predictor of academic ability then IQ (Alloway & Alloway, 2008[26]).

One argument appears to be that intelligence is more fluid and generally applicable than WM tasks, the number of operations possible being restricted by the task modality i.e. the verbal and/or visual input (Heinz-Martin, Oberauer, Werner, Wittmann, & Schulze, 2002). For our present purposes this position has been accepted; it being understood that the concept of memory as being separable from other systems is paramount when investigating influences on human reasoning. The measurement of WMC only being meaningful if its mediatory effects can be separated from

---

[26] Although this point is purely determined by ones definition of IQ.

systems associated with generic IQ processes. For instance, As Oberauer
(2005) states:

> "It would be a surprise and an embarrassment if one found that
> measures of WMC and measures of g were perfectly correlated. It
> would imply that measures of WMC do not come closer to measuring
> a theoretically well-defined parameter of the cognitive system than
> g[27] does." (Oberauer et al., 2005, p. 64).

This thesis will therefore consider WM as being separate from many non-
capacity based tasks. This means that every effort will be made to separate
effects of capacity in verbal, visual and global domains of WM, as defined
by Baddeley's quadripartite model (2007).   It is however acknowledged
that at the information-processing level no absolute discrimination is
possible, or particularly desirable.   For instance, whilst processes
underlying reasoning might be accounted for by the WM model, they may
also be explained for by additional individual differences which are generic
in nature and in some circumstances might be applicable to other systems,
such as processing speed (see p. 89).

---

[27] In this sense 'G' is implied to be a nondescript amalgamation of systems, processes and
abilities.

1.3.7 Working memory and reasoning. As had already been stated WMC is already well understood to be connected to the processes underlying non-analogically based reasoning tasks (Braine & O'Brien, 1998; Johnson-Laird & Byrne, 1991; Logie & Gilhooly, 1998; Rips, 1994), with individual differences in memory possibly accounting for erroneous solutions. This individual-difference perspective does not necessarily stem just from attentional or storage constraints. Another factor already highly associated with influencing reasoning ability is that of processing speed (see below).

Processing speed, reasoning and WM. Speed of processing (SOP) is an arbitrary term used to define the rate at which our cognitive functions operate. SOP is thought to constrain cognition by limiting the effectiveness of higher order faculties associated with both WM and IQ; arbitrating processes such as execution speed, the synchronisation of secondary tasks and the scheduling of operations. SOP is representative of the overall efficiency of neural networks and mental systems (Kail & Salthouse, 1994, Salthouse, 1996).

In terms of WM, increased fluency is understood to free resources for processing; SOP facilitating executive faculties involved in the parallel and sequential division and selection of attention (Fry & Hale, 1996, 2000) as well as determining the number of processes capable of being executed while information is held in a STS (Case, Kurland & Goldberg, 1982; Kyllonen, & Christal, 1990).

When reasoning a SOP constraint is believed to occur when the temporal requirements of parallel and sequential processing are exceeded by the demands of the task or problem, in which case task performance is expected to drop (Fry & Hale, 1996). Yet despite this predicted effect, it appears unlikely that SOP alone can completely determine reasoning outcome in isolation. Extensive modelling on the subject suggests that processing speed is a product of processes within WM and fluid intelligence, rather than a direct contributor to comprehension or deduction (Engle, Cantor & Carullo, 1992; Nettelbeck & Burns, 2010). This hypothesis is supported by a growing number of theorists (Colom et al., 2006; Colom et al, 2007) as well as those who point out that WMC measures contain specific processes which processing speed cannot account for such as inhibition of irrelevant schemata (Heinz-Martin et al, 2002).

Regarding plausible roles for SOP in AR, two mechanisms of cognitive speed proposed by Salthouse (1991) are directly applicable to inferential reasoning and the complexity-constraint theory (Halford, 1992; 1993; 1998). The time constraint limitation, where task complexity increases the time needed to form associations and the simultaneous mechanism where slower processing effects the abstraction of meaningful information by limiting what can be processed in parallel or retrieved from LTM. In either case Halford (1998) has suggested SOP as a possible age-related indicator of WMC's mediatory role in AR, while connectionist models of AR such as LISA (Hummel & Holyoak, 1997, 2003, see French 2002 for a review of

this form of programme) intrinsically demand that speed of processing may be thought of as processing efficiency, brought about by strong or weak neural connections between arguments.   These factors will be discussed in detail in the experimental chapters.

Mental workspace theory.   Broadly speaking, most theories of WM and reasoning fall into the mental-model category.   Developed from the original early work by Craik and Pierce (for a review see Johnson-Laird, 1983; 2000; 2004) this approach to reasoning postulates that problem solving is facilitated through the internal representation of (and experimentation with) the external world using "the meanings of assertions, together with general knowledge, in order to construct mental models of the possibilities compatible with the premises" (Johnson-Laird, 2001, p. 1).

From a WM perspective mental-models may be perceived as a cognitive workspace (Alloway, 2006; Baddeley 2007) where abstract rules relevant for immediate processing are stored in LTM and selected for relevance by the CE (Alloway, 2006; Baddeley 2007) to be combined with the results of recent processing (Gilhooly, 1998).   Such a workspace is analogous to the different facets of WMC, with individual differences in passive storage and active (executive) attentional constraints placing limits on how information is processed.

Theorists have hypothesized that holding representations in store in such a manner creates processing traffic-jams when the cognitive load is 'heavy'

(i.e. when more than one representation is held or when processing aspects requires a deeper manipulation of the problem), WMC limitations resulting in what Johnson-Laird calls a "bottleneck in the inferential machinery" (Johnson-Laird, 1983, p. 115). Bottlenecks overburden the reasoning mechanisms, possibly leading to erroneous thinking (Heinz-Martin et al, 2002; Kyllonen & Christal, 1990). Such circumstances often occurring if the processing demands exceed the resources available; or circumstances are not ideal (i.e. the reasoner is required to continuously visualize this problem without opportunity of refreshing it).

The episodic buffer as a workspace. Although theorists initially postulated that the CE might be entirely responsible for the maintenance and selection of information this is no longer the case. The adoption of the multiple resource model in 2000 meant that whilst the CE is still thought to be actively in control of selective information through attentional constraints, the EB and its unique multi domain store is now more likely to be responsible for the role of a workspace in its capacity as a mediator to LTM (responsible for accessing memories) and binder of cross-modal information (Baddeley, 2007).

Within the quadripartite model the slave systems act in tandem with the EB as short term stores for literal surface based code, holding either visual or phonological information as required by the EB, whilst the EB chunks this code into meaningful segments before passing it into LTM. This process is highly relevant to reasoning considering that Johnson-Laird and Byrne

(1993, p. 181) proposed in the 1960s that reasoning is not a form of

"syntactic process but a matter of understanding meanings" whilst as

mentioned in the previous section on the EB, Baddeley sees syntactic

parsing as being automatic in nature, whilst "The maintenance/binding of

several different propositions is effortful" (Baddeley et al., 2009, p. 640).

Experimental methodology of working memory based reasoning tasks.

There is strong evidence for a link between WM and performance in a

variety of reasoning tasks, (Baddeley, 1968; Barrouillet, 1996; Bull &

Scerif, 2001; Carpenter, Just, & Shell, 1990; DeStefano & LeFevre; Klauer,

Stegmaier & Meiser, 1997; Gilhooly, 1998; Johnson-Laird & Byrne, 1991;

Kyllonen & Christal, 1990).   Within this research the dual task

methodology has been used to show that burdening WMC gives rise to

performance errors in reasoning tasks (Baddeley, 1968; Gilhooly, Logie,

Wetherick, & Wynn, 1993; Klauer et al., 1997; Toms, Morris & Ward,

1993). Crucially, this effect does not appear to be a result of overloading a

purely storage based system. Modality specific studies using forms of

grammatical reasoning have typically shown only limited disruption by

concurrent articulatory suppression (Baddeley, 1995; Gilhooly et al., 1993;

Klauer et al., 1997; Toms et al., 1993), whilst (dual) visuo-spatial tasks have

been shown to interfere similarly with visuo-spatial reasoning (Klauer et al.,

1997). In most cases it appears clear that individual performance in

reasoning tasks suffers the most when the dual tasks overload the attentional

(executive) limits of the central executive (Baddeley 2007; Gilhooly et al.,

1993; Klauer et al., 1997; Toms et al., 1993).

Note that it does not necessarily follow that those individuals with high WMC are slower processors than their counterparts. In analysis of reasoning, Sternberg (1977) suggested that slower processors were actually more successful, instead implicating high WMC with the ability to hold more relations active for longer without the need for refreshing them, which may itself lead to more errors.

Performance factors. When considering the possible constraint effects in reasoning there is a danger of assuming that only processes directly involved with the solution are mediated by WMC. As already stated, in any problem solving task WMC may be implicated in processes beyond those primarily associated with the task, such as rule comprehension, meta-learning, interpreting experimenter demand, goal maintenance, inhibiting non-relevant stimuli etc. As such, although individual ability measures may appear to be associated with success/failure, there is a constant danger that a proportion of the results may be due to the effect of secondary performance factors. The best way to account for these during experimentation is to include measures which are less focused on capacity limits than their WM counterparts, i.e. IQ measures.

1.3.8 Analogical reasoning and working memory. To this date the majority of research focusing on WM's role in AR has been dominated by what is known as the complexity hypothesis (Halford, 1992, 1993, 1998, Halford, Maybery, O'Hare & Grant, 1994) developed from Gentner's work on structural mapping (Gentner, 1983, 1988; Gentner et al., 1993; Ratterman & Gentner, 1999).

Complexity-constraint theory. The complexity-constraint perspective has been described as a "construct invoked to explain the systematic decline of performance with increasing task complexity" (Oberauer; 2005, p. 368).

'Complexity' in this sense being defined by Gentner's (1983) taxonomy of relations (see p. 38) whose concept of similarity gave rise to the perspective that the more relationships (predicates and arguments) a base object shares with a candidate target, the greater the complexity and the greater quantity of processing power required represent the problem and calculate an appropriate response. Zelazo and Frye (1998) describing complexity as the more hierarchical rules that must be considered to accomplish a task

The 'constraint' complexity places on cognition has been interpreted differently, but in most cases may be summarised as the number of relationships that must be processed in (Andrews & Halford, 2002; Halford et al., 2002; Halford, 1992; 1993; 1998; Richland et al., 2006). The central idea is that there is a limit to the complexity of possible mappings due to WM limitations. WMC may therefore dynamically constrain analogical

89

thinking by what may be represented and processed. Halford (1998) argues

that as WMC develops across maturation, so does a child's ability to

represent relational arguments, with 50% of children being able to

manipulate ternary relationships by age five, rising to 100% by years 10-11.

It is argued that children who lack the processing capacity to manipulate an

analogical problem are less likely to provide an adequate response due to an

overloaded cognitive system (Andrews & Halford, 2002; Halford et al.,

1994; Halford, Andrews, Dalton, Boag, & Zielinski, 2002), resulting in a

possible increase in the selection of lower order (simpler) relations (Waltz et

al, 2000) which are more likely to be understood.

Over the last two decades a complexity-constraint hypothesis has been

successfully applied to computational models such as the STAR[28] (Halford

et al., 1994; Wilson et al., 2001) and the LISA[29] (Hummel & Holyoak,

1997, 2003) programmes which rely on humanistic limitations to what may

be maintained in conscious thought. However, within the information-

processing field of psychology the approach has been dominated by the

work of Morrison and his colleagues who have used two forms of analogical

problem, the scene-based analogy and the classical analogy to illustrate the

effects of complexity.

---

[28] Both the STAR and STAR-2 (Structural Tensor Analogical Reasoning) models assume
that each argument has multiple dimensions. Previously known statements regarding these
dimensions are held in a tensor array equivalent to LTM, the STAR models focusing on
only a limited number of these dimensions at a time, an act which the authors equate to
WM limitations.
[29] The LISA (Learning and Inference with Schemas and Analogies) model uses a system of
synchronous firing within neural networks to bind observable features to concepts in order
to identity relationships. The successful computation of any analogy relies upon keeping
the relevant relationships separate and out of synchrony. In this context WMC is
understood to be the limited capacity system that keeps all bindings active and independent.

In 2000 Waltz and colleagues examined the complexity constraint to the scene based analogy (cross mapping) paradigm. Like the Markman and Gentner (1993) study, Waltz and his colleagues observed that more relational answers were given in 'multiple' compared to single 'one shot' trials (see p.25), suggesting that people naturally attend to relational forms of response if cued to more than one base object. However, when a dual-task (phonological suppression repeating the word 'the' and executive interference from random-number-generation) was introduced alongside the analogical scene, this significantly reduced the tendency to identify relational similarities.

In accordance with complexity-constraint theory this implied that WMC played a role in the active mapping of relationships, the reduction of available WMC making it more difficult to compute relational mappings and increasing the proportion of less-complex attribute (featural) mappings which appeared to be either a default position in AR or a fall-back strategy, "in situations in which the mapping is ambiguous" (Waltz et al., 2000, p. 1206); featural mappings possibly being a form of response within a child's processing ability.

Although Waltz et al.'s (2000) experiment failed to discriminate between executive and phonological loading, Morrison and his colleagues (Morrison, Holyoak & Truong, 2001) suggested that the scene based analogy

methodology, might not have been sensitive enough to pick-up individual differences within working memory.

In a series of experiments using Sternberg's (1977) visual 'People Piece' analogies (a variation of the classical analogy format wherein the a:b::c:d relationship is represented by characters with one of 4 binary arguments (MAN/WOMAN, BLACK/WHITE, TALL/SHORT, FAT/THIN) as well as verbal classical analogy variants (as described by Sternberg & Nigro, 1980), Morrison et al. (2001) used dual tasks (articulatory suppression, spatial tapping, and random number generation) to tax the phonological and executive components of WM.

In both analogy formats the dual tasks significantly reduced accuracy (although in verbal classical analogies spatial-tapping had less of an impact compared to phonological and executive interference). Yet, the proportion of correct responses was significantly lower in the verbal compared to visual paradigm, leading to the conclusion that task modality existed within AR, and that verbal analogies were more vulnerable to phonological interference.

One of the additional observations Morrison highlighted was that accuracy in the random-number-generation condition was the lowest for each condition in each analogy format, this form of executive interference also eliciting the longest average response times. This lead to Morrison's suggestion that, as in other non-analogically reasoning based tasks, that

92

executive WMC mediated AR; a prediction in-line with Halford's (1992, 1993, 1998) predictions that processing capacity was limiting analogical performance.

In a follow up study Viskontas, Morrison, Holyoak, Hummel and Knowlton (2004) partly replicated the 2001 experiment. They introduced a new variation of the people-piece analogies which could be manipulated in order to increase the complexity of the problem, complexity being defined the number of irrelevant traits (between 1 and 4) favouring the incorrect response which were required to be attended to at once, e.g. in the simplest condition a participant had to make sure that the C:D terms possessed all 4 of the same arguments, specifically height, weight, gender and colour (in other words the answer was far more obvious).   In the most complex condition only one argument would be the same

A second change was the introduction of older participants, it being hypothesized (Hummel & Holyoak, 1997, 2004; Zelazo & Müller, 2002) that a developmental curve exists in the growth of available attentional resources (WMC and general cognitive functioning) that develops across childhood, but which declines in older age.

 As predicted relational responding was reduced in complex conditions for older participants. Younger adults (mean age 20) performed better, and at a faster rate, than older adults (mean age 75) in complex tasks, with the older adults being prone to distraction by the presence of irrelevant and

misleading information. The conclusion Viskontas came to was that a decline in analogical performance in older participants was due to attention and inhibitory deficits brought about through old age, implicating WMC in the need to maintain and manipulate the complex information.

In 2007, Cho, Holyoak and Cannon used a further variation of the PPA paradigm used by Viskontas et al. (2004) that introduced a gap between presentation and response to allow information to be properly encoded into memory and to add a further burden on the storage components of WM. They found that participants were slower at responding when they were, i) resolving conflicts caused by multiple possibilities and attribute based distraction, and ii) integrating multiple relations: effects which increased when all conditions were combined (conflict was presented in complex conditions with a delay).   This suggests a shared executive resource mediating complexity which the authors suggested was WMC.

Concurrent to Morrison et al.'s work on the people piece format of analogies, the concept that WMC arbitrates AR in scene based problems has been supported by Richland (Richland et al, 2004; 2006, Richland, Chan, Morrison & Au., 2010) using the Richland Picture Analogies (RPA). The purpose of these studies being to both challenge Goswami's (1992) proposal that domain knowledge alone is the sole mechanism for relational comprehension, and to offer a new hypothesis for a role WMC may play in constraining analogical performance.

94

The RPA used a 2x2 repeated design (see Figure 8) consisting of two conditions of complexity and two conditions of distraction. Crucially, Richland used Halford's (1998) concept of binary and ternary relations to determine complexity (*complex* ternary relations consisting of an object within a 3 object relationship, such as "chasing and being chased", binary *simple* relations consisting of a 2 object relationship such as "chasing" or "being chased", see page 38). They selected year groups above and below Ratterman and Gentner's (1998) and Halford's (1998) 5-year-old developmental floor (which Halford, 1998 cites as the age in which 50% of children are able to represent ternary relationships) in order to illustrate how changes in WMC influence relational responding under conditions of complexity.

For distraction, 'transformed' objects were chosen that were attributionally similar to the base query object (i.e. the looked the same but might be doing something new, like walking instead of sitting). The idea being that the growth of inhibitory skill (Diamond, Kirkham & Amso, 2002) could represent at least some of the maturationally developing WMC constraint hypothesized by Halford (1992, 1993, 1998).

x1 Relational Condition

x2 Relational Condition

(Above: Target analogy from a x1 Relational scene)

(Above: Target analogy from a x2 Relational scene)

(Condition R1 D) Base analogy from a x1 Relational scene, with distractor)

(Condition R2 D) Base analogy from a x2 Relational scene, with distractor)

(Condition R1 ND) Base analogy from a x1 Relational scene, No distractor)

(Condition R2 ND) Base analogy from a x2 Relational scene, No distractor)

Figure 8. The Richland Picture Analogies (Richland et al, 2004, 2006). Each Question has a 2x2 design with conditions of distraction and complexity. The RPA consists of 20 questions repeated in 4 conditions to make 80 in total. Participants are shown a page containing the target-scene on the top and a base-scene on the bottom. They are required to do one shot mapping by equating the target-object (shown with an arrow in the target scene) to an object in the base scene. Relational objects interact with one another in the same number of ways as the target object (such as the boy above). 'Featural error' objects resemble the target scene in appearance but do not interact within the scene (such as the cat above). Other forms of response are possible, including 'relational errors' (right similarity, wrong object: such as the girl in the pictures above); or 'other-errors' (objects not associated with the problem, such as the house or sandpit in the picture above).

96

When designing the RPA two important controls were introduced by Richland. Firstly, in order to prevent unintended loading of WM by extra objects not involved in the binary-ternary relations, the object per scene ratio in the RPA was set to five (complex conditions replacing a non-relevant background object such as a bush). Secondly, to account for domain knowledge all similarity-constraints used were designed to be recognizable by even the youngest age ranges. Analogies such as 'chasing' being understood by everyone at age 5.

In her experiments, Richland and colleagues (2004, 2006, 2010) consistently demonstrated reduced analogical performance in conditions of increased complexity and distraction, promoting (due to the controls set in place) her overall conclusion that domain knowledge alone could not be the sole mechanism for relational comprehension.

In her 2004 and 2006 studies, Richland et al. identified significant two way interactions for the number of relational responses elicited between age and distracter condition, distraction and complexity, as well as a significant three way interaction between age, distraction and complexity. The three way interaction showed that the effects of complexity and distraction differed within individual year groups. Children in Richland's youngest age group (age 3-4) showed a main effect of both distraction and complexity and a two way interaction between complexity and distraction. Children in the middle year group (ages 6-7) showed main effects of complexity and distraction but no significant interaction between them. Children in the

oldest year group (13-14 year olds) revealed main effects of complexity but not distraction, and no significant interactions between them. Whilst these results confirmed Richland's original hypothesis that younger participants performed less well at identifying relational responses than their older counterparts in the distraction conditions, they also suggested that the psychological construct of WMC (which was suggested by Richland as mediating the ability to illicit relational responses under conditions of complexity/distraction conditions and facilitate inhibitory skill) was a constraining factor in the youngest age range.

This supported Halford's developmental assumptions that maturation increased a child's ability to process complex relations, but also supported (through the complexity-distraction interaction) the Waltz et al. (2000) hypothesis that when the demand on WMC was high, reasoners were more likely to choose a featural response, if available. Richland proposes that individual differences in WMC underpin her findings and argues, through subsequent changes in the key-wording of the instructions[30] (Richland et al., 2006), that it was not task-understanding that was developing but the ability to process complex relations. That is, complex tasks represent a higher burden to the available resources.

In a further study using foreign participants, Richland underlined her position when Chinese pupils were found to perform at a higher rate under

---

[30] Richland replaced complex words such as "Patterns" with "The same as". See appendix A for the wording of the RPA.

conditions of complexity when compared to US pupils of the same age, a factor Richland suggests may be due to increased visual-semantic processing ability of Chinese children used to binding meanings to visually presented complex symbols (Richland et al., 2010).

Inhibition and AR. Like 'analogy' 'inhibition' is a troublesome term that means many different things to different people. Without wishing to be drawn too deeply on the subject what is being discussed here is not neural inhibition (that is to say the inhibition of action potentials in the brain), but what has been frequently termed 'cognitive', 'behavioural' or 'response' inhibition. That is to say: the ability to act by choice rather than impulse, resist non-relevant behaviour, and to 'quickly and flexibly adapt behaviour' (Davidson, Amso, Anderson & Diamond, 2006).

Such abilities represent our concept of what Davison et al. (2006) terms 'cognitive control' and are key components of the executive functions of working memory which frequently rely on the ability to switch attention resources between tasks or maintain a certain mode of thinking against interfering stimuli (Baddeley, 2007; Miyake et al., 2000).

Whilst for a number of years connectionist models such as LISA (Hummel & Holyoak, 1997, 2004) have proposed a form of neuro-cognitive inhibition as selectively pruning non-relevant semantic units, recently a number of theorists have started to consider inhibitory-control processes, as executive functions of WM (see Miyake et al., 2000, p. 73) may potentially be able to

99

explain individual differences in cognitive performance observed in AR (Krawczyk, Morrison, Viskontas, Holyoak, Chow, Miller & Knowlton, 2008; Morrison et al., 2006, 2010; Richland et al., 2004; 2006; 2010; Thibaut, French & Vezneva, 2010),

Such an account makes good theoretical sense. Like WM, inhibitory skill is known (Gathercole, Pickering, Ambridge & Wearing, 2004) to have "extremely slow, protracted developmental progression" (Diamond et al, 2002) while low WMC individuals have also been shown to be less effective at blocking intrusive thought than high-WMC individuals (Brewin & Beaton 2002). As Richland et al. (2004, 2006) indirectly asserts, it is likely that inhibitory control may represent part of Halford's (1992, 1993, 1998) concept of WMC.

The most prominent explanation for inhibitory control in AR is that competition may exist between relational and featural attributes (Morrison et al., 2010; Richland et al., 2004, 2006; Thibaut, et al. 2010). Just why this may be the case is currently being debated, although distraction errors are hypothesized to occur when irrelevant content is not removed from WM and when the information held in mind is contradictory to the present situation (De Neys & Van Gelde, 2008). For example, as might be found in scene-based analogies where an individual is presented with a problem where a relationship previously thought to be true (i.e. "the same" means an object looks similar) is no longer relevant given the demands of the task. Yet it remains to be seen whether featural objects present an intrinsic attraction

regardless of what the base query object looks like (which is opposite to what is predicted by Gentner's principle of Systematicity, see below), and/or whether inhibition is demanded because the query object held in mind is similar to one form of response.

Of course conflict between competing objects is not the only explanation for inhibitory skill mediating performance. Proactive inhibition in AR tasks may also play a role (Lustig, May & Hasher, 2001; Rowe, Hasher, & Turcotte, 2010; 2008), affecting paradigms with similar multiple trials and (in the case of the RPA) similar scenes and/or objects which are repeated across condition. The difficulty occurring either because of "…difficulty discriminating current items from those presented in earlier trials or difficulty maintaining suppression of prior trial information in the face of highly similar items." (Rowe et al., 2010; p. 804) or because external stimuli foreign to the test need to be inhibited. For example, background noise or sudden/unusual distractions causing wandering chain of thought (Chiappe, Siegel & Hasher, 2002; Zacks & Hasher, 1994)

However, as Oberauer (2005) points out, selecting a wrong relational object may also occur from a simple error in representing which content is deemed to be relevant to the task at hand (the ability to disseminate task requirements being one of Goswami's (1992) performance factors.

According to a strict interpretation of the relational shift and systematicity (Gentner, 1988; Gentner & Toupin, 1986; Ratterman & Gentner, 1998), we

should also expect featural objects to be *less* appealing than relational; at least in circumstances where we understand the required similarity-constraint and have the ability to process it, as according to Gentner we allegedly have an innate preference to prefer this form of relationship. However, it may also be true that, as Richland suggests, it is the development of inhibitory skill that represents this shift. At the time of writing this issue currently remains unresolved.

Relational primacy and WM. Relational primacy theory originated from structuralist claims that young children cannot reason analogically, and not from the assertion that relational processing does not take place beyond the recognition of similarity-constraints. As a result Goswami does not go into detail regarding her predictions for the theory, but believes historical evidence from the information processing perspective such as Sternberg's (1977) componential model support her claims.

One might expect any theory which is centred upon the recognition of the similarity-constraint to be dominated by a child's ability to interpret the demands of task (in which case you would expect fluid intelligence and executive decision making processes to play a role as one of Goswami's (1992) 'facilitation gradients' in the construction of schema for the task. However, it is important to note that, despite the term 'relational primacy' the appropriate analogy must also be recognized in the base as well as the target argument. If the appropriate response is not clear (what Goswami briefly describes as 'ambiguous' circumstances) then a role similar to that of

Gentner's mapping must take place in order to determine relevance, involving a further active processing component. Crucial to relational primacy is not the concept that mapping does not take place, just that it is not always necessary.

As mentioned earlier, the chief distinction between relational primacy and the complexity-constraint effect is Goswami's belief that once the similarity-constraint is held in LTM, its size is irrelevant, and that relational-complexity should not represent an extra cognitive load on the analogical process, either in situations of ambiguity or relational recognition.   This may be interpreted as meaning that relational mapping can load WM, but that processing a ternary relationship (as described by Halford, 1992, 1993, 1998) should not be more difficult to process than a binary one. Such an approach is grounded by recent research into binding which predicts that relevance testing is automatic and relatively unweighted (Baddeley, 2007) and that no greater level of attentional capacity is required to bind propositions than encode them separately. That is unless other conceptualizations are being bound at the same time (i.e. more than one similarity-constraint) or the capacities of the PL or VSSP are exceeded during recognition (Allen et al., 2006).

1.4 Research aims and rationale

As illustrated in the preceding section, the central argument for WM's involvement in AR is that capacity limitations[31] within the WM system mediate AR by placing constraints on the number of possible arguments that may be processed in parallel (Halford, 1993) and/or the number of hierarchical rules that must be maintained in order to solve the problem (Zelazo et al, 2004). This is termed the complexity-constraint theory.

The primary opposition to such an approach is considered to be Goswami's (1993) relational primacy theory, which postulates that the cognitive recognition, representation and manipulation of relationships does not place a recognizable burden on cognitive processes. Instead, in relational primacy theory 'performance factors' are thought to explain analogical performance, such as interpreting task demand.

Although (given the volume of research on WM and Reasoning) it is logical to assume that increased WMC aides the resolution of conflicts arising from performance factors (e.g. inhibition), the key difference between the two theories is that in relational primacy WMC is not increasingly involved in the representation of more complex relations. The loading on WMC is

---

[31] Although the term WM may be interpreted differently, 'capacity' in this respect refers to two resource-based mechanisms within the quadripartite WM model (Baddeley, 1996, 2000, 2007): the temporary maintenance of domain-specific information within the VSSP and PL (STSs) and the control of domain-general attentional resources which are defined by a number of core executive functions of the CE (p.57).

considered constant and minimal regardless of the number of rule/arguments the relationship may contain.

Aim. It is the aim of this thesis to better qualify the potential role(s) of WMC in the successful resolution of scene based analogical problems. It will achieve this by assessing the validity of relational primacy theory (Goswami, 1993) and the most commonly encountered interpretation of complexity-constraint theory proposed by Halford (Halford, 1992, 1993, 1998).   The central proposal of the thesis being that WMC mediates AR in scene based analogies by allowing the problem-solver to deal with increasingly complex analogies.

 The aim will be achieved by testing the effects of not only complexity but also distraction on AR - plus making direct comparisons to other/more generic cognitive processes traditionally associated with fluid intelligence.

It is understood that a complexity-constraint approach predicts that WMC will mediate the degree of complexity that may be processed accurately in an analogy; whilst relational primacy theory predicts a greater involvement of more general faculties such as fluid intelligence and/or WMC irrespective of the degree of complexity (faculties which Goswami terms the "facilitation gradient")

Direction of research. In order to achieve the above aim, a series of experiments will be reported comparing individual differences in various

aspects of cognition (i.e. those related to IQ/WM) to AR. Analogical ability is defined here as the ability to elicit a desirable form of relational response from a selection of competing alternatives.

For this to be achieved an analogical ability measure was required; the core requirement for this task being that it would allow for the manipulation and measurement of complexity in accordance with the thesis's stated aim of testing the relationship between WMC and simple/complex forms of analogical problem. Richland et al.'s (2004, 2006) Picture Analogy task (RPA) was judged to be the most suitable, the other most likely candidates being Sternberg's (1977) People Piece analogies.

- Not only would the RPA allow the ability to investigate factors of complexity, but also distraction (p. 99) allowing the investigation of a third component of AR: that of inhibition.
- Complexity in the RPA was defined by the *presence* of extra relations rather than their *absence* (as is the case of the PPA).
- The RPA uses a real-world visual setting, meaning that it is conceptually closer to a problem analogy than a classical analogy (The PPA being far closer to a classical A:B:C:D format).
- The rules are easier to explain to younger children and it was also considered much more dynamic and fun to play (thus requiring constant unrewarded attention).
- RPA is easier to transfer to a computer based paradigm, which would be necessary later on in the thesis.

Since (at the time of testing) the RPA was a relatively new paradigm in the U.S.A, it was necessary that the RPA be tested for its suitability with a UK sample to establish if a similar pattern of performance data to that reported by Richland was observed (Richland et al., 2004, 2006)[32].

Suitability was judged primarily on the ability to find main effects of age, complexity and distraction in the relational score across the age ranges chosen. It was hoped that interactions between complexity and distraction could also be reported in order to support Richland's assumption that a single underlying system (WMC) may be responsible for mediating these variables in AR. Furthermore, although Richland et al. (2004, 2006) reported that the strength of the interaction between complexity and distraction lessened in older age ranges (thus resulting in an age x complexity x distraction interaction), because in the present study the participants' age range was narrower, a three way interaction between year group, complexity, and distraction was not necessarily expected.

In line with this aim, Experiment 1 was a partial replication which critically appraised the findings from the Richland et al. series of experiments published at the time on a UK sample. Supporting evidence for the reliability of the RPA in a non-US sample has since been gained (Richland et al 2010); however, at the time of the present study it was considered

---

[32] Previously the RPA has shown cultural differences, specifically in the Chinese population where it was thought that a more (visual) written language base allowed Chinese children to perform better at the task (Richland et al, 2010)

important that main effects of complexity and distraction were replicated in a UK sample, as well as interactions between year group, complexity, and distraction.

Experiment 2 assessed the contribution of individual differences in the three WM domains (visual, verbal and global/attentional WMC) to analogical success (relational responses) and failure (featural responses); comparing the effect to that of other more generic (non-capacity centred) systems such as IQ. Crucial to the experiment was whether individual differences in cognitive ability could explain variance in the types of answer given when complexity and/or distraction was manipulated by the RPA.

Experiment 3 investigated the mediatory role of WM further, by attempting to establish if individual differences in STSs or WMC could be best related to task performance, and if either could explain the effects of complexity and distraction described by Richland et al. (2004, 2006, 2010).

Chapter 3 introduced a new approach, assessing the loading effects of relational objects (Experiment 4) and factors of complexity and distraction (Experiment 5). This was to test the complexity-constraint approach (i.e. relational-complexity burdens the WMC system) and to question whether any loading-data from the previous experiments could be interpreted in other ways. Both experiments utilized a (new) computerized touch-screen version of the RPA.

Experiment 4 was a reaction time study which examined whether differences exist between the time taken to select featural and relational responses. The results are discussed in terms of inhibitory responses, preferential selection of objects, and/or the (random) selection of incorrect responses. It also addressed the important question of whether featural responses can be considered *prepotent*.

Experiment 5 adopted a dual-task methodology which aimed to burden WMC beyond the load already imposed by the (computerized) RPA task. It has been claimed (Halford, 1992, 1993, 1998) that current WM measures may not be sensitive enough to discriminate between individual differences in AR tasks beyond the age of four- or five-years. Children beyond this age presumably having sufficient WM resources to process ternary relations as found in the RPA. Because of this argument it was proposed that a dual-task methodology might allow insight into whether the components of WM (visual and verbal STSs, and attentional WMC) are the same as those used to solve the RPA (performance theoretically dropping if a secondary task using the same processes/capacities as the main task is performed concurrently). A dual-task would also enhance the loading aspects of the RPA, meaning that effects of complexity and distraction are more visible (reducing the available WMC to a level where levels of complexity which are usually handled with ease, become difficult or impossible to process).

Finally, Experiment 6 examined whether individual executive processes (as defined by Baddeley, 2007, see page 65 onwards) could offer additional

insight into the potential role of WMC in AR. It being assessed whether

these processes could potentially explain some of the findings observed in

Experiments 1 to 5, and if so, how such an explanation might fit into

relational primacy and capacity constraint models of analogical thought.

Chapter two

2.0 Introduction

In 2004 a new scene based paradigm designed by Richland and her colleagues (Richland et al., 2004; see figures 8 and 9) was developed in America which purported to show a decline in analogical performance through the increase of complexity within an analogical scene. Using this new experimental methodology Richland et al. (2004, 2006) reported significant interactions in the number of relational responses elicited between factors of age and distraction (distraction being the presence of a featural object similar in appearance to the base item), complexity and distraction, but not age and complexity. However, the three way interaction between age, distraction and complexity was significant. The youngest children (ages 3-4) in her study showed the strongest effects of distraction and relational complexity, with the oldest (ages 13-14) the weakest. Richland attributed interactions between complexity, age and distraction to maturational factors constraining analogical performance; suggesting that individual differences in WMC and/or perceptual inhibition were possible candidates for this developmental element.

Richland concluded that, since all relations used in the RPA were both controlled (five per scene) and relatively simple (i.e. understood by the participants), processing errors brought upon by the complexity of the scene must be behind all incorrect (non-relational) choices: ergo domain specific knowledge alone, as proposed by Goswami (1993) could not be behind analogical success. Richland also observed that: (a) participants were

likely to choose featural objects when the 'perceptual distracter' (featural object) was present, and that (b) when available 3-4 year old children were more likely to choose featural than relational objects, a trend which did not continue beyond these age ranges, providing support for the relational shift theory (Gentner, 1988; Ratterman & Gentner, 1998).

2.1 Experiment 1

Experiment 1 investigates Richland's findings (Richland et al., 2004, 2006) by conducting a partial replication study with different year groups in the UK, appraising whether the RPA is a reliable paradigm to further investigate the relationship between WMC and AR - i.e. does it produce the same effects of complexity and distraction Richland et al. 2004, 2006 observed in their American sample?

Despite the aim to replicate some of the effects observed by Richland et al. (2004, 2006) it is important to note that the following experiment was not intended to be an exact replication. The current study uses four year groups (representing the development of children within the British Primary School educational system) with participants aged between 8-11 s (class years 3-6, mean number of participants in each group = 19.5, total participants 78). The original experiment used three year groups (from a North-American population) ranging between 3 and 14 years of age (mean number of participants in each group = 22.6, total number of participants 68). This change was instigated because the age gap between years 3 and 14 years was considered to be too great to capture the process of change and because

more participants within a narrower age range allow a closer view of changes in AR during the critical stage of development (i.e. relational-shift) described by Ratterman and Gentner (1998) and Halford (1993, 1998), and increases in WMC.

Although the number of participants in each group is roughly comparable, (an extra year group being included in the present study taking the total participants to 78 against 68 in the 2004 Richland et al.'s study) it is recognized that by reducing the age range of participants the interactions of age with the factors distraction and complexity reported by Richland may be less pronounced.

Using a sample size of 68 participants, Richland et al. (2004, 2006) [33] found larger effect sizes for the significant main effects of age and distraction (age, $\eta_p^2 = .71$; distraction, $\eta_p^2 = .29$) and smaller effect sizes in the significant main effect of complexity and in the significant interactions (complexity, $\eta_p^2 = .02$; age x distraction, $\eta_p^2 = .09$; age x complexity x distraction, $\eta_p^2 = .09$) in relational scores when looking across all year groups (note that the age x complexity interaction was not significant, $\eta_p^2 = .02$). Although some of these same significant main effects and interactions were also present when the year groups were analysed individually (Ages 3-4, complexity, $\eta_p^2 = .18$; distraction, $\eta_p^2 = .40$; complexity x distraction, $= 0.17$. Ages 6-7, complexity, $\eta_p^2 = .34$, distraction, $\eta_p^2 = .30$. Ages 13-14,

---

[33] Note that the effect sizes presented here are not reported in the original 2004 paper but they are available in the 2006 paper where the same set of data was re-analysed.

complexity, $\eta_p^2 = .42$), some were not statistically significant and showed an overall smaller effect size (Ages 6-7: complexity x distraction, $\eta_p^2 = .12$. Ages 13-14: distraction, $\eta_p^2 = .08$, complexity x distraction, $\eta_p^2 = .07$).

It can be noted that some of the effect sizes that Richland (2004, 2006) obtained for non-significant main effects and interactions were relatively large (e.g. main effect of distraction in ages 13-14, $\eta_p^2 = .08$; complexity x distraction interaction in age 6-7, $\eta_p^2 = .12$) and comparable to the effect sizes that were obtained in the same set of analysis for significant effects (e.g., age x distraction, $\eta_p^2 = .09$; age x complexity x distraction, $\eta_p^2 = .09$). This suggests that Richland's study was possibly under-powered due to a small sample.

Because there is no standard effect size criterion for interpreting partial eta squared (Levine & Hullett, 2002), and given the wide range of effect sizes found in Richland's et al. (2004, 2006), in the current study an *a priori* analysis of required power (G*Power software; Faul, Erdfelder, Buchner, & Lang, 2009) was based on Cohens f medium effect size of f= 0.25[34] (Cohen, 1977) and a power of 0.90 (Cohen, 1977). This analysis determined a minimum number of participants equal to 32 for the within-factor interaction (i.e. complexity x distraction) and equal to 44 for all other within-between factor interactions. The number of participants in the current experiment was 78, that is, above the minimum number required. With a

---

[34] This is equivalent to $\eta^2 = 0.06$.

sample of 78 participants and a power of 0.90 (see above), the smallest

effect size that the current study was powered to detect is f=0.15.

In Richland 2004 and 2006, age groups were chosen in accordance with

Gentner's relational shift (1983) but also Halford's (1992, 1993, 1998)

views on complexity-constraint which suggests that children below the age

of 5 have difficulty processing ternary relationships (such as those used in

the RPA) due to insufficient WMC. Interestingly, Gentner's original theory

predicts a more gradual change to relational responding over the

developmental course, which it was hoped would be visible through the

current year groups chosen in the experiment.

This study used a 2x2 repeated measures design using the factors of

complexity (one or two relations) and distraction (present or not-present)

defined by Richland et al. in the original 2004 paper (which utilized the

relational, featural and relational error scores as the dependent variables)

and the non-experimental (between subjects) factor of year group (4 levels).

The hypothesis was that the three main effects identified by Richland et al.

(2004, 2006): i.e. age, complexity and distraction, would be observable in a

similar UK based study in the relational score, whilst an effect of age and

distraction would be identified in the featural score. The experiment hoped

to observe some of the findings as reported previously in the relational score

by Richland et al. (2004, 2006), namely significant main effects of age,

distraction, and complexity. However, because of the differences in the age

range between the current study and Richland's study (see p.112), it was not

expected to find significant interactions between year group and the factors distraction and complexity (reported by Richland, 2004, 2006), but rather a two way interaction between distraction and complexity (i.e. with no differences among year groups). It was believed that such results might be indicative of WM resources being loaded by relational complexity.

2.1.1 Participants. Seventy-eight participants aged between 8 and 11 (UK class years 3, 4, 5, and 6) were recruited from three primary schools in the North East England. Participants were arranged into the following groups: 20 (participants) from year 3 (mean age = 102 months); 20 from year 4 (mean age = 112.6 months); 20 from year 5 (mean age = 123.95 months), and 18 from year 6 (mean age = 133.83 months).

All participants were screened prior to testing in order to establish that they were able to comprehend the English language sufficiently well enough for them to be able to understand instructions and complete the task and were free from severe mental health difficulties.

Additional requirements for the study were that all children involved were able to see the problems presented to them, were able to indicate the desired response, and were able to point and/or hold and use a pencil when required. All participants underwent strict ethical consent procedures (Appendix A).
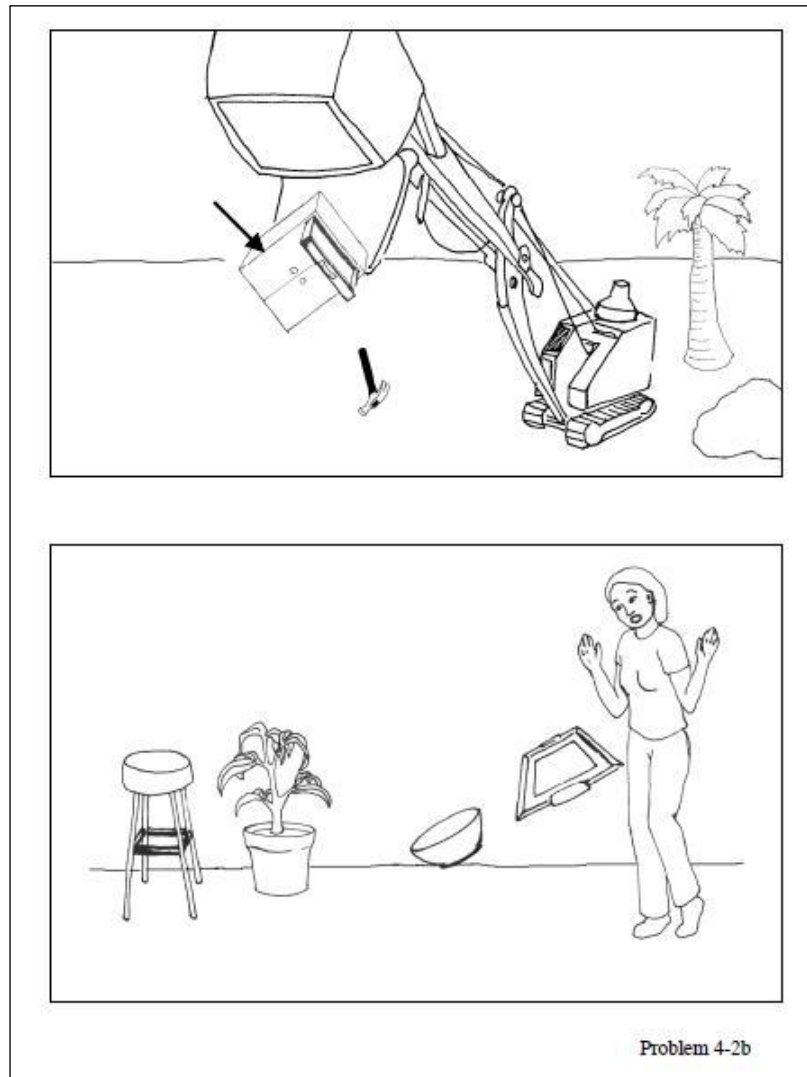
Figure 9. A single RPA problem. This image is problem 4-2b, that is to say the 3rd variation (2b) of the 4th question. Put simpler this is condition R2ND, i.e. it has two relations in it as the base object (the chest of draws) is "dropping something whilst also being dropped". If it was a one relational problem (R1ND [question 4-1a] or R1D [question 4-1b) then there would be no hammer, the relation just being "dropping". It is a no-distracter (ND) because there is not a chest-of draws in the bottom picture (see Figure 8 for an example of the 4 different types of problem). Objects in other conditions are replaced but not removed, extra relational-error objects taking the place of a non-interacting 'other-error' object (i.e. the plant), the number of objects always being 5 in each scene (in the target scene this is the woman, tray, bowl, plant, stool).

117

2.1.2 Materials and Design. A paper version of the Richland

Picture Analogies (or RPA: see Figure 9) was presented on a one-to-one

basis with the experimenter in a quiet environment.

Figure 9 shows an example of the RPA. Each page (i.e. a single problem)

contains a base (top) and target (bottom) picture. During the task the

experimenter points to the object with the arrow and asks the child to find

the object in the bottom picture that is the same (see appendix B for

standardized instructions).

The RPA follows a 2x2 design varying across dimensions of complexity

(distracter versus no-distracter) and complexity (complex or simple

conditions) to create four RPA conditions with 20 problems in each (80

randomly presented items in total): simple no distracter (R1ND), simple

with distracter (R1D), complex no distracter (R2ND) and complex with

distracter (R2D).

Responses are categorized into relational (the object that is doing the same

thing in the same sequential order as the base object, i.e. the tray in Figure

9), featural (objects that look the same as the base object), relational-error

(objects that are almost doing the same as the base object but in the wrong

order, i.e. the bowl in Figure 9) or unrelated objects (i.e. the plant in Figure

9).

Crucial to the design of the RPA the task uses both a controlled number of objects (five per scene) and analogies understood by the youngest participants such as [chased] or [in the middle of] and consists of objects appealing to younger participants (monkeys, dolls etc.).

Despite its origins, Experiment 1 was not a full replication of the Richland et al. (2004, 2006) studies, although it was understood that at the time of writing the RPA was still a relatively new paradigm and additional support for Richland's hypotheses were required. First and foremost, different age ranges were being investigated, it being understood that a broader picture of how analogical reasoning develops beyond the critical age of 5 years proposed by Halford (Halford, 1992; 1993; 1998; Halford et al, 1994, 2004) would contribute significantly to the field in that it would either suggest alternative explanations for AR that were not WMC dependent, or show how AR develops with the incremental increases in general cognition (including but not exclusive of WMC) experienced between Richland's extreme age-ranges of up to 16 years[35]. Secondly, Experiment 1 would conduct additional analyses not seen in the Richland et al. (2004, 2006) studies, such as meta-learning, which would look to see if RPA performance (relational responses) increased over the course of the task. It being possible that some groups of children, possibly older age ranges, were learning the required form of response as the task sequentially progressed;

---

[35] Halford himself suggests that Analogical thought develops between the ages of 3 and 12, but 100% relational responding is not likely until this age range is passed, implying a gradual development of AR ability beyond the time (age 5) in which he indicates that 50% of children are able to fully represent the ternary relationships used in the RPA.

an important point given that each RPA question consists of 4 visually
similar scenes presented throughout the battery, the relational answer being
the same in each.

2.1.3 Procedure. Standardized instructions (see appendix B) were
used following those used by Richland et al. (2004, 2006). Participants
were given 80 problems (20 questions with 4 conditions each), in a repeated
measures design (including all conditions) which had been placed in quasi-
random order using a web-based research programme (Urbaniak & Plous,
1997) which had been adjusted so that no problem was presented adjacent to
an identical question within a separate condition. In such cases conflicting
problems were swapped or the order re-randomised if this was not possible.

2.1.4 Results. The data from Experiment 1 was divided into three parts:
relational, featural, and relational-error responses.

First a correlational analysis was conducted to test for learning effects
(*meta-learning analysis*) within the task, correlating performance with
progression in the RPA. This was important as different forms of response
in the RPA follow patterns which can be learnt (i.e. the relational form of
response is always in the middle, whilst the featural form is always in
isolation); also to a lesser degree the RPA paradigm uses a repeated
measures design with each question being repeated four times- giving the
participant extra experience/the option of repeating the same response
without considering others.

Once the meta-learning analysis was complete a more detailed analysis was performed for three forms of response (relational, featural, relational-error [Rel-error], but not other-errors). Other error responses were very rare. In accordance with the RPA's design, the relational analysis consisted of a three factor mixed ANOVA with two experimental repeated measures factors, complexity and distraction - with the non-experimental between-participants factor of year group, whilst for the featural analysis only complexity would be used due to the lack of distraction data in no-distraction conditions.

Table 3. Mean responses (by year group).

	Mean Relational	SD	Relational%	Mean Featural	SD	Featural%	Featural%*	Mean Rel-Error	SD	Rel-Error%	Mean Other Errors	SD	Other Errors %
All	64.64	8.93	80.80%	5.33	4.04	6.67%	13.33%	7.01	5.11	8.77%	3.01	2.89	3.70%
Year 3	61.9	8.86	77.38%	6.25	3.64	7.81	15.63%	7.45	5.68	9.31%	4.4	2.89	5.50%
Year 4	61.85	10.24	77.31%	5.3	3.85	6.63%	13.25%	9.3	5.78	11.63%	3.55	3.44	4.44%
Year 5	65.5	8.33	81.88%	5.9	5.07	7.38%	14.75%	6.25	3.92	7.81%	2.35	2.60	2.94%
Year 6	69.83	5.61	87.29%	3.72	3.14	4.65%	9.31%	4.83	3.94	6.04%	1.61	1.54	2.01%

[†] Note: Within each experiment, four categories are used: relational (rel), featural (feat), relational-error (rel-error) and other errors (other)

* Featural responses are only present in approximately 50% of the questions; this column therefore represents the true percentage of featural answers in questions where they were present.

Table 4. Mean responses (by RPA condition and year group)

R1ND	Rel Mean	SD	%	Feat Mean	SD	%	Rel-Error Mean	SD	%	Other Mean	SD	%
All	17.24	2.13	86%	N/A	N/A	N/A	1.15	0.98	5.75%	1.62	1.68	8.10%
Year 3	16.35	2.06	82%	N/A	N/A	N/A	1.1	0.91	5.50%	2.55	1.70	12.75%
Year 4	16.75	2.38	84%	N/A	N/A	N/A	1.6	0.94	8.00%	1.65	1.90	8.25%
Year 5	17.6	1.93	88%	N/A	N/A	N/A	1.15	0.88	5.75%	1.25	1.52	6.25%
Year 6	18.33	1.61	92%	N/A	N/A	N/A	0.72	1.07	3.60%	0.94	0.94	4.70%
R1D												
All	15.4	2.55	77%	3.18	2.32	16%	1.13	1.00	6%	0.29	0.56	1%
Year 3	15.1	2.02	76%	3.6	2.06	18%	0.85	0.81	4%	0.45	0.69	2%
Year 4	14.8	2.63	74%	3	2.03	15%	1.8	1.11	9%	0.4	0.68	2%
Year 5	15.05	3.07	75%	3.65	3.05	18%	1.15	0.81	6%	0.15	0.37	1%
Year 6	16.78	1.96	84%	2.39	1.88	12%	0.67	0.91	3%	0.17	0.38	1%
R2ND												
All	16.81	2.56	84%	N/A	N/A	N/A	2.23	2.08	11%	0.96	1.20	5%
Year 3	16.1	2.85	81%	N/A	N/A	N/A	2.6	2.48	13%	1.3	1.26	7%
Year 4	16.1	2.95	81%	N/A	N/A	N/A	2.65	2.41	13%	1.25	1.48	6%
Year 5	17.1	2.05	86%	N/A	N/A	N/A	2.1	1.68	11%	0.8	1.06	4%
Year 6	18.06	1.83	90%	N/A	N/A	N/A	1.5	1.47	8%	0.44	0.71	2%
R2D												
All	14.45	2.90	72%	2.03	1.78	10%	2.38	2.15	12%	0.14	0.35	1%
Year 3	13.6	3.19	68%	2.55	1.61	13%	2.75	2.51	14%	0.1	0.31	1%
Year 4	13.4	3.05	67%	2.25	1.86	11%	3.1	2.38	16%	0.25	0.44	1%
Year 5	15	2.60	75%	2.05	2.06	10%	1.8	1.91	9%	0.15	0.37	1%
Year 6	15.94	1.98	80%	1.17	1.30	6%	1.83	1.30	9%	0.06	0.24	0%

Table 5. Mean relational scores (by RPA condition)

	R1ND	R1D	R2ND	R2D	R1	R2	ND	D
Mean	17.23	15.40	16.81	14.45	32.63	32.01	43.04	29.85
SD	2.13	2.55	2.56	2.90	4.280	5.271	4.317	4.957

Table 6. Mean relational scores (by year group)

	YEAR 3	YEAR 4	YEAR 5	YEAR 6
Mean	15.29	15.26	16.19	17.28
SD	2.23	2.50	2.04	1.40

Meta-learning analysis. A bivariate correlation comparing the proportion of responses within each response-category was run within age and RPA conditions against the sequential question number (1 being question 1, 80 being question 80 etc.), the intention being to show the patterns of responses as the participant progressed through the RPA task. In general, no effects of learning were observed within the task for any response category. However, an analysis within year groups did observe a significant negative correlation between the sequential questions order and the total number of relational responses in Year 5 ($r = -.242$, $p < 0.05$), and the total number of relational answers in Year 3 was also found to be approaching significance ($r = .193$, $p=0.056$).
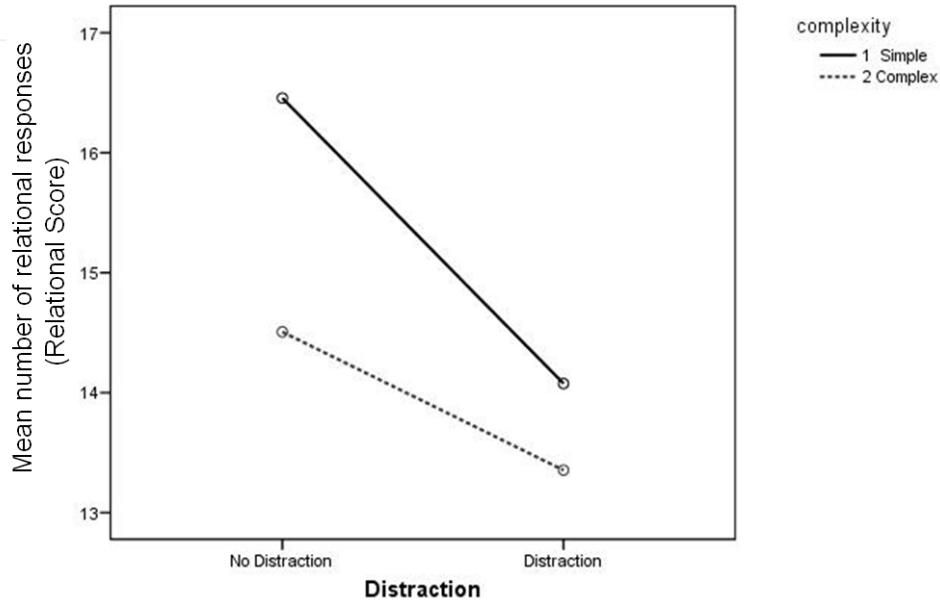
Figure 10. Complexity/distraction interaction in the relational score.

Relational analysis. The factorial ANOVA for the number of relational responses given (R^s) revealed a significant main effect of complexity (see Table 5, columns R1 and R2) ($F_{[1,74]} = 11.26$; $p<0.05$, $\eta_p^2 = .13$), distraction (see Table 5 columns D and ND) ($F_{[1,74]} = 175.004$; $p<0.001$, $\eta_p^2 =.70$), and year group (see Table 6) ($F_{[3,74]} = 3.84$; $p<0.05$, $\eta_p^2 =.14$). Pairwise contrasts revealed that the significant differences in year groups were found between year 3 and 6 ($p<0.05$), and year 4 and 6 ($p<0.05$). The interaction between complexity and distraction (see Figure 10) was also significant ($F_{[1,74]} = 4.77$; $p<0.05$, $\eta_p^2 =.06$), while there were no significant interactions between complexity and year group($F_{[3,74]} = 0.69$; $p>0.05$, $\eta_p^2 =.03$), distraction and year group ($F_{[3,74]} = 0.74$; $p>0.05$, $\eta_p^2 =.03$), or complexity, distraction and year group $F_{[3,74]} = 2.25$; $p>0.05$, $\eta_p^2 = .08$).

Table 7. Mean featural scores (by RPA condition)

	R1ND	R1D	R2ND	R2D
Mean	N/A	3.18	N/A	2.03
SD	N/A	2.32	N/A	1.78

Paired sample t-tests between RPA conditions across all year groups found significant differences between conditions R1D and R1ND (t(77) = -8.44, $p<0.001$, η² = .48), R2D and R2ND (t(77) = -13.06, $p<0.05$, η² =.69) as well as R1D and R2D (t(77) = 3.67, $p<0.001$, η² =.15), R1D and R2ND (t(77) = -5.17, $p<0.001$, η² =0.26). This illustrated the interaction between complexity x distraction in the relational score (see Figure 10) by demonstrating that participants gave more relational responses in no-distracter compared to distracter conditions whilst also eliciting fewer relational responses in complex distraction conditions.

Featural analysis. The same procedure for the number of featural responses (F^s, see Table 7) minus the distraction condition, found a significant main effect of complexity ($F_{[1,74]} = 49.11$; $p<0.001$, $\eta_p^2 =.40$) but no main effect of year group ($F_{[3,74]} = 1.65$; $p>0.05$, $\eta_p^2 =.06$).

Table 8. Mean relational-error scores (by RPA condition).

	R1ND	R1D	R2ND	R2D	R1	R2	ND	D
Mean (all years)	1.15	1.13	2.23	2.38	2.28	4.73	3.38	3.51
SD	0.98	1.0	2.08	2.15	1.728	4.063	2.70	2.66
Year 3 mean	1.10	.85	2.60	2.75	1.95	5.50	3.70	3.60
SD	0.912	.813	2.479	2.511	1.504	4.685	3.16	2.89
Year 4 mean	1.60	1.80	2.65	3.10	3.40	5.90	4.25	4.90
SD	0.940	1.105	2.412	2.382	1.729	4.734	2.86	3.08
Year 5 mean	1.15	1.15	2.10	1.80	2.30	3.95	3.25	2.95
SD	0.875	0.813	1.683	1.908	1.342	3.379	2.15	2.11
Year 6 mean	0.72	0.67	1.50	1.83	1.39	3.44	2.22	2.50
SD	1.074	0.907	1.465	1.425	1.787	2.770	2.26	1.82

Table 9. Mean relational error scores (by year group)

	YEAR 3	YEAR 4	YEAR 5	YEAR 6
Mean	1.82	2.28	1.55	1.18
SD	1.45	1.40	0.98	0.97

Relational-error analysis. For relational-errors given (the relational-error score, see Table 8) significant main effects of complexity ($F_{[1,74]} = 33.05$; $p<0.001$, $\eta_p^2 =.31$) and year group ($F_{[3,74]} = 2.72$; $p<0.05$, $\eta_p^2 =.10$) were observed revealing that after an initial rise at around class Year 4 (when more relational errors were made) - the number of relational errors decreased with age, presumably (given the relational data) as more children were able to correctly choose which form of relational response was correct. (see table 9) However, pairwise contrasts revealed that the differences in year groups were statistically significant only between Year 4 and 6 ($p<0.05$).

 No main effects were found for distraction ($F_{[1,74]} = 0.46$; $p>0.05$, $\eta_p^2 =.006$) and no interactions were found between complexity and year group ($F_{[3,74]} = 0.95$; $p>0.05$, $\eta_p^2 =.04$), complexity and distraction ($F_{[1,74]} = 1.16$; $p>0.05$, $\eta_p^2 =.02$), or complexity, distraction and class year ($F_{[3,74]} = 0.94$; $p>0.05$, $\eta_p^2 =.04$).

Relational vs. featural responses. In the Richland et al.'s (2006) study, relational-errors and featural responses were compared in R2D conditions (i.e. a condition where both complexity and distraction were present) using a one-way ANOVA, this was done to see which form of response represented the greatest attentional appeal/distraction. A variation of this methodology was performed here in order to see which form of response was preferable for each year group, and how this preference developed.

Here an ANOVA was performed on condition R2D with the two experimental measures of the relational and featural scores and the non-experimental between groups factor of age. It was predicted that a main group effect of type of response would be found due to the large difference in size between relational and featural responses.

As expected there was a significant main effect of response type ($F_{[1,74]} = 538.39$, $p<0.001$, $\eta_p^2 = .88$) meaning that relational responses were preferred over featural responses, but no interaction between response type and class year group ($F_{[3,74]} = 1.72$, $p>0.05$, $\eta_p^2 = .07$) meaning that although there was a difference, some class years did not prefer certain types of response.

2.1.5 Discussion. Richland et al.'s (2004, 2006) study makes two basic assumptions regarding her data. The first regarding *complexity* is that performance was better in non-distraction/complexity conditions, and that due to control of the number of objects (five per scene) and (simple to understand) relations used, domain specific knowledge alone could not

account for the complexity findings. Secondly regarding *Maturational Factors and the Relational shift,* Richland postulating that relational shift hypothesis was supported in that, beyond ages 4-5 (the key years highlighted by Ratterman & Gentner; 1998), a trend was observed wherein the proportion of relational responses chosen over featural was greater in the older class years. As a result of both of these claims Richland hypothesized that the interaction between age/complexity and distraction was mediated by maturational factors, which were hypothesized as being WMC and inhibitory skill. Both of these claims will be reviewed in light of the present data.

(1) Complexity and Distraction

Relational score. Children on the present experiment gave more relational responses in the no distracter (R1ND and R2ND) than distracter conditions (R1D and R2D) and gave more relational responses in simple (*r1*) compared to complex (*r2*) conditions. As predicted by the relational shift (Gentner, 1988; Ratterman & Gentner, 1998), the youngest participants (Year 3 and 4) also elicited the lowest proportion of relational responses and the oldest (Year 6) the highest.

The main effects of year group, distraction, and complexity found in Richland's study (2004, 2006) were all replicated here, with the effect size for year group and complexity being smaller in the current study (year group: $\eta_p^2 = .13$ in the current study vs. $\eta_p^2 = .71$ in Richland's study; complexity: $\eta_p^2 = .13$ in the current study vs. , $\eta_p^2 = .22$ in Richland's study),

and the effect size for distraction being larger in the current study ($\eta_p^2 = .70$

vs. $\eta_p^2 = .29$ in Richland's study). The two way interaction between age and

distraction and the three way interaction between age group, distraction, and

complexity found in Richland et al.'s study (2004, 2006) were not replicated

here. This was assumed to be due to the age range chosen in the study,

which offered a restricted age range for any age-related change to be

identified. However, it should be also noted that the effect sizes obtained by

Richland et al. in the interactions involving the factor year group were small

and partially comparable with those found in the current study (year group x

complexity: $\eta_p^2 = .03$ in the current study vs. $\eta_p^2 = .02$ in Richland's study;

year group x distraction x complexity: $\eta_p^2 = .08$ in the current study vs. $\eta_p^2 = .09$ in Richland's study; but note that the effect size in the year group x

distraction interaction was smaller in the current study: $\eta_p^2 = .03$ in the

current study vs. $\eta_p^2 = .09$ in Richland's study)

In strict accordance with Gentner's (1989) view of similarity space (where

featural 'distracter' objects are described as relational objects with less

bound descriptive information) the interactions observed within the data can

therefore be broadly understood to be in accordance with Richland et al.'s

(2004, 2006) interpretation of SMT[36], an effect of complexity apparently

having been demonstrated.

---

[36] e.g. if the participants already held the similarity-constraint in LTM, then lowered levels
of performance in complex and distracter conditions reflect increased task difficulty.

Intriguingly conditions of distraction, above those of complexity, were considered the most difficult[37], providing additional support for Richland's idea that inhibitory mechanisms may underlie AR, even though the original study did not find the same effect.

Yet despite these conclusions, other interpretations of the data do exist beyond the complexity-constraint perspective. For instance, the rising proportion of relational over other forms of response across year groups may not arise from the relational shift or the increased ability to process relations per se, but from participants being primed in standardized training to answer relationally, the effect of age not being from a growing dominance of relational information, but developing cognitive abilities used in to interpret/follow that training.

Other issues regarding relational responding surround how complexity may be measured. In the RPA relational objects are the only form of objects that interact, the idea of complexity being to increase the descriptive dimensions of the target object (i.e. turning the similarity-constraint from 'chased' in the R1ND condition into 'chasing and being chased' in R2ND). However, target objects do not always have to be described by their function, their existing the possibility that they are being identified by a single comparatively low-level relation which focuses on their position, i.e.

---

[37] In terms of a sequential order for highest proportion of relational responses the order was R1ND, followed by R2ND, R1D and R2D. 70.3% of the variance in the relational score was accounted for by the main effect of distraction, with 13.2% accounted for by the main effect of complexity.

'in the middle.' Secondly, although the number of objects in each scene is technically always five, not all objects are relevant, so it is entirely possible that *r2* scenes are increasing the number of relevant response.

In order to show the validity of these latter prospects we must look at the patterns of other forms of response within the RPA.

Featural score. For featural responding a main effect of complexity and distraction was identified as well as an interaction between complexity and distraction, but no effect of year group, the effect of class year group on the F^s being too small to detect given the present experiments sample size. This could indicate that the F^s is uninfluenced by maturation, however given the low variance in the score a more likely conclusion was that the RPA lacked power for this form of analysis, especially when the sample was separated into small groups.

This observation appeared to be unique to the present study as in 2004, 2006, Richland (with fewer participants but a lower age-floor) had previously identified a main effect of age within this form of response, as well as an interaction between age and complexity. Despite this claim, age effects were not entirely absent as a significant negative correlation between age and F^s that had been identified in the R2D but not R1D conditions, a finding which at first glance, made good theoretical sense if condition R2D required more developmentally accrued processing power because of its complexity.

Yet one aspect of the analysis came as a surprise: that of the effect of complexity (on F^s). In a finding that directly conflicts with Richland et al (2004, 2006) a higher proportion of featural responses were consistently reported in condition *r1* than *r2* across the majority of year groups. This is directly contradictory to Waltz et al.'s (2000) interpretation of the complexity-constraint hypothesis (Halford, 1992, 1993, 1998), which states that:

"…any manipulation that reduces available working-memory capacity will make it more difficult for reasoners to compute relational mappings, and hence will increase the proportion of less complex attribute mappings in situations in which the mapping is ambiguous." (Waltz, 2000, p. 1206)

One possible reason for this was general task difficulty, R1D somehow being more difficult than R2D, a factor which seems unlikely given Richland's original hypothesis. Since *relational responding* was slightly but not significantly higher in condition R1D than R2D, this was dismissed, instead it being assumed that condition R1D somehow placed unique demands on the reasoner, which made featural responses more attractive than other forms of error (non-relational response). This could be because of the forced-choice dynamic in such problems which clearly highlighted two forms of response which are potentially correct according to two opposing logical paths, a theory which has some providence given Richland

134

et al.'s (2004, 2006) proposed (but unexplained) effect of inhibition within the RPA. Inhibitory skill possibly kicking in situations where a) less processing time is required, b) where a type of response is more prominently detectable or c) where one form of response is proponent.   In such circumstances it is common sense to assume that if one of the visible forms of response is similar to the object being compared and held in mind, rapid selection of that response would require some form of base inhibition.

Interpreting the data more generally however, the number of featural responses elicited in any condition/class year group was considered small (the ratio of featural to relational answers across the whole of the RPA approaching 1-4, the mean percentage of featural responses per participant being just 6% or 5 out of a possible 80), indicating that the sample chosen was not particularly prone to this form of response. At this stage it was not known whether this was potentially due to the age range chosen (as predicted by Richland et al., 2004, 2006) and/or (as suspected given the standardized instructions) the paradigm naturally primed participants to avoid featural forms of response. In either case it was apparent that the cognitive development of individual differences across childhood could be explained by individual differences in ability, through increased ability to process relationships and/or to comprehend/follow instruction, or indeed both.   Yet whilst this clearly underlined a future direction for research, the interpretation of the present featural effect was less clear cut.

Here the low level of responses brought F^s within range of other forms of response (the mean for relational-error and other-error responses being 7 and 3 respectively). Elucidating further on this, no significant difference was found between the proportion of featural and relational-error responses within condition R2D (half of all conditions involving featural responses), highlighting the possibility mentioned earlier that non-relational (incorrect) responses are chosen in complex conditions not because a processing error has arisen due to *object* complexity, but because of *scene* complexity, Condition R2D having more (potentially) relevant forms of responses - increasing the task ambiguity.

Relational-error score. Within the relational-error score a main effect of age was found as well as complexity, although again no interactions (between age, complexity or distraction) were observed. Once more a lack of power was indicated for this form of error analysis; however unlike the F^s the variance within the score appeared large enough to consider that older participants may make less relational errors (a hypothesis supported by the correlation analysis). At first the main effect of complexity also looks meaningful but for this analysis it was hypothesised *a priori* that despite the control of five objects per scene the higher proportion of relational-error responses available in *r2* compared to *r1* problems would mean that the number of relational-error responses would almost certainly be greater in complex conditions. This effect being presumed to be likely to occur because there were more relational-error responses than any other form of response and because relational-error objects were more task-appropriate

than (i.e. they shared some of the similarity-constraint observed in the base item).

In complex *r2* problems, although marginally more relational-error responses were elicited in distraction conditions compared to no distraction conditions, the difference was not significant indicating that the presence of a distracter generally did not increase the likelihood of a relational-error response in either complex or simple RPA conditions. This in turn suggests that the children understood what was being asked of them in the task (i.e. to select relational responses) but in complex conditions were having difficulty deciding which relational response represented a best fit.

Maturational factors (the relational shift). In her study, Richland proposes that the age based interactions with factors of complexity and distraction were due to maturational factors which "may interact to constrain children's capacity to perform successfully on picture analogies that require more WM or perceptual inhibition" (Richland et al., 2004; p. 153). As such, the 2004, 2006 study was thought to provide evidence for the relational shift wherein, according to some researchers (Morrison et al., 2006, 2010), at around 4-5 years old children begin to incrementally prefer relational responses over featural.

Although no such age based interactions were observed, the current study supported this developmental claim in that a significant main-effect of age was identified in relational responding. Additionally, although the featural

data crucially failed to replicate the same maturational assumptions it was still somewhat consistent with Richland's hypothesis, especially given that no significant interactions were observed in the relational score for Richland et al.'s (2004, 2006) 6-7 year old year group.

This failure to precisely replicate Richland et al.'s findings was almost certainly because floor/ceiling effects were present in the original trial, wherein a disparity clearly existed between the lowest year group (3-4 years old) and the highest (13-14 years old), the youngest participants eliciting a far higher proportion of featural responses than was observable in this experiment in the 8-9 year old group. Such a lack of featural-age data in the present sample was of course not entirely unexpected, whilst Richland's middle year group (6-7 years old) was always likely to under-perform compared to the youngest age presented here (8 years old). There has been additional evidence to suggest that beyond 4-6 years children reason by relational-relations rather than attribute-based relations (Morrison et al., 2006, 2010), however (due to the age ranges chosen here) in this case insufficient evidence has been provided to support this. It therefore seems likely that young children lack the cognitive processes or capabilities necessary to adequately complete the RPA, presumably because they have less executive resources with which to actively manipulate and map complex analogical problems.

Maturational factors (relational primacy perspective). As we have seen in chapter 1, an abrupt naturalistic change in thinking styles is not necessarily

138

supported by knowledge-based theorists who predict that children can answer relationally even at early ages (Goswami, 1992; Goswami & Brown, 1989). Or indeed by some relational-shift theorists who postulate a more gradual change, the relation shift occurring "at different ages in different domains, depending on domain knowledge" (Gentner & Ratterman, 1991, p. 456). This leaves the possibility that the relational shift at these ages is not occurring because of the inability to process relations, but because they are failing to understand the task.

Following this perspective age based changes in WM resources or cognitive abilities (capacity based or otherwise) may still be applied to the RPA, albeit indirectly through such spheres as meta-learning, interpreting task ambiguity (i.e. selecting relevant responses from less relevant options) or the ability to comprehend the instructions used, interpret experimental demand and/or maintain goals in the face of highly demanding parallel processes/distractions.

Intriguingly Richland et al. (2006) refute the idea that children were unable to understand the RPA, speculating that instead the effect is mainly due to the capacity based limitations of WM in the active processing of complexity. However this claim was focused almost purely on the words used (language in analogical tasks being notoriously vague) and is directly contrary to Goswami's speculation (Brown 1989; Goswami, 1992; 1994) that given the right instructions and similarity-constraint, most children can complete an analogical problems by answering relationally.

Meta-learning. In general, no effects of learning were observed within the task for any response categories, meaning that the RPA was considered to be mostly free of meta-learning, however Year 5 did demonstrate an effect wherein more relational responses were elicited near the end of the task than the beginning, meaning that some learning was taking place in older participants. This was not entirely unsurprising, as learning is almost certain to take place in one form or another within most, if not all, sequential reasoning tasks such as the RPA (lending limited credence to the idea that the maturational factors Richland described, may not entirely be due to the loading of WMC effective active processing). However, the low-strength of this one correlation was not considered strong enough to represent a major effect, and so the concept of WM effecting meta-learning within these class years in the RPA was rejected.

Efficacy of the RPA. The RPA offers the almost unique potential to explore the effects of complexity and distraction on WM. Generally the data presented here supported Richland et al.'s (2004, 2006) claims that such factors mediate AR, the relational analysis in particular indicating age-based interactions which were in-line with SMT. Yet even non-complexity based interpretations of the findings promote the concept that further psychological insights into WM and AR may be achieved through the paradigm.

As ever, concerns regarding the RPA were raised, particularly in regard to the prospect of using the task to measure age based effects on featural responding. For this reason it is suggested that although further assessment be conducted in this area (in order to investigate the lack of featural data observed here) future analyses may benefit more from a relationally focused assessment.

2.1.6 Conclusion. In conclusion, although these findings differ from the Richland et al. (2004, 2006) study in a number of important ways, this experiment generally supports the hypothesis that factors such as complexity and distraction increase task difficulty in the RPA. Like the earlier Richland experiment the data also seems conclusive in that it provided strong support against Goswami's (1993) relational primacy perspective, participant's particularly high success rates suggesting that they were able to successfully identify the required similarity-constraint whilst still making errors in the task, suggesting that processing difficulties were occurring during later mapping, rather than initial recognition.

However sufficient questions have been raised in regards to possible alternative interpretations of this data to warrant further investigation. One area in particular which requires additional attention is the maturational factors discussed by Richland et al. (2004, 2006) in the original paper. Given our knowledge of the existing relationship between STM and reasoning it seems self-evident to propose that WM is in some way mediating analogical thought. Yet if we wish to infer that this mediatory-

role is due to a constraint-effect of relational-complexity (Halford 1992, 1993, 1998) in accordance with Gentner's taxonomy (Gentner, 1983, 1988, 1989) - then it should be a requirement for the research to establish a connection between WMC and the proportion of relational responses elicited.   The alternative is to adopt a more general perspective wherein children with higher WMC have an inherent advantage over others as they may be better able to maintain a problem during processing regardless of view of structural complexity (Gilhooly, 1998; Kyllonen & Christal, 1990) a theory which is entirely in accordance with the claims of Goswami (1993).

This experiment therefore paves the way for a series of experiments designed to test both the hypothesis that WMC plays an important role, and Richland's (2004, 2006) argument that RPA complexity data could potentially provide evidence for the involvement of WM and its processes in AR. It is clear that whilst the findings observed here support the idea of a 'relational shift' in that the data may be seen as an extension of Richland's observations in the 2004, 2006 study, sufficient doubt exists in the current experiment (due to the relatively high proportion of relational over featural responses) to suggest that the role of working memory/intelligence may not be as clear-cut as previously suggested. The next experimental step should therefore be to investigate the validity of each perspective by directly comparing WMC and non-capacity based measures of ability to the performance in the RPA, to attempt to further clarify the role of each in AR.

2.2 Experiment 2

Experiment 2 was designed to assess the involvement of WMC in AR whilst controlling for generally ability (hereby referred to as IQ[38]).

As stated in the literature review, the most widespread account of AR-WM is the complexity-constraint theory. This states that any cognitive manipulation that reduces WMC makes it more difficult to successfully map analogical problems. Increased object complexity makes additional demands on these WM resources, possibly leading to an increased selection rate of less-complex attribute based objects during the mapping processes (Waltz et al., 2000). Individuals with higher WMC may also have additional resources with which to devote to analogical tasks, meaning that their reasoning functionality may be less affected by complex relationships than their low WMC counterparts (Gentner, 1983, Richland et al., 2004, 2006).

It is understood that WMC increases throughout childhood (Alloway et al, 2004; 2005). Halford (1992, 1993, 1998) and Richland et al. (2004, 2006) claimed that WMC may act as an age-centric constraint on the number of relations that can be processed in parallel. This assertion was partly supported in Experiment 1: it being established that older participants responded relationally at a higher rate than their younger counterparts,

---

[38] The word IQ is used here cautiously. As described in the previous chapters, 'IQ' or 'intelligence' may be seen in some instances as being synonymous with WM, however in this instance it is used to refer solely to processes which do not rely on capacity as their chief component. It includes tasks which have WM components or which primarily rely on speed of processing, the latter being known to be associated with WMC.

specifically under conditions of distraction or high complexity. Yet despite this observation, a complexity-constraint effect has not been practically demonstrated using WM measures.

It was discussed at the end of Experiment 1 that other factors besides WMC might potentially explain the results observed both in the previous experiment and the Richland et al. (2004, 2006) study, and that WM did not necessarily have to rely on a complexity-constraint model to mediate AR[39]. More critically it was also reported that increased complexity resulted in less featural responses, not more; possibly implying that individual conditions in the RPA place different demands on cognition other than those anticipated by a complexity based approach

These observations provided two clear directions of research for Experiment 2. Firstly, it was seen as necessary to associate WMC with performance in the RPA in order to test the validity of any WM/AR account. Secondly, precisely how WMC might influence this form of thinking should also be clarified; particularly investigating Richland's (2004, 2006) proposal that the maturational factors reported in the experiment could be explained by capacity constraints in the WM system.

A new experiment was therefore designed to directly compare the involvement of WM and IQ in the RPA in an attempt to gain theoretical

---

[39] A WM effect might also represent the application of mental ability to interpret experimenter demand, or the overall functioning efficiency/speed of cognitive systems, neural networks and selective executive processes.

insight into AR.   To this end a battery of eight tasks, consisting of four

measures of IQ and four measures of WM, were presented alongside the

RPA.

The IQ tasks were chosen in order to be representative of core aspects of

commonly encountered and widely available intelligence tests for children:

focusing on both visual and verbal tasks- it being a requirement that each IQ

task was widely accepted in the research community.   The tasks being

selected on the grounds that they would appeal to young children as a

puzzle, game or familiar form of school administered vocabulary test- whilst

also being well within the ability of children younger than those being tested

here (in case of future research pursuing these year groups). It was reasoned

that these tasks should be not be overly representative of a single factor,

therefore two tasks required a verbal response; a single word reading task

(Wechsler, 2005) and a vocabulary task (Wechsler, 1991), and two required

a visual or 'hands on' response; a block design task and a coding task

(Wechsler, 1991). All the tasks required the ability to think quickly and

accurately (see materials and design for a review of these measures below).

The WM measures were chosen in order to represent the different

modalities of Baddeley's traditional model of WM. Taken from the

Automated Working Memory Assessment programme or 'AWMA'

(Alloway, 2007), two tasks were visually based (the Odd One Out task and

the Mr.X task) and two were verbally based (the listening recall task and the

backwards digit task). The battery of WM tasks selected were considered

not too difficult for young children to understand and they could easily be explained by the experimenter during an instruction phase (see materials and design below).

The strength of the relationships between the resulting eight variables and the R^s (the central measure of AR in the RPA) was then to be examined through linear regression modelling and correlational analyses.

In a further attempt to answer more directly how covariates may interact with factors such as complexity and/or distraction, a series of ANCOVAs were performed to investigate whether high/low IQ and/or WM could account for individual differences in relational responding within the different conditions. These, and all subsequent ANCOVAs, were carried out in accordance with the recommendations by Delaney and Maxwell (1981), means centring the covariate in order to protect it from altering the main effect of the repeated measure (see referenced paper).

The R^s was chosen as a measure of AR in strict accordance with the traditional perspective within this field (chapter 1) which states that only relational responses typically dichotomize analogical thought. However Waltz et al.'s (2000) recent claims that attribute based responses may be a side effect of reduced WMC warrants further investigation. For this reason the featural score (F^s) was also included in the main analysis.

Since no age based interactions were observed in the first study, this experiment would extend beyond the lower age range tested in Experiment 1 to include Year 2 children. This decision was made primarily because of analogy's cornerstone relationship with cognitive development (Gentner, 1988, Halford, 1992, 1993, 1998; Piaget et al., 1977) but also because the previous years tested may have been above that suggested for the relational-shift (Morrison et al., 2006). By lowering the age range to 6/7 the change from featural to relational responses should hopefully be more obvious in the data and more comparable to Richland's original study, without going so low as to risk a greater level of task misunderstanding.

2.2.1 Participants. Seventy-four children aged between 5 and 11 years (UK class Years 2, 3, 4, 5 and 6) were recruited from three schools in the North East of England. Participants were arranged into three year groups: 20 aged 6-7 years-old (mean age = 78.75 months), 31 aged 8-9 (mean age = 101.19 months) and 23 aged 10-11 (mean age = 124.87 months) and were recruited in accordance with the same ethical recruitment/consent criteria previously mentioned for the earlier study.

2.2.2 Materials and Design. Eight subtasks in addition to the paper version of the RPA used in Experiment 1 were administered to each participant on a one-to-one basis: four measures of IQ, representing processing speed, visual ability, verbal ability and vocabulary; and four measures of WM.

147

Cognitive ability tasks.

Basic reading test (verbal measure of ability). The basic reading test is a

subtask taken from the standardized WIAT battery (Wechsler, 2005), the

task consisting of 55 sequentially administered written words which the

child is required to correctly read aloud.

Words systematically increase in difficulty as the child progresses

(difficulty being defined by both word length and frequency within the

English language), a maximum of ten seconds being allowed for the correct

pronunciation of each item. The task is concluded when a child reaches the

end of the task or when five consecutive errors are made. Task reliability for

the Basic reading test has been reported as being .87 (Wechsler, 1992).

Vocabulary (measure of verbal fluency). Taken from the WISC-IIIR

(Wechsler, 1991) the vocabulary subtask is a list of 30 words for which the

children are asked to provide definitions. Words are presented verbally one

at a time and are of gradually increasing complexity, the task being

discontinued if the child reached the final item or if four consecutive failures

(scores of 0) are made. In accordance with the WISC manual, each item is

scored 0, 1, or 2 depending on the level of understanding, with a score of 2

representing full understanding of that particular word.

Although also a measure of verbal ability, the vocabulary test is also thought

to be an excellent measure of general intelligence (Prifitera et al., 2005;

Flanagan & Kaufman, 2004, Sattler & Dumont, 2004) and as such is

perhaps the most broadly applicable subtest to a variety of cognitive abilities. This association with other intelligence tasks is thought to due to the increased efficiency of neural networks demonstrated in verbal fluency measures (Kail & Salthouse, 1994; Salthouse, 1996). Verbal fluency is a temporally-centric skill known to place demands on the CE through the rapid recall and selection of relevant information (the speed of processing effect being observable in its frequent correlation with other IQ tasks) and is also known to be associated with WM (Chuah & Maybery, 1999), the selection processes of which are crucial to the WM system.

Reliability scores for the subtask have been favourable with a mean reliability coefficient of .87 having been reported (Goldstein & Hersen, 2000).

Coding (measure of processing speed). The coding subtask from the WISC-IIIR (Wechsler, 1991) uses a code-key of numbers, each representing a simple geometric shape, and a list of 119 numbers. Going from left-to-right children were asked to draw in the correct shapes for as many numbers as they could within two minutes; time bonuses being given for faster performances. Children aged seven and under, were also asked to complete an additional easier page of 65 items, children over this age being automatically accredited the points for the full completion of this easier task as a method for controlling for age.

The coding task is understood to be a measure of motor and processing speed, general processing ability, STM and attention (Weiss et al., 2006) and has been cited as having a moderate-high reliability coefficient of .79 (Goldstein & Hersen, 2000). It was chosen for this task because it adequately represented a mix of functions from core cognitive areas (IQ and WMC): processing speed being seen as synonymous with processing power through its enhancement of individual systems (Kail & Salthouse, 1994).

Block design (visual ability measure). The final subtask from the WISC-IIIR (Wechsler, 1991) is the Block-Design task, which is thought to be a good measure of visual/spatial ability (Flanagan & Harrison, 2005; Kaufman, 1994;) in particular executive functioning (Lezak, Howieson, & Loring, 2004) and visual assessment and planning skills (Brown, Brockmole, Gow & Deary, 2012).

The block design task consists of 12 individually administered patterns of increasing complexity (number of blocks, complexity of design) which the child is asked to reproduce uses coloured plastic blocks. Each pattern has a time limit within which the child must complete the task, points being scored the faster the child can complete the task. If the child does not complete the task then zero points are scored, two failures being allowed before the task is discontinued. Reliability for the block-design task is good, with an average reliability coefficient of .87 (Goldstein & Hersen, 2000).
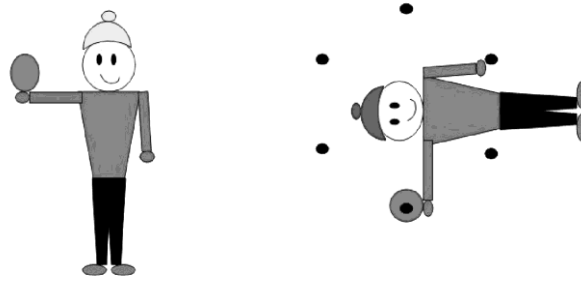
Figure 11. The Mr.X task. As each sequence appears participants are asked to determine whether the ball in the hand of the Mr.X on the right is in the same hand as the Mr.X on the left; before recalling (un-cued) the location of all the balls from the sequence that were on the right.

Working memory tasks:

Backwards-digit task (verbal). The Backwards-Digit task is a verbal WM subtask from the computer based Automated Working Memory Assessment (AWMA; Alloway, 2007). For this task a child was given a list of digits and asked to recall them in backwards order. For the first trial, the span would consist of 2 digits, with an additional digit being added each time a child got four or more spans (out of a possible six) correct. If less than four out of six were answered, the task was concluded. For the Backwards-Digit task a low-moderate reliability coefficient of .64 has been reported (Alloway, 2008).

Listening-recall task (verbal). The Listening-Recall task is a computer based measure of verbal WM taken from the AWMA battery (Alloway,

2007). Children were given a sentence to remember and then asked to say whether it was true or false before recalling the last word of the sentence, the number of sentences increasing by one each time a child answered four out of six questions correctly. For this subtask the average reliability coefficient has been reported as being .81 (Alloway, 2008).

Mr.X task (visual). A subtask from the AWMA (Alloway, 2007), the Mr.X task uses two cartoon figures on the left and right hand side of the screen, both of whom are holding a ball in one hand. Whilst the Mr.X on the left (identified continuously as having a yellow hat as opposed to blue on the right) remains at the same orientation across all Mr.X trials, the figure on the right is rotated through one of six pre-determined points, with the ball assigned to either the right or left hand (Figure 11).

For each problem the child is required to say whether the ball in the hand of the Mr.X on the right is in the same hand as the ball in the hand of the left Mr.X. The picture is then removed and the child is then asked to recall the (right Mr.X) ball location from the six possible locations. Once four sequences were complete, an additional Mr.X was added, meaning that more ball locations had to be remembered. The average reliability coefficient of the Mr.X task is stated as being .77 (Alloway, Gathercole, Kirkwood, & Elliott 2008).

Odd-one-out task (visual). A subtask from the AWMA (Alloway, 2007), the *Odd*-One-Out task requires the child to view three shapes before identifying

which one the odd one is. The child then recalls the location of each odd-one-out shape presented in the trial, in order. For the first trial the child is given one batch of three shapes, in accordance with the rest of the AWMA program, once the child had successfully completed four of six trials, an extra batch was added. The reliability coefficient of the odd-one-out task is stated as being .82 (Alloway, 2008).

From the above eight experimental measures, composite scores were calculated to represent IQ, WM and VWM and VSWM in the modelling of the results. Also included was age (maturational factors previously having been flagged as a core determinant of analogical ability).

2.2.3 Procedure. This study was divided up into two phases lasting (on average) 30 minutes per participant. In Phase 1, the first half of the RPA (consisting of R1ND and R2D) was administered, followed by the IQ subtasks. In Phase 2 (administered with a gap of at least 24 hours in between sessions) the AMWA subtests were given first, followed by the second half of the RPA (consisting of the R1D and R2ND conditions). All tasks were administered on a one-to-one basis in a quiet location. As the RPA follows a repeated measures design, with 4 variations (conditions) for each question, both sections of the RPA were presented in a standardized quasi-randomised order using a web-based research programme (Urbaniak & Plous; 1997). Here, any questions of the same condition appearing sequentially were re-randomised so that they were always at least two questions apart.

153

Table 10. Varimax rotated factor
loadings for WM and IQ measures.

	Component	
	I	II
	IQ	WMC
Block-Design	**0.824***	-0.108
Vocabulary	**0.789***	-0.379
Reading	**0.769***	-0.278
Coding	**0.678***	-0.515
Listening-Recall	**0.578***	**0.476***
Mr.X	**0.557***	**0.338***
Odd-One-Out	**0.525***	**0.568***
Backwards-Digit	**0.332***	**0.567***

* = Component loading above the .30

level

Table 11. Correlation matrix for individual difference measures

	Listening Recall	Backwards Digit	Odd One Out	Mr..X	Block Design	Vocab	Coding	Reading
Age	-.153	-.157	-.085	.022	.588**	.607**	.847**	.696**
Listening		.410**	.394**	.358**	.303**	.237*	.024	.182
Backwards			.347**	.255*	.154	.023	.045	.051
Odd One				.430**	.249*	.108	.064	.104
Mr.X					.356**	.214	.13	.145
Block						.633**	.652**	.666**
Vocab							.631**	.634**
Coding								.769**

* = *p<0.05*

**= *p<0.001*

2.2.4 Results.

Factor analysis. Before any analysis was conducted, a factor analysis, followed by post-hoc correlational testing was performed on the individual difference measures used in order to observe any multicolinearity effects that may be evident.

The Principal Axis Factor (PAF) used a Varimax (orthogonal) rotation of all 8 individual difference measures was conducted (Table 10 shows the varimax rotated loadings of each measure). An examination of the Kaiser-Meyer Olkin measure of sampling adequacy suggested that the sample was factorable (KMO =.765). Prior to analysis, an arbitrary criterion figure for

deciding how strong a loading must be in order for it to be included in component models was set at 0.3.

The PAF showed a predicted degree of separation, with two variables (assumed to be WM and IQ) being shown to account for the 8 measures (λ >1). Because factor I was shown to contain a mixture of WMC and IQ components, it was named "IQ". Because factor II continued purely WM components, it was named "WMC". Table 11 shows the post-hoc correlational matrix.

In summary, each WM measure was found to be correlated with one another at the <.001 level of significance (with the exception of Mr.X/Backwards digit, which was significant at the $p<0.05$ level). Each IQ measure was also correlated with one another at the same level (<.001).

However, the listening-recall task was correlated with both the block-design ($r=.303$, $p<0.001$) and Vocab tasks ($r=.237$, $p<0.05$); whilst the odd-one-out was correlated with the block-design task ($r=.249$, $p<0.05$). The Mr.X was also correlated with the Block design ($r=.356$, $p<0.001$).

Mean scores. Table 12 shows the mean scores for R^s/F^s respectively, by RPA condition and by year group. Table 13 shows the mean scores for IQ and WM measures by year group, whilst Table 14 shows the mean scores categorized into experimental factors.

Table 12. Mean relational and featural responses (by RPA condition and year group)

		All RPA Conditions			R1ND			R1D			R2ND			R2D		
		Mean	SD	%	Mean	SD	%	Mean	SD	%	Mean	SD	%	Mean	SD	%
Relational	All Ages	58.39	11.08	72.99%	16.46	2.86	82.30%	14.08	3.87	70.40%	14.51	2.89	73%	13.35	3.69	66.75%
	Year group 1	48.62	9.76	60.78%	14.81	2.58	74.05%	11.24	3.73	56.20%	11.81	2.68	59%	10.76	3.35	53.80%
	Year group 2	59.11	8.54	73.89%	16.19	2.75	80.95%	14.48	3.58	72.40%	14.89	2.06	74%	13.56	2.88	67.80%
	Year group 3	64.39	9.39	80.49%	17.81	2.52	89.05%	15.65	3.18	78.25%	16	2.39	80%	14.94	3.67	74.70%
Featural	All Ages	6.61	5.43	16.53%*	N/A	N/A	N/A	3.76	3.22	18.80%	N/A	N/A	N/A	2.84	2.76	14.20%
	Year group 1	9.33	6.29	23.38%*	N/A	N/A	N/A	5.19	3.52	25.95%	N/A	N/A	N/A	4.14	2.96	20.70%
	Year group 2	6.15	3.99	15.38%*	N/A	N/A	N/A	3.78	3.2	18.90%	N/A	N/A	N/A	2.37	1.62	11.85%
	Year group 3	5.16	5.39	12.9%*	N/A	N/A	N/A	2.77	2.72	13.85%	N/A	N/A	N/A	2.35	3.17	11.75%

* = Adjusted for the 40 trials in which featural objects occur

Table 13. Mean individual difference scores

Year group	Listening Recall (Verbal simple)		Digit Recall (Verbal WMC)		Odd One Out (Visual simple)		Mr X (Visual complex)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	96.77	15.86	102.08	16.77	102.23	15.63	98.97	17.22
1	87.43	14.48	101.43	15.971	103.38	17.99	98.38	16.7
2	100.19	15.52	109.52	15.57	104.11	15.38	97.56	17.83
3	93.35	16.82	96.03	16.23	99.81	14.3	100.61	17.45

	Vocab		Reading		Coding		Block Design	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
All	20.16	7.19	27.52	13.93	83.53	22.6	27.04	12.15
1	13.67	5.05	11.52	10.85	55.52	19.89	14.14	7.92
2	20.07	5.15	31.48	11.01	85.48	12.77	30.67	8.43
3	24.65	6.65	34.90	8.47	100.81	7.85	32.61	10.94

Table 14. Mean individual difference scores (categorized)

	VSWM Mean	SD	VWM Mean	SD	WM Mean	SD	IQ Mean	SD
All Age groups	99.42	13.7	100.6	13.89	100.01	11.84	158.25	49.13
1	99.43	13.05	100.88	14.62	100.15	11.35	94.86	36.84
2	104.85	12.77	100.83	13.99	102.84	11.9	167.7	27.99
3	94.69	13.55	100.21	13.76	97.45	11.92	192.97	23.53

Multiple regression analysis. The factor analysis had revealed two variables that could be contributing to the R^s, "IQ" consisting of all IQ and WM measures, and "WM", which contained just the WM measures. It was therefore necessary to investigate the predictive strengths of these components. As a result a Multiple Linear regression analysis was performed with the aim of constructing a global model of the data which offered the strongest explanation of the data from the variables which contributed to the greatest proportion of variance.

IQ model 1. The first variable tested was IQ. The factor analysis had already indicated that this variable contained the components from both WM and IQ, so two models were compared, one which contained just IQ components (model 1) and one which dealt with both (model 2).

For model 1 ($p<0.001$) the association between the R^s and IQ measures was moderately strong (Multiple $R = .53$), Vocabulary, Block-Design, Coding and Reading tasks accounting for 50% of the variance. The regression

coefficient for block design was .20 (95% CI = -.012 to .414); for

vocabulary it was .30 (95% CI = -.050 to .644); for coding it was .14 (95%

CI = .005 to .335) and for reading it was .12 (95% CI = -.093 to .335).

IQ model 2. For model 2 ($p<0.001$), using WM and IQ components, the

association between the R^s and IQ measures was increased (Multiple R =

.60), now accounting for 55% of the variance. The regression coefficient

for the listening recall measure was -.06 (95% CI = -.187 to .078), for the

backwards-digit it was .003 (95% CI = -.111 to .117), for the odd-one-out it

was -.04 (95% CI = -.164 to .090), for Mr.X it was .19 (95% CI = .078 to

.309), for the block-design it was .11 (95% CI = -.106 to .334), for the

vocabulary it was .29 (95% CI = -.052 to .624), for coding it was .13 (95%

CI = -.003 to .264) and for reading it was .16 (95% CI = -.049 to .365)

WM model 1. ($p<0.001$) Although IQ model 2 gave us a good indication

that WM measures played a (in some cases) lesser role, the second WM

variable was now tested as means to compare the results. Here the

association between WM and R^s was weak (Multiple R = .162). All WM

measures accounting for 12% of the variation (adjusted r^2).

The regression coefficient for the Listening recall task was .006 (95% CI = -

.167 to .180), for the Backwards-digit it was -.013 (95% CI = -.171 to .144),

for the Odd-one-out it was -.028 (95% CI = -.205 to .149) for Mr.X it was

.27 (95% CI = .114 to .425).

IQ model 3. (*p*<0.001) The above components were then reduced to see if a better fit for the data could be attained for both models, this meant removing all components not contributing to the model, specifically the listening-recall, odd-one-out and backwards digit measures, leaving the Mr.X measure as the only measure of WM.

For this model the association between IQ and R^s was again moderate (Multiple R = .59), but accounted for 56% of the variation (adjusted r^2). The regression coefficient for the Mr.X measure was .17 (95% CI = .064 to .27).

IQ model 4. (*p*<0.001) The model which best accounted for relational responding was therefore IQ models 2 and 3. Since the partial correlation analysis has already shown that age may interact with the measures used, and since both of the central theories of AR (Gentner, 1983; Goswami, 1992) has already indicated maturational factors may be involved. The IQ model 2 (thought to be the more comprehensive of the two) was re-run to include age.

For this model the association between IQ and R^s was again moderate strong (Multiple R = .60), accounting for 56% of the variation (adjusted r^2). The regression coefficients for the components were also improved, for the Mr.X measure it was .18 (95% CI = .073 to .284) for block design it was .08 (95%

Table 15. Mean relational responses (by RPA condition)

	R1ND	R1D	R2ND	R2D
Mean	16.46	14.08	14.51	13.35
SD	2.855	3.862	2.891	3.69

CI = -.131 to .294) for vocabulary it was .24 (95% CI = -.097 to .568), for

coding it was .09 (95% CI = -.072 to .256) for reading it was .14 (95% CI =

-060 to .344) and for age it was .95(95% CI = -.731 to .2.629)

ANCOVA analysis. One of the aims of this chapter was to explore the roles

of various cognitive systems in the response-selection processes under

conditions of complexity and distraction; it being anticipated that any

interactions observed involving WM and complexity/distractions could be

interpreted as being a mediatory capacity effect (High WMC individuals

being better than their lower counterparts at responding relationally in these

conditions).

For this reason ANCOVAs with two experimental repeated measures

factors– complexity and distraction, and the non-experimental (between

subjects) factor of year group were conducted with the covariate measure of

'Mr.X task- which, out of all WM tasks, had been found to account for the

largest proportion of variance within R^s in the regression analysis. In order

to compare this to non-capacity centric abilities, 'IQ' was also introduced as

a separate covariate. Table 15 shows the mean scores in each RPA condition

for R^s. In order to account for the possible change in the main effect caused by the covariate, mean centring of the covariate took place according to the recommendations made by Delaney and Maxwell (1988).

IQ. The first of these ANCOVAS was conducted with the experimental factor of IQ. Due to the number of individual IQ tasks contributing similarly towards the R^s - IQ is included here as a summary only, IQ representing all non-capacity based task used in Experiment 2.

The ANCOVA using IQ as a covariate of the proportion of relational responses given revealed main effects of IQ ($F_{[1,77]} = 84.87$; $p<0.001$, $\eta_p^2 =.52$); complexity ($F_{[1,77]} = 43.30$; $p<0.001$, $\eta_p^2 =.36$) and distraction ($F_{[1,77]} = 68.33$; $p<0.001$, $\eta_p^2 =.47$); and interactions between distraction x IQ ($F_{[1,77]} = 5.94$; $p<0.05$, $\eta_p^2 =.07$) and complexity x distraction ($F_{[1,77]} = 4.08$; $p<0.05$, $\eta_p^2 =.50$). But no interactions were found between complexity x IQ ($F_{[1,77]} = 0.41$; $p>0.05$, $\eta_p^2 =.005$) or complexity x distraction x IQ ($F_{[1,77]} = 1.22$; $p>0.05$, $\eta_p^2 =.01$)

Mr.X task. The ANCOVA using the Mr.X task as a covariate observed main effects of Mr.X ($F_{[1,77]} = 14.63$; $p<0.001$, $\eta_p^2 =.16$), distraction ($F_{[1,77]} = 67.23$; $p<0.001$, $\eta_p^2 =.47$), complexity ($F_{[1,77]} = 44.02$; $p<0.001$, $\eta_p^2 =.37$) and an interaction between complexity x distraction ($F_{[1,77]} = 4.12$; $p<0.05$, $\eta_p^2 =.05$), as well as an interaction between distraction x Mr.X ($F_{[1,77]} = 3.90$; $p=0.052$,
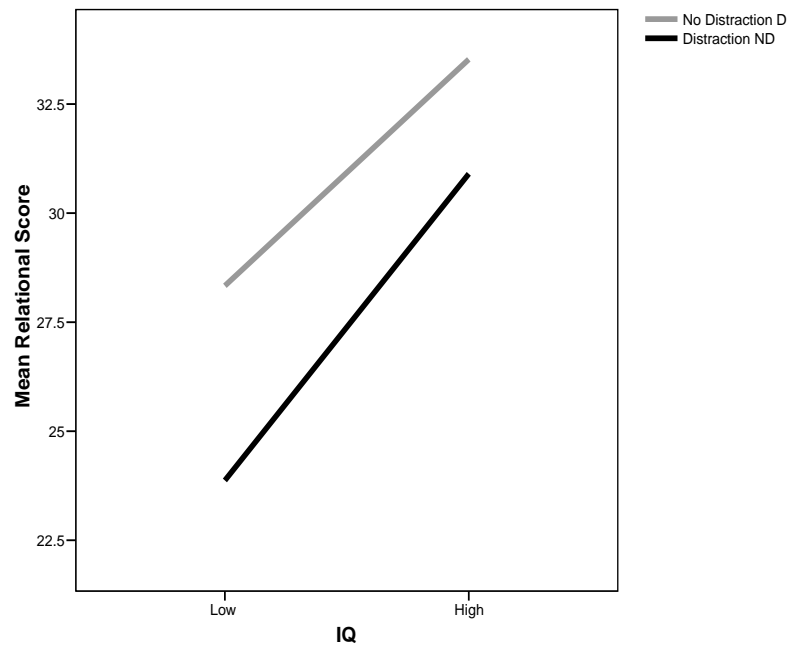
Figure 12. Interaction between IQ and distraction. The median split for

High/Low performers in IQ was introduced at 158.25
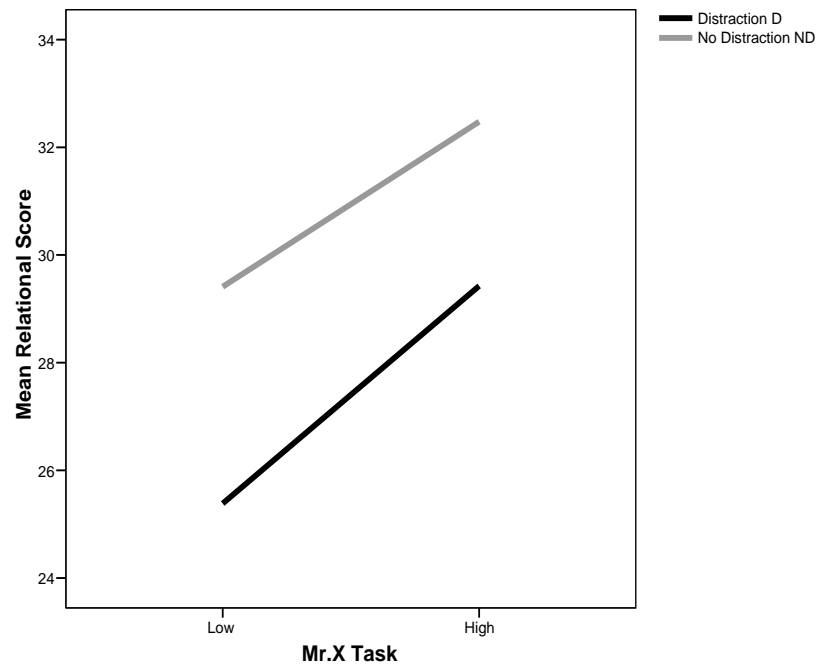
Figure 13. Interaction between Mr.X and distraction. The median split for

High/Low performers in Mr.X was introduced at 98.

Table 16. Correlations between featural responding and individual difference measures (by RPA condition)

	Listening Recall	Backwards Digit	Odd One Out	Mr.X	Block Design	Vocab	Coding	Reading	IQ	VWM	VSWM	WMC	Age
All	-140	-0.19	-0.17	-.375**	-.457**	-.486**	-.387**	-.356**	-0.463**	-0.093	-0.223*	-0.184	-0.333**
R1D	-0.085	0.026	-0.025	-.431**	-.456**	-.466**	-.348**	-.372**	-0.446**	-0.033	-0.281*	-0.184	-0.329**
R2D	-0.169	-0.066	-0.65	-.235*	-.367**	-.413**	-.360**	-.267*	-0.392**	-0.138	-0.109	-0.144	-0.28*

* = significant at the <.05 level

** = significant at the <.001 level

$\eta_p^2 = .05$) that was approaching significance. No interactions were found between complexity x Mr.X ($F_{[1,77]} = 1.69$; $p > 0.05$, $\eta_p^2 = .02$); complexity x distraction x Mr.X ($F_{[1,77]} = 1.79$; $p > 0.05$, $\eta_p^2 = .02$).

Analysis of the featural score. Because this thesis is concerned with successful analogical performance, we are less concerned with non-analogical forms of response. Never-the-less, Waltz et al. (2000) have shown that featural responses may be useful to the analogical researcher as low WMC may be reflected in a greater proportion of attribute based responses. As a result the following limited assessments were carried out.

T-test. A paired samples t-test revealed that significantly more featural results were obtained in condition R1D then R2D (t(78)= 3.258, $p < 0.05$, η^2 =.12), mean R2D F^s = 2.84, SD = 2.757, mean R1D F^s = 3.76, SD = 3.219.

Correlational analysis. As was conducted with R^s, correlational analyses were run assessing the statistical relationships between the individual difference measures and F^s (see Table 16).

Multiple regression analysis.
IQ Model 1. $p < 0.001$. The association between the F^s and all the WM and IQ measures was below moderate strength (multiple $R = .38$). Together IQ and WM accounting for 31% of the variation in F^s (adjusted r^2). The regression coefficient was .001 the listening-recall (95% CI = -.080 to .081), .008 for the backwards digit (95% CI = -.062 to .077), .08 for the odd-one-

out (95% CI = .005 to .159), -.11 Mr.X (95% CI = -.181 to -.041), -080 for the block design (95% CI = -.2145 to .053), -.224 for the vocabulary (95% CI = -.430 to -.018), -.021 for the coding (95% CI = -.1.00 to .059) and .016 for the reading (95% CI = -.110 to .142).

WM model 1. p<0.05. The association between the F^s and all the WM measures was weak (multiple R = .19). WM accounting for just 14% of the variation in F^s (adjusted r^2). The regression coefficient was -.03 for the listening-recall (95% CI = -.115 to .053), .02 for the backwards digit (95% CI = -.059 to .094), .08 for the odd-one-out (95% CI = .006 to .165) and .14 for the Mr.X task (95% CI = -.219 to -.069).

IQ model 2. p<0.001. Each item not contributing to the F^s was sequentially removed from the analysis; this reduced model resulted in only two significant components being found to be associated with featural responding, age and the vocabulary measure. The relationship being the strongest yet for the featural score, but still weak (multiple R = .24), accounting for 22% of the variance in the F^s (adjusted r^2). The regression coefficient for the vocabulary score was -.33 (95% CI = -.520 to -.139) and for age it was -.23 (95% CI = -.936 to .474). The conclusion being that neither IQ nor WM could adequately account for featural responding.

2.2.5 Discussion. The results in this study have indicated that WMC and IQ are both important components for success in scene-based analogical reasoning problems. Whilst perhaps common-sense given the

well-established connections between WMC and other forms of reasoning, these findings offer additional degrees of insight into the involvement of WM in AR.

Overall, IQ tasks were consistently seen to be better associated with, as well as contributing more to the R^s than WM tasks, yet WM was also reported to be closely connected with the concept of IQ. For the most part the involvement of WM was restricted to the Mr.X task which was observed to be a powerful-contributor to analogical success and the *best* contributor when age was controlled for; providing the possibility that WM might not be the maturational factor implied by Richland et al. (2004, 2006, 2010). Critically, no interaction was observed between any WM task and conditions of complexity/distraction.

This latter point was an important omission. Whilst the fact that an interaction between Mr.X and distraction was approaching significance was encouraging, if support were to be provided for the complexity-constraint theory (Andrews & Halford, 2002; Halford, 1992, 1993, 1998; Halford et al., 2002; Hummel & Holyoak, 1997) then it was a necessity that WMC be concretely related to increased complexity. This was not established; however concerns had been raised in the analysis that the experiment lacked sufficient power to successfully analyse three-way interactions, so these observations were viewed with caution.

Maturation. One interpretation of lack of maturational effects similar to those described by Richland et al (2004; 2006; 2010) are the floor/ceiling effects discussed in Experiment 1, e.g. the year groups used. Individual differences in WMC between year groups were not sufficient for significant effects to be observed. This perspective is supported by Halford (1992, 1993, 1998) and Morrison et al (2010) who suggest that by age 5 children have enough WMC to handle the complexity of ternary relationships (i.e. those used in the RPA, such as "being chased x being in the middle x chasing") leading to a less observable effect in younger children when the complexity level is raised.   Yet this explanation swings both ways, and given the findings presented here (that IQ may mediate AR) it could be argued that the only reason such differences were detected in the original 2004, 2006 study was because of the poor performance of the youngest year group- which tells us little other than the fact that young children aged 3-4 are poor at the RPA when compared to 14 year olds, and which could equally be attributed to factors involving IQ as well as WMC.

A further issue regarding the involvement of WM in Experiment 2 was that WMC is a domain general resource. If higher WMC is related to better analogical performance, then an equal cross-domain effect should have been observed in either VWM tasks (sentence-recall or backwards-digit task)? It could be argued that Mr.X, like the RPA, is a visual paradigm and therefore involve visual processing. However, in this case a relationship between the odd-one-out task (the other VSWM measure) and AR should also have been reported.

The solution to this dilemma is that the Mr.X task may be more sensitive to the demands placed on WMC then the other WM measures, but this raises the issue that Mr.X may be also be tapping specific processes necessary for completing the RPA; processes which could conceivably be more in-line with IQ than WM (see "IQ", below). Further research, determining what this Mr.X effect might relate to, is indicated.

IQ. Although Experiment 2 has been noted as lacking power, it should be observed that no such limitations seemed present in the IQ measure; IQ being established as interacting with both distraction and, to a lesser degree complexity.   This presents the important notion that functions of IQ may be able to describe the data purported to relate to the complexity-constraint hypothesis both in Experiment 1 and the original Richland et al. (2004, 2006) study: possibly through the identification and selection of 'correct' responses without any demands being placed on WMC by object complexity (Goswami, 1992).

As mentioned above, it is possible that what the IQ tasks are measuring may be functions which might be shared by highly-demanding tasks such as the Mr.X. This may be storage and/or processing capacity, in which case a further experiment (Experiment 3) detailing the relationship both may have to the RPA is a prerequisite for further assessment.   Cognitive abilities already well known to be mediate various aspects of thinking may also be involved, such as processing speed, increased executive functionality and cognitive efficiency (Kail & Salthouse, 1994; Lezak et al, 2004; Salthouse,

1996)[40]. Such processes potentially arbitrate IQ as well as WM, explaining the data from the regression modelling and factor analysis (which suggested that aspects of the IQ measures may contain components associated with WMC). It is hoped that Experiment 3 will be able to resolve this issue.

Featural discussion. The purpose of the featural analysis was to see if Waltz et al.'s (2000) predictions regarding WMC and featural responses were true. Overall the analysis appeared to show a similar pattern of findings to the relational analysis in that higher IQ children performed better at the RPA, in this case by choosing less featural responses. Again the mediatory effect of WM was limited entirely to the Mr.X, but was surprising in that the direction of the effect was the reverse of what was predicted. The effect being similar to what was observed in Experiment 1 in that the F^s was greater in complex conditions than simple- this being counter-intuitive to theories which suggest that increased complexity requires greater amounts of processing power resulting in a larger proportion of featural responses in these conditions (Hummel & Holyoak, 1997; Waltz et al., 2000).

Therefore, it had to be considered whether featural responses either might not be as distracting within the RPA as previously suspected, and/or that condition R1D was placing additional demands on WM that were different

---

[40] Such faculties have already been suggested as being intrinsic to AR through connectivist models of AR like LISA (Hummel & Holyoak, 1997). Faster and more effective processing being hypothesized to result in a greater level of inferential abstraction where more descriptive axes may be processed in unison (Salthouse, 1991).

to, or in excess of, those experienced in R2D; making R1D more prone to this type of mistake.

Evidence for the former is starting to emerge through eye tracking experiments. Whilst younger children are already known to spend more time on distracters then pertinent items in analogical problems (Thibaut, French, Missault, Gérard & Glady, in press) - in the RPA paradigm it has been reported that participants actively search the pictorial target scene for meaningful similarities having first spent time ascertaining the relation in the base (Gordon & Moser, 2007). Previous studies have already shown that fixation is highly correlated with meaningfulness of regions within a scene (Henderson, Weeks, & Hollingworth, 1999), meaning that complex relations in the RPA may in fact attract attention more than distracters.

Another way in which these findings might be explained is through what may been termed here as 'signposting.' This theorizes that relational forms of response (relational or relational error) are more likely to occur in conditions R^2 than R^1 not because increased complexity makes the problem harder, but because extra relational objects offer meta-level indicators that this form of response is preferable, making it easier. This may be explained thus: In condition R^2 the likelihood that relational forms of response will be selected is larger than others, as 3/5 of the objects in condition R^2 are relational in nature compared to 2/5 in R^2. Although in the RPA the 5 objects per scene ratio does not change, the presence of two relational objects, which usually are visually predominant either in terms of size or

positioning in the centre of the scene, may highlight that this form of response is the one required, or at the very least act as a memory cue to the RPA training session[41].

Just like the hypothesis presented previously, eye tracking studies also support the signpost effect. It is already known that larger objects are more likely to fix attention, and are easier to locate at high speeds (Djamasbi, Siegel, & Tullis, 2011), however recent research has also shown that scenes with low object to space ratios are generally less preferred whilst clusters of objects attract fixation, particularly if the objects in question are interesting or unusual (Henderson, & Ferreira, 2004), as in the case of the RPA.

A third explanation is that by offering a more direct (essentially forced) choice between featural and relational objects, more demands are placed on a discriminatory or inhibitory system in R^1 conditions then R^2. This may cause a problem for the reasoner if they are holding the base object in mind (the base object identical to or similar to the distracter object) as it could require inhibitory skill to dismiss that form of response, resulting in more featural objects being chosen in R^1 than R^2 conditions.

All of the above may be considered fair interpretations of the data yet one major concern which arose during the testing of the F^s was that of power- it

---

[41] Think of the signpost theory as turning the RPA into multiple choice questions rather then a measure of AR. In such cases having three forms of response of a certain type makes you consider that an element contained by each of the three may be the required answer.

being consistently shown that a large number of participants were required to illicit meaningful data. These circumstances were almost certainly due to the low proportion of responses, which in most RPA conditions was less than three. Given the higher F^s in Experiment 1, it was suspected that continued use of an experimental population size similar to the one used here might result in a measure with too high a degree of inter-experiment variability (i.e. intrusive random or artifactual factors have a greater effect on small scores).

A second issue regarding variance within the F^s was more serious. Experiment 1 had already indicated that there was not enough variance between year groups in the F^s for a maturational factor to be identified. Given that such a change had been predicted by the much more extreme (broader) age ranges of Richland et al. (2004, 2006), and difficulties had already risen in detecting interactions in the larger R^s. The small effect size of the F^s made it much more difficult to elicit meaningful data, thus questioning the efficacy of using featural data with such a narrow age range.

2.2.6 Conclusion. Although WMC, and not a STS, has been indicated as possibly constraining AR, evidence has been provided in Experiment 2 to suggest that IQ may account better for relational responding than WM, especially in conditions of distraction, although neither IQ or Mr.X are likely to entirely explain this condition and Mr.X alone is unlikely to entirely explain complexity. Additional questions

174

remain as to whether the current methodology may be sufficient to establish

a similar relationship with WMC.

As discussed previously, Halford's constraint hypothesis (Halford, 1992;

1993; 1998; Halford et al, 1994, 2004) predicts that by the age of 5 years

children may have sufficient WMC to process the RPA's 'complex' ternary

relations. Yet such an explanation does not clarify the complexity effects

demonstrated in relational responding within the year groups in Experiment

1.

Given that that it remains to be seen whether bound information such as

"chased and chasing" represents an increased cognitive load over "chasing",

as predicted by Halford (1992, 1993, 1998), it instead may be considered

that processing differences beyond relational representation may resolve the

dilemma illustrated by the presence of extra candidate objects. As reported

in chapter 1, this process is traditionally described in the field of AR as

mapping (Gentner, 1983), where the relevance of each argument is

compared for application with the task goals.

One such approach which takes this into account is relational primacy

(Goswami, 1992). As stated in chapter 1, Goswami's theory does not rely

upon the assumption that increased relational-complexity also represents

increased WMC load[42]. Instead, in situations where the similarity-

---

[42] Beyond what is required in the processing and maintenance requirements of relational
objects regardless of complexity.

constraint is known, it is the ambiguity (relevance) of candidate arguments which dictates performance.

Accordingly, increased complexity in the RPA increases problem ambiguity by providing additional objects for consideration. If the correct similarity cannot be identified or a decision is made without considering the appropriate compliment of competing responses, then errors could potentially be made that replicate complexity data, leading to reduced relational performance in complex conditions

In such ambiguous circumstances, performance factors- what Goswami calls the facilitation-gradient (Goswami, 1992) - mediate task performance, specifically the ability to form and abstract schemas appropriate for the resolution of the problem. Although how these schemas are formed is not described in detail, such processes are of course roughly analogous to fluid intelligence and the executive decision making faculties of the CE (which will be investigated further in Experiment 3).

A second relevant theory is that of Richland (Richland et al., 2004, 2006, 2010) and Morrison (Morrison et al. 2010), this too prescribes executive processing as underlying situations where the similarity-constraint is known but its application is not immediately obvious, and where a decision must be made from competing stimuli. The main difference between such theories being that mapping in relational primacy is optional and that for Richland et al. (2004, 2006) relational-complexity may also increase the need for

176

additional processing resources even if the problem is adequately represented in WM.

It is important to note that neither theory refutes the claim that WMC is involved in AR. The executive control aspects of the CE have long been thought as being synonymous with/representative of WMC (Baddeley, 2007), so it is reasonable to assume that the hypothesized constraint-effect of WMC in Experiment 2 may also represent the basic processing demands of a one-to-one comparison process or a mental workspace approach where more WMC allows more efficient manipulation of the problem (which can be illustrated with or without arguments of complexity).

It should therefore be considered a direction for future research in chapter 3 that the WMC effect demonstrated in chapter 3 should be better explained, and that the CE's ability to select appropriate information from that held in mind could potentially explain the reported mediatory effect of WMC in relational performance in the RPA. In Experiment 2 the Mr.X task was reported to be a predictor of AR, but was considered to be composed of attentional (WMC) and storage components (STS). Experiment 3 will develop our picture of the role of WM in AR from this assumption, looking at simple/complex span and verbal/nonverbal WM tasks (Experiment 6) in order to deconstruct the contribution of reportedly 'execute-heavy' tasks such as Mr.X.

2.3 Experiment 3

The previous experiment presented the possibility that some processes in WM may be mediating analogical performance in the RPA. However, it remained ambiguous as to what the effect might represent. Experiment 3 was therefore designed to assess the involvement of STS and WMC in AR, to deconstruct the relationship observed between WM and AR in Experiment 2.

The term 'memory' is often used to refer to stored information that is required for concurrent recall, such as retaining a telephone number in conscious thought. Information such as this does not necessarily depend upon processing ability, but may include such aspects if we are required to protect the memorized information from competing (non-relevant) 'intrusive' data, or perform a mental operation on the information being stored and/or retrieved (such as dividing a number).

The WM model (Baddeley, 2000) is modular, containing both processing and storage components. The term 'STS', or *storage capacity* is primarily used to define information stored for later processing (Halford, 1998). 'WMC' or *processing capacity* is used to describe computational *and* maintenance aspects of STM; i.e. "information that is currently entering into some kind of reasoning, decision-making, or other computational process." (Halford, 1998, p. 142). According to Alloway and her colleagues (Alloway, 2009; Alloway et al., 2008) WMC may be further conceptualized as a domain *general* resource which is applied to both visuo-spatial and

verbal domains, whilst STS's are domain *specific,* and applied to their prospective domains only.

As AR is believed to involve active reasoning processes requiring the manipulation of information for meaning, WM measures are assumed to predict analogical success better than those measuring STS. A factor exemplified by the fact that the RPA has both target and base scenes presented simultaneously, reducing visual STS loading by allowing the participant the ability to refer back to the base at any time rather than remember the query object.

WMC and reasoning are well known to be intercorrelated (Baddeley, 2007; Barrouillet, 1996; Bungel, Wendelken, Badre, & Wagner, 2004; Cho et al, 2007; Cho, Moody, Fernandino, Mumford, Poldrack, Cannon, Knowlton & Holyoak, 2010; Halford et al, 1994; Krawczyk, Hanten, Wilde, Li, Schnelle, Merkley, Vasquez, Cook, McClelland, Chapman & Levin, 2010a; ; Johnson-Laird & Byrne, 1991; Krawczyk, McClelland, Donovan, Tillman & Maguire, 2010b; Logie & Gilhooly, 1998). Children with greater WMC are predicted to perform better at the RPA because they can allocate more processing resources to a problem and avoid processing errors caused by incorrect assessments. In terms of AR, the most popular explanation for the involvement of WMC is the complexity-constraint hypothesis which relies on interactions between relational objects. However, it may also be true that scene-complexity (i.e. the overall number of relevant responses regardless of interactions, and the ability to recognize the correct response

from these), may make demands on the reasoning system. A third possibility raised in Experiment 2 is that one-to-one mapping naturally requires a baseline of computing power regardless of object complexity (i.e. the ability to compare two sets of information no matter how many descriptive axes they may possess). Increased WMC presumably still being beneficial to processing by allowing strategies to be better maintained against non-relevant interference, as well as assessing whether these are a good fit given the situation (Logie & Gilhooly, 1990).

Despite the presumption that WMC is behind successful analogical thought, knowing whether visual/verbal capacity plays a major or minor role in solving the RPA is critical to our understanding of AR. Even with the concurrent presentation of target and base scenes, a STS could be involved through the maintenance of the similarity-constraint or potential object candidates whilst attention is devoted to the target. When faced with ambiguous problems/conditions which may take longer to solve, or during prolonged testing, a greater STS span may also be advantageous (Gordon & Moser, 2007), keeping the demands of the task in mind as well as the analogy being used. A further factor to consider is that the RPA is a repeated measures paradigm using 4 conditions- similar scenes where the answer is always the same. Remembering previous responses and subsequently adapting behaviour to suit the current situation may therefore also be profitable good strategy.

For these reasons WM processes potentially benefit from greater
visual/verbal STS spans by providing resources for mental visualization of
the problem (i.e. a mental workspace), and the retention of verbal
instructions.   The current aim is not necessarily intended to draw a line
completely between STS/WMC; however understanding the contributions
of each is important in developing our concept of AR.

In Experiment 2, even though a methodological distinction was made
between VSWM and VWM it can be argued that the WM tasks used were
measures of WMC rather than STS (Alloway et al., 2008).   It was decided
that in order to best explore what the reported WM (i.e. Mr.X) effect
represented, an attempt should be made to deconstruct (as far as is possible)
the reported WM-AR relationship into STS and WMC aspects, assessing the
contribution of both, and whether either could explain the complexity-
effect[43] reported thus far in Experiments 1 and 2.   In a continuation of the
search for an interaction between WMC and complexity, the contribution of
both measures to the R^s would then be assessed within conditions of
complexity and distraction.

In order to reduce floor/ceiling effects, participants were recruited from a
single UK year group (Year 5) between Richland's original upper and lower
age limits. Although children aged between 10 and 11-years were more
likely to have sufficient WMC to represent and manipulate the similarities

---

[43] Increased R^s performance in simple, no distractor conditions, first reported in Richland et
al. (2004).

used in the RPA (Holyoak, 1992, 1993, 1998; Morrison et al., 2010; Richland et al., 2004, 2010)[44] , it was considered that the hypothesis that increased WMC associated with older year groups could still aid AR in other ways. Possibly through the ability to apply increased levels of computational power to the problem, WMC representing increased processing efficiency or the ability to attend between multiple strategies and therefore select more relevant solutions to the problem (Heinz Martin et al, 2002; Logie & Gilhooly, 1990).

2.3.1 Participants. Thirty children aged between 10 and 11 years (Year 5; mean age= 121.83 months) were recruited from three 'new' primary schools in accordance with the ethical criteria previously used in the earlier studies.

2.3.2 Materials and Design. In addition to a paper version of the RPA, two measures of *IQ* were used from the previous experiment, namely the block-design-task and the vocabulary-task (Wechsler, 1991). As a further measure of domain general knowledge and verbal fluency, the relational-task, was also included, as it was more relevant to the ability to recall associations.

---

[44] It has been argued that age based interactions are unlikely to be observed in previously tested ranges given that Children less than 6 years old may have already experienced a relational shift and/or possess the necessary WMC resources to process the ternary relations found in the RPA (Halford, 1993; Morrison, Doumas & Richland, 2006; 2011).

For the WM subscales four measures of WM were used: two measures of WMC and two measures of short-term memory capacity, (STS) each of which were visual and verbal tasks respectively. For the WMC tasks the Listening Recall task (Alloway 2007) and Mr.X task (Alloway, 2007) were again used for in order to try and replicate the earlier findings (see Experiment 2). For the measures of visual/verbal STS, two new tasks were introduced, both of which were intended to represent minimal executive loading: Digit recall (verbal; Alloway, 2007) and the Just Noticeable Difference Task (visual; Thompson, Hamilton, Gray, Quinn, Mackin, Young & Ferrier, 2007).

Just noticeable difference task (JND). The visual JND task is a computerized variation of the Thompson et al. (2006) task. Because it is assumed that all tasks designed as WM capacity measures are in some way 'contaminated' by executive processes (Phillips & Hamilton, 2001), the JND task was conceived to minimize this executive effect (even if it could not be removed entirely). This would be done by constructing a simple visual recall task with no reasoning demands, relying instead on a yes/no question relating to whether an object was the same size or not.

The JND consists of five trials, each containing 30 questions. For each question, a yellow base square appears for 1300ms to the top left of the mid-point (a jitter being included to make sure that sizes could not be judged by comparing the distance between squares). The base then vanishes and after a delay of 4000ms a target square appears, which remains until participants

respond. Participants answer using a colour coded button box, pressing the left button marked with the green word "same" if they think the target square is the same size as the base, and pressing the right button marked in red "different" if they think the target square is different.

The size of squares, as well as the order they appear in, are initially randomly generated, but then standardized for each participant. The difference between target and base squares are incrementally reduced by 10% for each level, so that in trial four (the second most difficult) the percentage difference is just 10%; whilst trial one (the first and easiest) differs by 40%. For trial 5, the difference is moved down to just 6%, five per cent or lower being considered too difficult for this age range.

After a practice trial with an automated teaching programme (which would correct the participant if they answered correctly or incorrectly) the children were instructed to answer as quickly as possible.   In order to make sure participants were answering properly, and to make sure that the results were not due to visual difficulties in focusing attention, half the questions (appearing alternatively) were controls within which the base and target appeared simultaneously. A cut-off was also introduced so that if the participant was answering at chance level (50%) the task would stop after 15 questions, the score being zero for that trial. The overall score was determined by the number of responses from all conditions. Due to time restrictions no reliability data was available, although similar paradigms had previously been successfully applied (Thompson et al., 2006).

Digit recall.   A measurement of verbal STS; the digit recall task is a subtask from the AWMA (Alloway, 2007).   Numbers are read out and the participant asked to repeat them. Six questions make up each trial, with the participant advancing only if they answer four or more questions correctly. For each extra trial, an additional digit is added for the participant to remember. Reliability for the digit recall task in the AWMA is .84 (Alloway, 2007).

Similarities task.   One of the most widely used measures of relational reasoning; the similarities subtask is taken from the WISC-IIIR (Wechsler, 1991) as a direct measurement of domain general knowledge and understanding of relations. Here 19 word pairs are given and the participant is asked to describe what makes them the same. Answers are divided into high level (2 points), low level (1 point) and inappropriate (0 points) responses which are marked according to strict criteria in the WISC manual[45].

 Reliability for the similarities task has been shown to be .81 across all year groups (Goldstein & Hersen, 2000).

---

[45] Although some leeway is given for experimenter interpretation, this is a classic example of the form of relational reasoning task where the participant must respond with an answer the experimenter has previously decided is "correct" and where the experimenter has decided that some forms of answer are cognitively "better" then others. It was chosen because it represented a commonly encountered form of relational paradigm.

2.3.3 Procedure. This experiment was divided up into 2 sessions lasting 35 minutes (on average) each. The first session consisted of RPA conditions R1ND and R2D, followed by the: Block Design, Vocabulary and Similarities subtasks (in that order). After a minimum of 24 hours the second session was conducted, consisting of RPA conditions R1D and R2ND, followed by the Mr.X, Listening Recall, and JND subtask.

All tasks were conducted on a one-to-one basis in a quiet location within the school, the order of presentation of the RPA problems being counterbalanced using the same procedure and web based programme as experiments one and two (Urbaniak & Plous, 1997): i.e. in a quasi-random order. Here problems from both conditions in a session were randomly determined, the order being manipulated when a problem was presented adjacent to an identical question within a separate condition (i.e. question 1, RND and question 1 R2D).

Table 17. Mean relational scores (by RPA condition).

		Mean score	SD	%
Relational Score	Total	64.3	5.22	80.375
	R1ND	18.4	1.25	92
	R1D	15.13	2.6	75.65
	R2ND	14.93	1.61	74.65
	R2D	15.87	2.36	79.35

Table 18. Mean individual difference scores.

Task	Mean score	SD
Similarities	17.6	4.62
Block Design	37.27	12.36
Vocab	26.47	8.2
Digit Span	94.2	3.37
Listening	91.1	22.28
Mr.X	94.97	23.25
JND (Span)	4	0.98

2.3.4 Results. Table 17 shows the mean relational scores from Experiment 3, whilst Table 18 shows the mean scores from each of the tasks.

Between condition analysis. t-tests conducted between conditions of the R^s reported that there was a significant difference between conditions R1ND and R1D (t (29)= 6.60, $p<0.001$), R1D and R2ND (t(29)=11.192, $p<0.001$, η^2 =.81), R1ND and R2D (t(29)= 6.517, $p<0.001$, η^2 =.59) and R2ND and R2D (t(29)= -2.177, $p<0.05$, η^2 =.14). No significant difference was found between conditions R1D and R2ND (t(29)= -0.388, $p>0.05$, η^2=.01) or conditions R1D and R2D (t(29)= -1.553, $p>0.05$, η^2 =.08), meaning that it was unlikely that effects of complexity could be observed.

Correlational analysis. Two phases of correlational analyses were carried out. The first looked at the relationships between task scores and the R^s (see Table 19). The second compared task scores and the R^s within individual RPA conditions (see Table 20).

The R^s was found to positively correlate with the Similarities (r = .417, $p<0.05$), Block Design (r = .363, $p<0.05$), Vocabulary (r = .422, $p<0.05$), and Mr.X (r = .392, $p<0.05$) tasks, whilst the Listening Recall Task was approaching significance (r = .355, $p=0.054$) tasks. The Digit span and JND measures were not found to be related to R^s ($p>0.05$).

For condition R1ND the Similarities (r = .507, $p<0.01$), and Vocabulary (r = .422, $p<0.05$) subtasks were found to be positively correlated with the proportion of relational responses in this condition.   For condition R2ND the Mr.X subtask was correlated with the proportion of relational responses (r = .399, $p<0.05$). For condition R2D the Similarities (r = .495, $p<0.01$), Block Design (r = .390, $p<0.05$) and Vocabulary task (r = .574, $p<0.01$) were all correlated with R^s.

Table 19.   Correlational matrix for IQ/WM measures and relational responding

	Similarities	Block Design	Vocab	Digit Span	Listening	Mr.X	JND
Relational	.417*	.363*	.422*	.320	.355†	.392*	-.102
Similarities		.469*	.469**	.729	.174	.311	.298
Block			.352†	.248	.349†	.498*	.310
Vocab				.297	.212	.324	-.201
Digit span					.786**	.652**	-.125
Listening						.623**	-.112
Mr.X							.119

†Correlation is approaching significance (between the 0.05 and 0.06 level)
** Correlation is significant at the 0.01 level
* Correlation is significant at the 0.05 level

Table 20. Correlations between relational responding and individual difference measures (by RPA condition)

	Similarities	Block	Vocab	Digit span	Listening	Mr.X	JND
R1ND	.507*	.310	.453*	.223	.312	.228	.084
R1D	-.010	.117	.097	.191	.189	.174	-.243
R2ND	.342	.244	.101	.319	.337	.399*	.065
R2D	.495*	.390*	.574**	.187	.217	.333	-.074

** Correlation is significant at the 0.01 level
* Correlation is significant at the 0.05 level

ANCOVA analysis. Repeated-measures analyses of covariance (ANCOVAs) were conducted using the same means centred methodology as before (Delaney & Maxwell, 1988). These were performed for all experimental measures found to be correlated with R^s (Similarities, Block Design, Vocabulary, Listening Recall and Mr.X). This was done in order to see if they mediated conditions of complexity and distraction.

Similarities task. For the similarities-task covariate analysis, significant main effects of Similarities task ($F_{[1,28]}= 5.92$; $p<0.05$, $\eta_p^2 =.18$) complexity ($F_{[1,28]}= 27.06$; $p<0.05$, $\eta_p^2 =.49$), and distraction ($F_{[1,28]}= 10.02$; $p<0.05$, $=0.26$) were found. Also, the interactions between complexity x Similarities ($F_{[1,28]}= 4.34$; $p<0.05$, $\eta_p^2 =.13$), complexity x distraction ($F_{[1,28]}= 61.76$; $p<0.001$, $\eta_p^2 =.69$) and complexity x distraction x Similarities ($F_{[1,28]}= 5.47$; $p<0.05$, $\eta_p^2 =.16$) were significant. No interaction between distraction x Similarities ($F_{[1,28]}= 0.004$; $p>0.05$, $\eta_p^2 =.000$) was reported.

Figure 14. Three-way interaction between complexity, distraction and the similarities-task. The median split for High/Low performers in the similarities task was introduced at 18.5.

Although R^1 conditions were easier than R^2, there was less reported difference between simple and complex conditions in participants who had a higher similarities score- meaning that high-similarities participants were better able to handle complexity effects.

In order to represent the three-way interaction, a median split was introduced after removing outliers, to illustrate high and low similarities-task scores (Figure 14).

An independent samples t-test with the grouping variable of High/Low similarities was then performed in order to gauge the strength of the

interaction and why it was taking place (i.e. the difference between RPA

conditions in high/low peforming similarities particiapamts may appear

great in the figure, but may in fact be non-significant, disguising why an

interaction was detected) .

A significant difference was observed between high/low similarities groups

in condition R1ND (t(28)= -2.803, $p<0.05$, η^2=.22), R2ND (t(28)= -3.112,

$p<0.005$, η^2 =.26); whilst the difference In the R2D Condition was

approaching significance (t(28)= -1.973, p=0.058, η^2=.12). No significant

difference was observed between low/high groups in condition R1D (t(28)=

0.718, $p>05$, η^2 =.02).

This analysis appears to show that participants with high similarities scores

perform better than people with low scores, particularly in the no-distracter

conditions (on both levels of complexity). Participants who perform better

at the similarities task eliciting a higher R^s in complex no distracter

conditions (R2ND); an effect which was also true to a certain degree of

complex conditions in general (the difference between low/high peformers n

condition R2D approaching significance).

Summary of the 3-way interaction: Participants in both similarities-task

groups peformed at the same level in distracter conditions, but generally

high similarity peformers found conditions of complexity easier when they
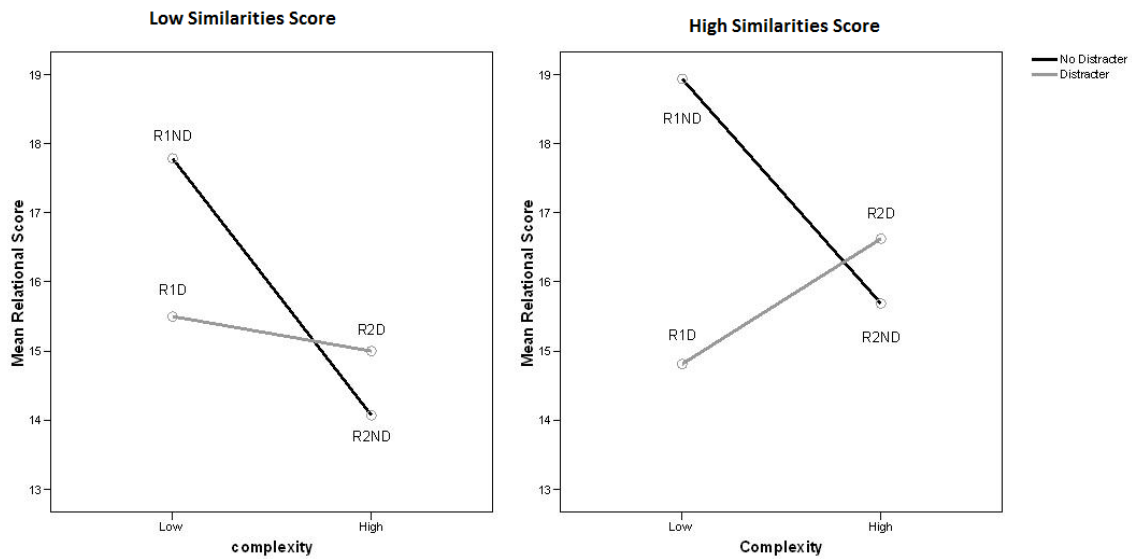
did not have a distractor present in them.
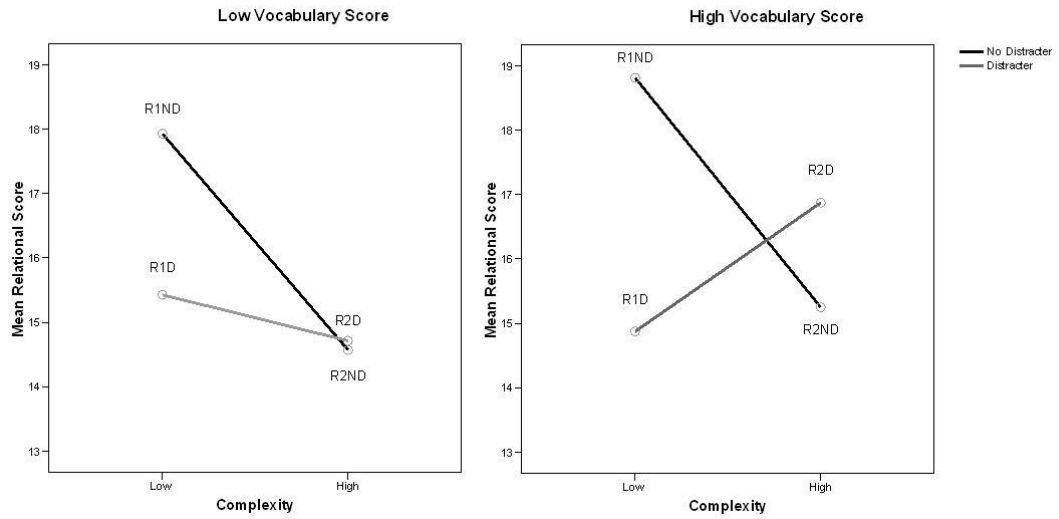
Figure 15. Three way interaction between complexity, distraction and the vocabulary task. The median split for High/Low performers in the similarities task was introduced at 18.5.

Vocabulary task.　For the Vocabulary-task covariate analysis, main effects of Vocab ($F_{[1,28]} = 0.021$; $p<0.05$, $\eta_p^2 = .18$), complexity ($F_{[1,28]} = 24.74$; $p<0.001$, $\eta_p^2 = .000$) and distraction ($F_{[1,28]} = 10.53$; $p<0.05$, $\eta_p^2 = .27$) were found. Interactions were also reported between complexity x distraction ($F_{[1,28]} = 66.96$; $p<0.001$, $\eta_p^2 = .71$) and complexity x distraction x Vocab ($F_{[1,28]} = 8.29$; $p<0.05$, $\eta_p^2 = .23$) but not between complexity x Vocab ($F_{[1,28]} = 1.57$; $p>0.05$, $\eta_p^2 = .05$) or distraction x Vocab ($F_{[1,28]} = 1.44$; $p>0.05$, $\eta_p^2 = .05$).

Once again, in order to interpret the three way interaction, a median split was introduced to the vocabulary score, separating high and low scoring participants (Figure 15).   An independent samples t-test with the grouping variable of High/Low vocabulary was then performed. This showed a

significant difference between high/low groups in condition R2D (t(28)= 2.777, *p*<0.05, η² =.22) whilst the difference between groups in condition R1ND was approaching significance (t(28)= -2.037, *p*=0.051, η² = 0.13). No significant difference was reported between groups R1D (t(28)= -0.576, *p*>0.05, η² =.01) or R2ND (t(28)= -1.153, *p*>0.05, η² =.05).

Summary of the 3-way interaction: The interaction centered on condition R2D, children with higher vocabulary scores eliciting more responses in this complex distracter condition then their low vocabulary counterparts.

Block-design task. For the Block-Design covariate analysis, main effects of complexity ($F_{[1,28]}$ = 24.46; *p*<0.001, η^2_p =.47), distraction ($F_{[1,28]}$ = 10.14; *p*<0.05, η^2_p =.27) were reported, whilst a main effect of Block design was approaching significance ($F_{[1,28]}$ = 4.18; *p*=0.051, η^2_p =.13). Interactions were reported between complexity x distraction ($F_{[1,28]}$ = 53.68; *p*<0.001, η^2_p =.66) but not complexity x Block design ($F_{[1,28]}$ = 1.24; *p*>0.05, η^2_p =.04), distraction x Block design ($F_{[1,28]}$ =0 .353; *p*>0.05, η^2_p =.01) or complexity x distraction x Block Design ($F_{[1,28]}$ = 1.09; *p*>0.05, η^2_p =.04).

Listening-recall task. For the Listening-recall covariate analysis, main effects of complexity ($F_{[1,28]}$ = 23.51; *p*<0.001, η^2_p =.46) and distraction ($F_{[1,28]}$ = 10.02; *p*<0.05, η^2_p =.26), as well as an interactions between complexity x distraction ($F_{[1,28]}$ = 51.76; *p*<0.001, η^2_p =.65). No main effect of Listening Recall ($F_{[1,28]}$ = 3.85; *p*>0.05, η^2_p =.12) was reported, nor were

any interactions between complexity x Listening Recall ($F_{[1,28]} = 0.097$; $p>0.05$, $\eta_p^2 =.003$) distraction x Listening Recall ($F_{[1,28]} = 0.008$; $p>0.05$, $\eta_p^2 =.000$) or complexity x distraction x Listening Recall ($F_{[1,28]} = 0.05$; $p>0.05$, $\eta_p^2 =.002$)

Mr.X task.   For the Mr.X covariate analysis, a main effect of Mr.X ($F_{[1,28]} = 4.99$; $p<0.05$, $\eta_p^2 =.15$), complexity ($F_{[1,28]} = 24.72$; $p<0.001$, $\eta_p^2 =.469$) and distraction ($F_{[1,28]} = 10.08$; $p<0.05$, $\eta_p^2 =.27$) was indentified, as well as an interaction between complexity x distraction ($F_{[1,28]} = 51.67$; $p<0.001$, $\eta_p^2 =.65$). No interactions between complexity x Mr.X ($F_{[1,28]} = 1.54$; $p>0.05$, $\eta_p^2 =.05$), distraction x Mr.X ($F_{[1,28]} = 0.17$; $p>0.05$, $\eta_p^2 =.006$) or complexity x distraction x Mr.X ($F_{[1,28]} = 0.002$; $p>0.05$, $\eta_p^2 =.000$).

2.3.5 Discussion.   Experiment 3 sought to assess the relationship between STS, WMC and AR. Whilst the involvement of STS remained inconclusive (no significant correlation being observed between this form of measurement and AR) the results of this study provided further evidence for the involvement of WMC in AR. Experiment 3 underlined the observation from Experiment 2 that Mr.X was an important contributor to the solution of the RPA whilst also suggesting that other WMC measures may be important in the successful resolution of the task (the correlation between the R^s and second WMC task, the listening recall task, was approaching significance).

Despite this, IQ measures were again shown to be better predictors of analogical success than WM, and whilst no WMC-complexity interactions

were reported, measures of IQ were once more observed to interact with factors of complexity and distraction.

It was suggested in Experiments 2 and 3 that the lack of WMC and complexity interaction data may be due to floor/ceiling-factors; that is to say participants were finding the task too easy for their year group, and that the level of variance in the R^s was perhaps too small to elucidate significant conclusions from when comparing covariates such as WM and IQ in conditions of complexity *and* distraction (even in a single year group such as Experiment 3). Yet whilst this is acknowledged[46], significant interactions were established between the R^s and IQ tasks - providing important evidence that such interactions could be observed, despite the low power and variance of the score.

This observation that IQ contributed more to the R^s than WMC presented a dilemma for the researcher. Either because a) these IQ measures were better indicators of the memory processes underlying the resolution of complex problems then the WMC tasks of the AWMA (unlikely). b) The same systems were facilitating an unrepresented mediatory effect of WM in complexity. Or c) IQ processes could account for the resolution of

---

[46] This was reflected in both the between-condition analysis, which showed no significant difference between conditions R1D and R2ND, or R1D and R2D (meaning that participants found conditions of distraction equally as difficult regardless of levels of complexity) and the power analyses which suggested that 30 participants for a 2x2 analysis might not be sufficient in the RPA given that in all three experiments an average of 80% of the responses in the year five year group were analogical (relational) in nature and therefore 'correct'. This high level of correct responses is highlighted by the work of Krawczyk et al (2010a) who showed that despite administering the RPA to traumatic-brain-injury patients with "severe" executive difficulties, aged between 3 and 17 years, performance in any condition was still up to 90% in some conditions.

complexity problems solely through non-capacity 'fluid intelligence' processes, i.e. the ability to choose an appropriate response regardless of relational-complexity

In order to answer what processing systems could be supporting AR; the tasks found to be significantly associated with AR were appraised. These were IQ and the Mr.X task

.

IQ measures. The three IQ tasks (block-design, similarities and vocabulary tasks) were chosen as measures of general problem solving ability that may include, but which did not rely upon, storage components of WM (i.e. fluid intelligence, see section below). It is notable however that unlike the block-design-task, both the vocabulary and similarities tasks (which were found to be mediate factors of complexity/distraction) require very little active manipulation of visual stimuli, both having being chosen for their known association with 'general' cognitive functioning (Prifitera et al, 2005; Flanagan & Kaufman, 2004, Sattler & Dumont 2004) as well as their representation of what Goswami (1992) terms 'domain general knowledge'.

The prospect of domain knowledge being represented through increased vocabulary and knowledge of relations in the relations task, was however unlikely: the high rate of participant success in Experiments 1-3 supporting Richland et al.'s (2004, 2006) original claim that the similarity-constraints used in the RPA are already represented in LTM by the age ranges being tested, and that the children understood the requirements of the task.

199

It was therefore considered feasible that what was being demonstrated by IQ measures in Experiments 2 and 3 was an efficiency quotient consisting of increased functionality (Prifitera et al., 2005) and processing speed (Kail & Salthouse, 1994; Salthouse, 1996).   This makes good etiological sense. Both the vocabulary and similarity tasks require strong associations between bound 'chunks' of semantic information, as well as the rapid access of task-relevant information stored in LTM (e.g. verbal fluency, see section below). Even the block-design task is a timed measure, requiring faster, more efficient executive functions (Brown, Brockmole, Gow & Deary, 2012), and where children who perform faster are rewarded with a greater score, regardless of whether they both complete the task.

On this understanding, the arguments for the involvement of the concepts of fluid intelligence and processing efficiency (described here as 'verbal fluency') in AR were laid-out. Although it was considered that both might in fact represent aspects of the same processes, clear divisions were never the less suggested.

Verbal fluency. "Verbal fluency" (see Salthouse 1996 for a review) has been described as the efficiency of neural networks and the speed in which cognitive operations associated with fluid-intelligence, planning and problem solving may be carried out (Fry & Hale, 1996).   Increased functionality is also beneficial in processes correlated with STS as faster, more efficient processing could reduce the amount of time a similarity-

constraint or other representation of relational problem needs to be held in mind. WMC may also potentially benefit through the number of cognitive manipulations (i.e. assessments) that may made within that reduced period, lowering the demands on capacity constrained processing resources.

Verbal fluency has already been associated with the resolution of visually based reasoning paradigms (Bryan, Luszcz & Crawford, 1997); facilitating task-specific processes such as retrieval, encoding and rehearsal, all of which are understood to aid the abstraction of meaning in relational problems through increased speed as well as the ease in which bound information is recalled (Carpenter et al, 1990).

This later aspect is of particular importance to the analogical researcher as theorists such as Holyoak (Cho et al., 2007; Gentner et al., 2000; Holyoak & Thagard, 1989; Hummel & Holyoak, 1997; Hummel & Holyoak 2003) have proposed a connectionist view of analogical thought where concepts may be represented through the strength of bindings between semantic nodes held in mind. Here stronger connections (indicated here as being measured by verbal fluency tasks) equate to easier and more readily availability associations attached to those concepts- leading to fewer incorrect responses in AR tasks and better overall conceptual clarity.

 Although an approach such as this relies on aspects of cognitive control (specifically the inhibition of non-relevant relationships) it remains consistent with the maturational factors observed in both Experiments 1 and

2, as well the Richland series of experiments. This is because, whilst

increased real-world experience and interaction may lead to stronger

bindings, speed of processing has also been shown to develop across

childhood (Salthouse 1996, Chuah & Maybery, 1999).

Fluid intelligence. "Fluid intelligence" is defined as the ability to understand

complex relationships and solve novel problems (Martinez, 2000).

Currently there is a great deal of debate as to whether fluid intelligence is a

separate entity from WM (see chapter 1) and processing efficiency or

whether it is independent.   In this instance it may be thought of as a

reasoning system which is not dominated by a storage component.   A

high fluid intelligence may be seen as being advantageous in AR as it would

allow us to select and dismiss the correct (most relevant) relational object

from a number of possible candidates, regardless of their levels of

complexity. In the RPA increased difficulty may conceivably be measured

by either the number of relevant objects[47] or the degree of similarity a

candidate object shares with another. Thus the difficulty is in recognizing

the correct response, fluid intelligence equating directly to the relational

primacy theory described by Goswami (1992).

Mr.X task.   Adapted from an earlier ("Mr.Blobby") measure by Hamilton

et al (2003), the Mr.X task is a visuo-spatial measure that was selected due

---

[47] This is not relational (object) complexity but scene complexity

to its greater loading of attentional resources compared to other STS measures.

As described in Experiment 2, in the Mr.X task children are required to store, and maintain information (in this case the location of a ball) whilst performing a secondary (not dual) task which impedes or at least diverts attention (by using up storage and processing components) from the primary. Children with an increased WMC are thought to be able to complete the task through greater processing and storage capacities by being able to apply more resources to the problem whilst under increasingly heavy load. The more problems a child is able to solve sequentially, and the more they can hold in mind whilst conducting secondary processing, the greater their Mr.X score.

Despite representing overall capacity, the Mr.X task also engages a number of important executive systems as defined by Baddeley (2007). For instance, the participant must inhibit non-relevant data, switch between two modes of thinking, and be able to divide attention- holding the location of a number of balls from a number of different problems whilst conducting extraneous tasks. As with the IQ measures, performance of Mr.X is also likely mediated by an efficiency/speed of processing component: the ability to recall the visual locations of the Mr.X 'ball' being determined not only by the ability to maintain visuo-spatial information against temporal decay, but also by the speed in which the participant recognizes and processes the primary and secondary tasks, and how quickly/flawlessly they are able to

switch between the two modes of thinking. Dual and secondary task studies have already shown the benefits of returning to a primary task at faster speeds before it has had time to decay (Miyake, Friedman, Rettinger, Shah, Arbor, & Hegarty, 2001).

It was therefore considered that the Mr.X measure could represent three essential systems which are crucial for the resolution of the RPA. Fluid intelligence and/or increased functionality executive processes, or WMC.

Since the latter could be considered to be the product of the second (and had already been shown to be associated with AR in the previous experiments) and the first had been dealt with in IQ, attention was paid to the possible advantages of seeing executive processes as skills advantageous to the resolution of the RPA.

Executive functions. Executive control is known to be associated with other forms of inductive reasoning (De Neys & Van Gelder, 2008; De Neys, Schaeken, & Ydewalle, 2005; Gilhooly, 1998; Gilhooly et al, 1993). Yet recent research has provided increasing evidence that it also has specific relevance to analogical reasoning, through neuroimaging (Crone et al, 2009; Cho et al, 2009; Krawczyk et al, 2010) and classical forms of experimentation (Iroise, Houlton, Kalina & Blakemore, 2010; Morrison et al 2010; Richland & McDonough, 2010; Thibaut et al., in press)

It is suggested that Baddeley's (2007) executive functions -being able to coordinate incoming information, switch between strategies, update old and inhibit new information- have obvious applications in tasks such as the RPA. Not only must a child choose relevant responses whilst holding others in mind, a core requirement of scene based analogies is for a child to be able to bind previously existing rules to novel objects whilst consistently switching between featural (base) and relational (target) selection strategies.

A fourth option has also been presented by Richland and colleagues (Richland et al., 2004, 2006; Morrison et al., 2010) who hypothesize that inhibitory control may play a key role in relational success at the RPA when two forms of response conflict and previously rule-sets (such as 'what is the same as') require suppression (Davidson et al., 2009; Diamond et al., 2002; Indre, Viskontas, Robert, Morrison, Holyoak, Hummel & Knowlton, 2004).

However, inhibition may also conceivably play other role in AR where prepotent information (such as the appearance of a base object) may require inhibition because it is either forefront after being recognized in the base (Oberauer, 2005) or because selective pruning of semantic bindings controls relevancy (Hummel & Holyoak 1997; 2007).

2.3.6 Conclusion. Given the small sample size, caution must be made in interpreting many of the observations reported here; however these findings still offer important insights into the processes underlying AR.

In summary, by demonstrating a relationship between AR and WMC, this study has supported theories which implicate capacity based restrictions on analogical thought, such as the complexity-constraint hypothesis (Halford, 1992, 1993, 1998). However, it is proposed that an open mind be kept when reviewing these findings.

Reduced relational responding in conditions of complexity (conditions R2ND and R2D) has previously been inferred both by Richland et al. (2004, 2006) and this thesis (Experiment 1) as representing an increased load on WM brought about by the presence of extra relational objects. A theoretical relationship existing between processing power used and the number of descriptive dimensions within an analogical argument (in the case of the RPA this would equate to the argument 'chasing' requiring less WM resources to successfully compute than 'chasing and being chased'). However perspectives other than Gentner and Halford's complexity-taxonomies may be applied to such findings.

Goswami's relational primacy theory (1992) is one such approach, relying on the formation of schemas and learning sets which encourage a child to respond in a specific way without it being automatic and more importantly: without a demand being placed on WMC by the presence of extra relations. In other words the manipulations performed by fluid intelligence and executive control of relational concepts may determine relational responding without the size of the concept being relevant.

'Relational-error' objects in this way may camouflage the goal of the task without changing the similarity-constraint itself by diverting attention away from the relational object and exposing weak schemas that are more readily corrupted by distractions, temporal decay, or otherwise open to interpretation. Whilst such ambiguity may be described by a complexity-constraint approach (the more similar relationships an object holds to the base object, the more likely it is to be the right answer) it may also be described by incorrect assumptions in the rule-set required to solve the problem i.e. the incorrect 'looking for objects involved in a chase' as opposed to the correct 'looking for an object in the middle'.

This could potentially elicit the same pattern of data as presented earlier (reduced relational responding in complex conditions), it being argued that extra relations increases the number of admissible options- not because they are more complex, but because the reasoners understanding of the similarity-constraint is poor.

Thus far in the thesis, the involvement of WM, and subsequently WMC has been indicated as arbitrating AR. However, as reported in Experiments 2-3, IQ was shown to be a better predictor of analogical success.

With such arguments firmly established, chapter 3 will investigate the nature of WMC's involvement in AR. It will focus on the questions as to how much of a load the RPA may be placing on the WM system (which will be investigated in this instance through reaction time and dual task

experiments) as well as assessing what role the executive functions

described by Baddeley (2007) may be playing in AR. The latter question

being postulated in an attempt to see if EFs can describe the relationship

between WM and AR reported in Experiments 2 and 3.

Chapter three

3.0 Introduction

Chapter 3 is intended to expand upon the findings of the previous chapter.
Its main objective is to examine the function that WMC, as indicated in
Experiments 2 and 3, may play in analogical thought in the RPA. Using a
Reaction Time (RT) paradigm, Experiment 4 will explore the possible
loading effects of relational and featural response formats in the RPA;
hopefully providing insight into Richland et al.'s (2004, 2006) suggestion
that featural inhibition, as a process of WMC, may be a central component
of successful relational responding in the RPA by suppressing irrelevant
forms of attributional response. Effects of complexity will also be
illustrated through RT's in this experiment in consideration of the
possibility that differences in response timings may affirm complexity
effects not visible in the proportion of relational responses.

Experiment 5 will use a dual-task paradigm to see if processes known to be
involved in WMC can be associated with performance in the RPA,
attempting to demonstrate visible capacity constraint effects by increasing
the loading of the task beyond its original format.

Experiment 6 will continue from Experiments 3 and 5. The intention being
to potentially imply a new role for WMC in AR through the involvement of
specific Executive Functions (EFs) in AR, the primary role of which is to
divide and focus attention (Baddeley, 1996b) as well as mediate information
from LTM (Baddeley, 2000) in WM.

3.1 Experiment 4

Chapter 2 showed how the presence of a featural distracter and extra relational objects increased task difficultly in the RPA, and how WMC may play a significant role in arbitrating analogical thought- although precisely why remained unknown. Previously in the thesis interpretation of the relationship between WM and AR has focused on the complexity-constraint effect perspective of analogical thought (Andrews & Halford, 2002; Halford, 1992, 1993, 1998; Halford et al., 2002), however until now, no interaction between WM and relational-complexity in the RPA has been evident. One possible explanation for this concerns executive control and the ability to suppress non-relevant material in the analogical scene (Richland et al., 2004, 2006, 2010).

Cognitive inhibition has been described as an executive function involved in the "stopping or overriding of a mental process, in whole or in part, with or without intention" MacLeod (2007, p. 4), its componential architecture being understood to be closely entangled with that which underlies WMC (Brewin & Beaton 2002; Redick, Heitz, & Engle, 2007). Inhibitory control is one of Baddeley's (1996) core processes of the CE, tasked with the executive control, division and selection of attentional resources (see chapter 1).

Whilst Experiment 6 will look at the executive processes in detail, this experiment will prime subsequent analysis by endeavouring to provide an understanding of how age related processing factors may work in the RPA.

Experiment 4 will utilise reaction times to investigate the potential loading

effect on the reasoning process instilled by factors such as complexity and

distraction in the RPA. Experiment 4 will additionally attempt to provide

insight into the potential relationship shared between whatever agents

underlie the increased level of difficulty associated with these factors and

inhibitory skill- which remains the most widespread theory of how

executive control may constrain analogical thought.

Over the last two years various theorists have suggested an inhibitory role in

visually based analogical reasoning tasks through neuroimaging studies on

normal (Cho et al, 2010) and abnormal patients (Krawczyk, Henten, Wilde,

Li, Schnelle, Merkley, Vasquez, Cook, McClelland, Chapman & Levin,

2010a[48]). These have indicated the involvement of the prefrontal cortex and

parietal regions of the brain known to be related to the executive selection

and maintenance of information.

Other theorists have taken a psychometric approach, using inhibitory

paradigms in young adults/adults to show age based changes in the ability to

suppress irrelevant information (Chuderska & Chuderski, 2009; Viskontas

et al., 2004), the functions of inhibition being known to decline with age

(Zelazo, Muller, Frye, Marcovitch, 2004), while some have inferred the

---

[48] Krawczyk and his colleagues running an fMRI study on the RPA paradigm with a small
sample of Traumatic Brain Injury patients

effect from data patterns which are purported to represent inhibitory demands (Richland et al., 2004, 2006, 2010; Thibaut et al. 2010).

Inhibition is thought to work in analogical thinking by blocking prepotent responses and strategies or by pruning inappropriate semantic connections, thereby establishing relevant meaning (Dumontheil et al., 2010; Hummel & Holyoak, 1997; Krawczyk et al., 2008, 2009; 2010, Morrison et al, 2010; Oberauer et al., 2005; Thibaut et al., in press; Viskontas et al., 2004).

Featural responses may be considered prepotent for a number of reasons. They could represent novel strategies or forms of response which are based on prior beliefs and intuition rather than novel (current) situations that are contrary to the prior inclinations of the child (De Neys, 2006a; Diamond et al., 2002, a concept exemplified by heuristic[49] and analytical dual-reasoning processes (De Neys, 2006b). Alternatively the visual similarity of the (base) query object to a (target) featural distracter could also mean that the concept held in mind needs to be overridden in order to consider other relational objects (Oberauer et al., 2005).

Fortunately, this prepotent hypothesis is relatively easy to test in a simple RT experiment. Gordon and Moser's (2007) eye tracking-RPA paradigm suggesting that participants (in this case undergraduates) spend significantly

---

[49] In such a manner, 'heuristic' featural responses may be thought of as being innately preferable to 'rational' relational responses which are engaged only when time is taken to consider their relevance, heuristic objects presumably being subsequently chosen at a faster rate.

longer times looking at informative regions of the problem (relational or relational error objects), and spend longer on complex trials compared to simple; whilst using the PPA[50] Cho et al. (2007) predicts that as relational-complexity increases, so the required inhibition necessary to resolve the problem.

Although it is important to highlight the fact not every aspect of inhibition is purported to be explained by such a paradigm (i.e. the continuous inhibition of responses during protracted reasoning processes will not appear prepotent), it is postulated here that if erroneous attributional based objects are a prepotent response then featural responses should require less processing and be elicited faster than relational responses. Diamond, Kirkham and Amso (2002) proposed that anything under 3 seconds may be prepotent, with the likelihood increasing as RTs decrease.

Of course whether such proposed differences are observable in a RT study remains to be seen; however such a methodology does afford the additional advantage of further investigating the role WMC may play in analogy. In addition to this it also allows the additional testing of the complexity-constraint methodology; it being argued that the added processing power required for assessing the relevance of relational objects may potentially equate to longer selection times for relational objects than featural (relational objects being more complex than featural).

---

[50] The People Piece analogies are a visual form of the classical analogy.

In this experiment, a younger year group (class Year 2) was included to investigate the development of inhibitory skill and processing speed, while possibly explaining the floor effects reported earlier. It was accepted from the beginning that a RT experiment using the RPA would be susceptible to the repeated measures design of the scene based paradigm, where each question is repeated across four conditions and where each response is in a similar if not identical location.   In order to alleviate extraneous factors such as practice effects across conditions (which are identical in the RPA), only one RPA condition was administered to each child. This created year groups with four independent conditions (representing each of the RPA conditions), which were combined to form experimental conditions of 'no-distraction' (R1ND + R2ND), 'distraction' (R1D + R2D), 'simple' (R1ND + R1D) and 'complex' (R2ND + R2D).

The first question Experiment 4 aimed to address was whether relational responses in distracter and/or complex conditions took longer to elicit than no-distracter and/or simple conditions- providing additional evidence for/against the concept that complex/distracter conditions require greater degrees of processing power (WMC) than their counterparts (simple/no-distracter). The second question was to investigate whether featural responses are a result of prepotent stimuli, it being considered that if featural responses were the result of response-inhibition failure, then they would be elicited at a faster rate than thought-out relational forms of response. Richland et al. (2004, 2006) assumed that younger children had less

214

resources with which to inhibit prepotent stimuli, and would therefore perform worse at the RPA and inhibition tasks.

3.1.1 Participants. In total, 96 children were recruited from class Years 2 (ages 5-6) and 5 (ages 10-11) from two primary schools from the North-East of England in accordance with the criteria adopted in the earlier studies. Although no analysis would be conducted between individual RPA conditions and year groups due to the concerns raised in Experiment 2 about power, these children were divided up as equally as possible in the following manner: 9 Year-2 children in condition R1ND, 10 in condition R2ND, 12 in condition R1D and 10 in condition R2D; 12 Year-5 children in condition R1ND, 12 in condition R2ND; 14 in condition R1D, and 13 in condition R2D.
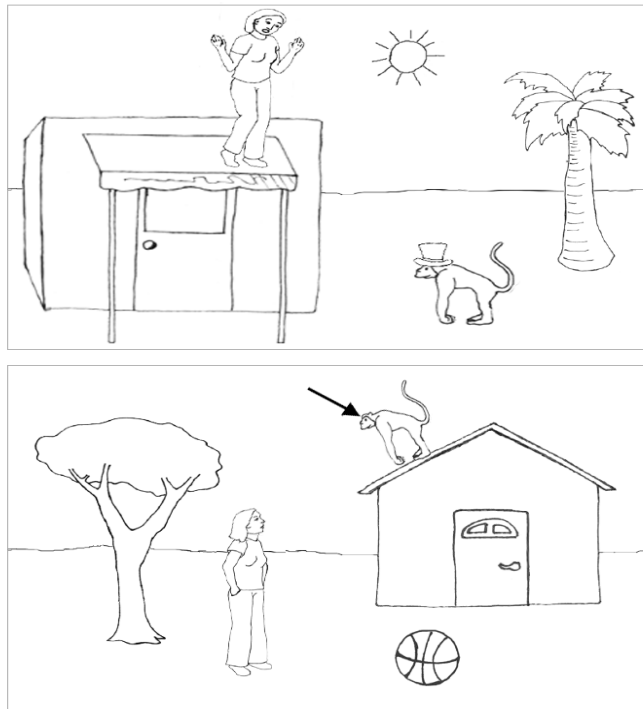
Figure 16. New 'Example C' problem added to the RPA. Constructed by the author this was introduced in order to encourage children to respond faster without detracting from the original instructions.

3.1.2 Materials and Design. Using a between-subjects design, the mean response times for relational and featural answers in Years 2 and 5 were compared using four experimental conditions (participants in each group received only one condition, experimental conditions being constructed from two independent RPA conditions): no-distracter (R1ND + R2ND), distracter (R1D + R2D), simple (R1ND + R1D) and complex (R2ND + R2D). Further comparisons were also made between types of response in all RPA conditions and in both year groups combined.

For this a computerized version of the RPA (Richland, 2008) was created specifically for use in this experiment. The same analogy problems as the paper version were presented on the screen, one at a time. Responses were recorded using a touch-screen monitor, the data including the child's response (i.e. featural or relational), as well as the amount of time taken to make the decision.

3.1.3 Procedure. This experiment was conducted in a single one-to-one session, lasting on average 15 minutes.

In addition to the standard instructions, an extra novel practice question (Figure 16) was created from the original object-drawings used by Richland et al. (2004, 2006) where the children were taught to respond as quickly as possible. This was done by inserting the following two paragraphs to the normal read-aloud standardized instructions:

 "We are about to do the last practice before we start. The final rule to the game is that we need to answer as quickly as possible. The quicker we answer, the more points we get… so it is very important to touch the correct answer as soon as you think you see it. Are you ready? Good, now as quickly as you can, touch the correct answer on the bottom picture. On your marks, get set, go.

[Child is shown the final practice question. Regardless of the child's response, the next paragraph is read out].

Remember to be as accurate as possible. If you select the wrong part of the pattern, or miss, you'll get no points! Have a closer look here, [experimenter, indicates by pressing an open space on the bottom half of the screen]. You see? That would be wrong.

Now again, the arrow shows us part of the pattern we have to find in the bottom picture. The arrow is pointing at the…. [child responds]. Very good, so what is the answer here? Remember as soon as you see it, touch the right answer! [Wait for child's response]. Here, the woman is correct. Again they're the same, but they look different. They are doing the same thing. The woman is on the roof, the monkey is on the roof."

Despite the apparent time limit, children were given as long as they needed to answer each question in accordance with the original RPA design, however they were frequently reminded to be as quick as they could (after they had responded) if they seemed to be taking too long.

Table 21. Mean relational and featural response times (by RPA condition and year group)

| | Year groups | No Distracter | | | | Distracter | | Simple | | Complex | |
		All	SD	R1ND+R2ND	SD	R1D+R2D	SD	R1ND+R1D	SD	R2ND+R2D	SD
Rel	Both	3892.58	1075.59	3602.73	792.02	4188.73	1243.61	3823.9	1065.59	3959.8	1092.54
	Y2	4219.97	1250.31	3772.32	890.79	4624.98	1404.14	4243.09	1277.47	4196.85	1255.3
	Y5	3645.5	854.21	3487.66	711.15	3822.28	974.87	3501.46	744	3784.2	940.79
						R1D+R2D	SD				
Feat	Both					4715.44	1993.41				
	Y2					5498.75	2211.79				
	Y5					4057.46	1544.7				

3.1.4 Results. Prior to analysis, strict procedures were put into place regarding data outliers. Touch-screen RT data was seen as being particularly sensitive to extreme scores from accidental presses or from the child's desire to select any response based on speed rather than the need to select a correct response.   Participant RT data was therefore rejected from each RPA trial that a) was above two standard deviations from the mean (i.e. not following instructions to select a response as quickly as possible) and b) less than 200ms, which was classed as anticipation.

In addition to this, three participants were excluded from the distracter analyses as they gave no featural responses in their distracter condition. These five participants were as follows: two Year 5 participants from condition R2D, two Year 5 participants from condition R1D and one Year 2 participant from condition R1D.

Table 21 shows the RT mean scores for Experiment 4.   In the following analysis two analyses would be conducted, one comparing featural and relational RTs in conditions of distraction, and one comparing relational RTs between complexity conditions. No 2x2 analysis of the four RPA conditions or 2x2x2, year x RPA conditions, would be conducted on the assumption that the division of responses in each RPA condition was too low to achieve significant computational power, this experiment being concerned with factors of complexity and distraction, not how these conditions interacted.
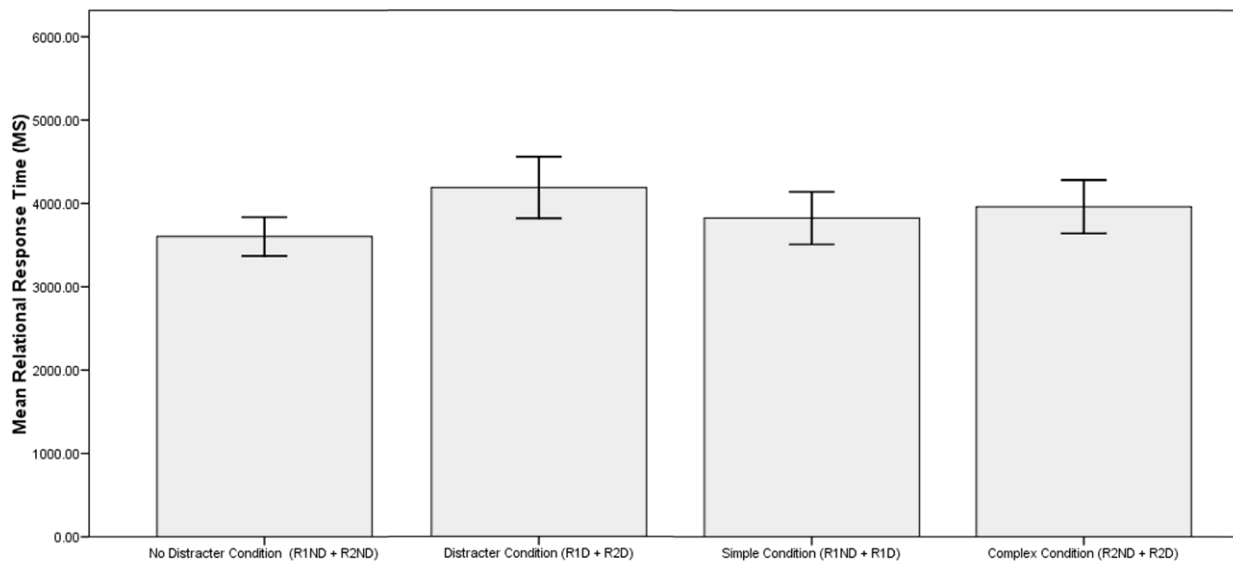
Figure 17. Mean reaction times (by RPA condition). Bars represent standard errors.

Distraction. Figure 17 shows the mean RTs for Experiment 4. Figure 18 shows the reaction times by each group. A paired samples t-test across all conditions reported that the mean RTs for relational responses was significantly less than featural (t(45)=-2.315, *p*<0.05, η²= .11), featural responses taking longer to elicit.

A paired sample t-test comparing relational RTs to featural RTs in both class years showed that Year 2's featural RTs were significantly greater than its own relational RTs (t(20)= -2.249, *p*<0.05, η² =.20), however no significant difference was observed between relational and featural RTs in Year 5 (t(24)= -0.921, *p*>0.05, η²=.03); older children taking the same average time to select a response. An independent samples t-test on the relation and featural RTs with the between subjects factor of class year showed that relational RTs (t(65.24)= 2.499, *p*<0.05, η² = .09) and featural

RTs (t(34.86) = 2.515, $p<0.05$, $\eta^2 = .15$) were greater in Year 2 than Year 5. That is, younger children took significantly longer on the RPA for either form of response.

Relational RTs were significantly longer in distracter than no-distracter conditions in both years combined (t(76.09) = 2.704, $p<0.05$, $\eta^2 =.09$), Year 2 alone (t(34.23)= 2.315, $p<0.05$, $\eta^2= .14$), but not for Year 5 (t(43.5)= 1.413, $p>0.05$, $\eta^2 = .04$).

Complexity. An independent samples t-test within the within subjects factor of complexity showed that relational RTs were the same in either condition (t(91)= -0.607, $p>0.05$ $\eta^2= .00$). The same analysis for Year 2 (t(37.99)= 0.115, $p>0.05$, $\eta^2=.00$) and Year 5 (t(49.19)= -1.215, $p>0.05$, $\eta^2 =.03$) also showed no effect of complexity

An independent samples t-test on relational RTs in complex conditions with the within subjects factor of class year showed no significant difference between class Years 2 and 5 (t(33.82)= 1.235, $p>0.05$, $\eta^2=.04$).
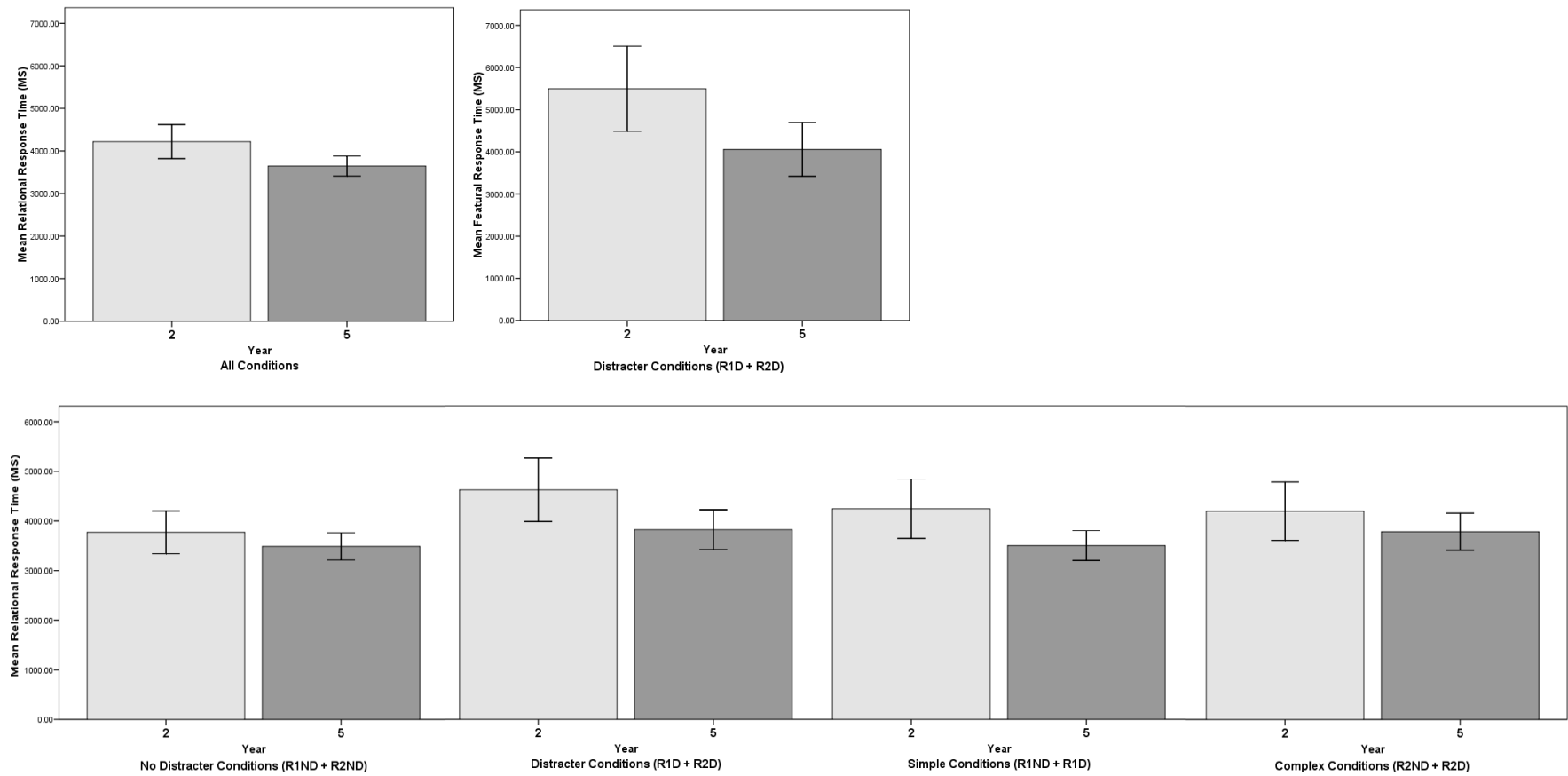
Figure 18. Mean reaction times (by year group and RPA condition). Bars represent standard error

3.1.5 Discussion. This experiment has demonstrated similar findings to that of Experiment 1 in that it has been shown that conditions of distraction increase task difficulty, increasing the time spent in the solution of the problem. While complexity factors have not been identified, this study provides practical evidence for the Richland et al. (2004, 2006) hypothesis that, despite understanding the similarity-constraint, younger children are less able to process conditions of distraction.

The developmental difficulty younger children appear to be demonstrating when processing problems within which a distracting featural object is present is thought to illustrate an important aspect of the floor effect reported in chapter 3. Representing cognitive growth in fluid intelligence, WMC and/or the increased efficiency of executive systems facilitating higher level processes (including processing speed).

As discussed earlier, one factor suggested by Richland et al. (2004, 2006) that may contribute to maturational analogical skill is inhibitory ability; younger children being less able to actively suppress alternative inappropriate methodologies, such as those presented by the attribute based featural distracter.

While this experiment further endorses such an approach, it is suggested that the potential role cognitive inhibition may play in AR may be centred less around the suppression of prepotent visual stimuli rather than the ability to cognitively select appropriate rationale/thought processes. I.e. being more

computational in orientation (Conway & Engle 2004; Friedman & Miyake, 2004[51]) and requiring more in-depth processing than simply dismissing something that simply looks the same- occurring regardless of whether a relationship is featural or relational.   Such an inhibitory role is indicated by spreading activation models of analogical thought, such as LISA (Hummel & Holyoak, 1997) where the synchronous binding of partially distributed representations in WM and LTM is determined by a selective activation process wherein a method of executive control inhibits competing node. Pruning semantic connections (in accordance with the goal) until a matrix of strong associations remain, representing our understanding of object relevance. Indeed, none of data reported here may be considered to suggest that featural responses are prepotent, with featural RTs in younger children considerably higher than relational RTs at almost 5.5 seconds. Diamond, Kirkham and Amso (2002) suggesting that anything up to 2,500 MS may be considered to be prepotent in 4 year olds, with prepotent responses typically falling between 1,000 and 1,500 MS in 5-6 year olds.

Although such a long featural delay suggests that young children are not selecting responses at random (as suggested by Thibaut et al, 2010) or because of anticipation, it also offers a counter explanation to that of inhibition- that in situations where the child is unsure of the answer they are taking longer to resolve the problem and are therefore more likely to make an incorrect response. Whilst pointing at a difficulty in processing which is

---

[51] For other non-processing roles of inhibition, see Rowe et al, 2010; Chiappe, Siegel & Hasher; 2002; Zacks & Hasher,1994.

similar to what Richland et al. (2004, 2006) is arguing, it may also be seen that what is being manipulated in distraction conditions is not forms of prepotent response, but the ambiguity of the problem (Goswami, 1992; Oberauer et al., 2005).   In which case considerable processing resources may have to be used in addition to Gentner one-to-one mapping comparisons in order to determine what the experimenter is demanding and what response is therefore most appropriate. How this processing may be thought of, and whether it can be explained by the concept of fluid intelligence or executive functionality of WM remains, including inhibitory skill, remains to be seen in the Experiment 6.

3.1.6 Conclusion. As in Experiment 1, Experiment 4 failed to associate effects of complexity with relational responding; a finding directly contrary to the Cho et al. (2007) study which found marked complexity differences and which illustrates the difficulties in establishing the concept of analogy in a field governed by multiple measures of AR.

However distraction, which in Experiments 2 and 3 has been demonstrated as being associated with aspects of fluid-intelligence and verbal fluency, was shown to be associated with what was hypothesized to be an increased processing load.   An important finding from Experiment 4 has been that maturation plays an important role in the solution of the RPA, and that older children are much less hampered by featural distracter then those who are considerably younger. This could imply that the effectiveness of the RPA to discriminate large distraction effects is limited when utilizing children

226

above a certain age ceiling (that of 5) and that the complexity/distraction interactions reported by Richland whilst using in older age ranges in the RPA (Richland et al., 2004, 2006, 2010) may be associable with factors other than the proposed WMC constraints. Children above the age of 5 appear to have sufficient processing power to complete the RPA at a high enough level to make meaningful error data virtually impossible to gain.

In conclusion, this experiment suggests that greater levels of processing may underlie the erroneous selection of featural responses than is predicted by a prepotent explanation.   Two assertions are believed to explain the pattern of data: that the processing difficulties exist due to either limited WMC (explained in the form of reduced processing power or inhibitory skill) or increased ambiguity, which has resulted in the selection of erroneous featural responses do to a lack of clarity regarding experimental demand. Experiment 6 will attempt to resolve this issue by assessing how the reported association between executive WMC and AR in Experiments 2 and 3 may potentially be represented according to the former: deconstructing the CE into individual executive functions described by Baddeley (2007) and Miyake (2000).

3.2 Experiment 5

The complexity-constraint theory (Halford, 1992, 1993, 1998) suggests that that capacity, defined by the number of conceptual chunks and segmentations of information in WM, may constrain AR. As children develop, so their capacity matures, expanding to a landmark developmental stage where they can begin to process complex relations previously out of reach (Halford, 1998). At the age of five-years it has been hypothesized that children are able to process ternary relations (Ratterman & Gentner, 1998), such as those used in the RPA, potentially explaining why WMC-complexity interactions have not been observed thus far.

To resolve this problem, the current experiment was designed to load the RPA by using Baddeley's (1986) dual task paradigm. This methodology requires the simultaneous application of a task alongside the primary measure of reasoning. If either task uses the same processing resources then an interference effect should be visible in the data, with a reduced level of task performance potentially being observed in either measure. As WM is modular (Baddeley, 2000), tasks can be run which use visual-spatial, phonological or executive (attentional) resources, or combinations of all three.

To this end a battery of three dual tasks was assembled for the RPA, the aim being to see if the scene based analogies created by Richland et al. (2004, 2006) were loading WMC enough for factors of complexity or distraction to be detected. The dual-tasks chosen were the articulatory suppression task,

the visual spatial tapping task, and the random number generation task: all three of which have traditionally been associated with the functioning of the CE (Baddeley, 1986; see section 1.3.1) and which represented the visuo-spatial/verbal/executive modularity of the WM model. The idea being that if AR strongly relied upon any of the individual modules storage/processing resources then creating additional competition for those specific visual/verbal or executive resources should be detrimental to the analogical process.

Again, a touch-screen version of the RPA is used, this time to allow the dual task to be performed more easily whilst conducting the main task. In order to reduce the loading effect of remembering previous trials, a between participants design was adopted with an increased number of participants in order to mitigate the difficulties extrapolated from the previous study.

To increase the statistical power of the paradigm, Experiment 5 would only utilize a single complex condition (i.e. R2, collapsing conditions R2ND and R2D into a single variable, complex conditions being indicated in complexity theory as representing a larger cognitive load). It would not compare the effect of no distraction/distraction or high/low complexity[52].

Since a relational analysis only would be conducted, both conditions would be combined in order to maximise the power of the design.

---

[52] As will be seen in the following section, the variance between these was not great, and was considered too small to ascertain significant results. Therefore, the experiment focussed on the effects of dual tasks.

In accordance with the observations from Experiments 2 and 3 which suggested an association between WMC and AR, the Hypothesis for Experiment 4 was that 'executive' WMC (storage and processing capacity defined by attentional control), rather than STS (storage components only), could be shown to be further associated with analogical performance by reducing the WMC available to reasoners, through dual task interference.

Given the results of the previous experiments we might expect articulatory suppression or spatial-tapping tasks to have little impact on relational performance, whilst more executive heavy tasks such as Random Number Generation may impede relational accuracy. This is because the latter requires the ability to manipulate stored information - hypothetically using the same WMC resources that are required to solve the analogical problem. Similarly, if a distracter condition in an RPA question uses more WMC than a no-distracter condition then it is conceivable that a drop in performance (in distracter conditions) may be visible in dual-task conditions which use similar resources: possibly providing evidence for the Richland assumption that attentional resources are necessary in the inhibition of featural stimuli.

3.2.1 Participants. For this experiment, due to difficulties observed testing younger participants using a touch-screen in Experiment 4, 119 children from Year 5 only (ages 10-11) were recruited from nine new primary schools in the North East of England using the ethical consent and recruitment procedures as previous studies.

3.2.2 Materials and Design.   Using a between-subjects 2x4 dual task design, normal analogical performance (measured by the number of relational responses elicited) in the two complex conditions of the RPA (with and without a distracter present: i.e. R2ND and R2D) was compared to performance under visuo-spatial, verbal and executive interference (the fourth condition being the control).

Each of the 8 conditions (including both controls) contained 15 participants with the exception of the random-number-generation trial within condition R2ND, which held 14 participants. For the primary task the experiment was conducted using the computerized version of the RPA under the same procedure as Experiment 4. Each dual task was chosen in accordance with the measures of EF popularised by Baddeley et al. (Baddeley, 1986; 2007) and that were known to reduce reasoning performance in certain domains (Baddeley, 1986, 1996b): each task covering one of the three primary slave systems of Baddeley's (2000) modular multiple resource model (visuo-spatial, verbal and CE).

Dual Tasks.

Control: No dual task was used.

Articulatory suppression (AS): Children were required to articulate the word "rhubarb" continuously at a rate of 1 per 1500ms.   After every problem a break was provided before progressing onto the next screen in order for the child to regain their breath.

Visual spatial tapping (ST): Children tapped out (unseen tapping) a pattern/sequence of 4 buttons (up, down, left right) on a USB numeric keyboard at a rate of one button press per 1500ms. The keyboard has hidden from view by the use of a dividing screen. A break was provided half way through the trial to allow the children to rest.

Random number generation (RNG): The random-number-generation task (RNG) was the original measure of executive functioning used by Baddeley (1966; 1987) to form the concept of executive attention as a resource. RNG is believed to measure the ability to suppress automatic responses in favour of directed possessing; it also is a measure of WMC.

For this task children were given a number-line diagram consisting of the numbers between one and nine. In accordance with the original design (Baddeley, 1966; 1996) they were then asked to vocalize as many numbers as they could in the space of two minutes, at a rate of one every two seconds. To aid them in this, a computerized metronome was introduced; giving a low background beat every 2000ms.

Although they could say the same numbers as many times as they liked (so long as they were between one and nine) they could not say the same number immediately after itself or say numbers adjacent to one another on the number line (e.g. if 3 was chosen they could not say 2 or 4 immediately afterwards). To help facilitate this, children were told to be as "random" as possible, the concept of which was aided by analogy that they were pulling

random numbers from a bag whilst blindfolded.   An additional rule was that they could not use 'easy' patterns to 'cheat' such as alternating between numbers or going up in two's or threes (in which case they were warned and asked to choose another, different number).

The task was measured by the amount of numbers successfully generated within two minutes. Answers which broke the rules did not count. For patterns the first number of each identified pattern was allowed whilst the others in the string were disallowed until the pattern had been abandoned.

Children were first given an unlimited practice time, at the end of which they were required to be able to articulate numbers randomly. During this practice they were first allowed to see the number line, and once they were able to generate numbers fluently, were then asked to generating numbers without it, before progressing to the main task.   No reliability data for this task was available.

After the experiment it was intended that the number of randomly generated responses in the task be measured for order- using a computerized measure of randomness (Towse & Neil, 1998). However this was not performed.

3.2.3 Procedure. Testing was conducted in a single session lasting 20 minutes on a one-to-one basis with the experimenter in a quiet location within a previously untested school. Each dual task condition was conducted simultaneously with the computerized RPA programme.

Due to time constraints imposed on this experiment study, only conditions R2D and R2ND were used out of the four conditions available. This, it was believed, would most likely provide the best chance to detect loading effects as both were complex (theoretically representing the largest cognitive load in the RPA) whilst including both a distracter and a non-distracter condition. Theoretically (according to complexity-constraint theory) condition R2D should also represent the largest demand on WM possible in the RPA.

3.2.4 Results. Table 22 shows the mean proportion of relational responses in each RPA and dual task condition in Experiment 5. Prior to the experiment an independent samples t-test was performed on the two RPA conditions to make sure that a significant difference was reported between the relational and featural responses.

In order to see if the dual tasks were effectively loading the RPA, a one-way ANOVA with the factor of 'dual task condition' (Control, AS, ST, RNG) was conducted for the proportion of relational responses elicited within both RPA conditions (R2ND and R2D). The post-hoc analysis would therefore allow a comparison between the control condition and the three dual task conditions.

Table 22. Mean proportion of relational responses (by RPA and Dual-Task condition)

	Control	SD	Spatial-Tapping	SD	Articulatory Suppression	SD	Random Number generation	SD
Condition R2ND	13	1.65	13.13	1.81	12	2.07	11.36	1.39
Condition R2D	12.27	2.52	11.07	2.87	10.27	2.74	11.67	3.18

For condition R2ND the ANOVA revealed a significant main effect of dual task condition for the relational score ($F_{[3,55]} = 3.37$; $p<0.05$, $\eta_p^2 =.16$), meaning that, overall, relational responding decreased when dual tasks were implemented alongside the analogical task.   However, post-hoc testing using the Bonferroni analysis revealed no significant difference between the number of responses within any of the dual tasks conditions (all $p>0.05$)[53]; the dual-task effect being prescribed to the difference between RNG vs. ST which was approaching significance ($p=0.051$). The second one-way ANOVA for condition R2D showed no main effects of the relational score ($F_{[3,56]} = 1.37$; $p>0.05$, $\eta_p^2 =.07$).

3.2.5 Discussion. Compared to the previous experiments, Experiment 5 proved to be inconclusive. The inability to find a dual task loading effect, so-often reported in other reasoning tasks, points to either experimenter

---

[53] In order to correct for multiple comparisons, in SPSS the p value corresponding to each comparison is multiplied by the number of comparisons (i.e. 6).

effects, (i) poor design methodology (e.g. the number of children, between methods design), or (ii) the fact that the RPA does not load WMC to a sufficient enough level for any effects to be observable.

(i) Poor design methodology comes into consideration when it is observed that the expected loading effect of dual tasks did not have the effect on relational responding that was anticipated. This could indicate that the dual-tasks chosen were not conducted in an appropriate manner and that, despite best efforts of the experimenter (which including the use of a metronome and practice sessions) the children were not articulating words frequently enough in the RNG or AS conditions. Not being random enough in the RNG condition, or not regularly following the correct sequence in the ST condition.

Tasks such as the RNG might also have been too *hard* to conduct as a dual task for this age range alongside the RPA- even with a touch screen. Conversely, another explanation is that that the low between-condition variance of the RPA was insufficient for differences to be observed between executive and domain-specific capacities. That is to say RPA was too *easy* for dual task effects to be shown. In order to judge whether this may or may not be the case it is proposed that in the future more attention be focused on dual-task as well as primary-task performance.

A final point regarding the methodology used is that the executive loading effect anticipated was not found because the RNG task failed to utilize

faculties which were similar or identical those used in relational reasoning. In such a manner it might be a potential avenue for further research to either increase the difficulty of the dual tasks or find those tasks which might conflict more during processing (such as retrieving names from a semantic set, such as the animal naming task).

(ii) It might be considered that the failure to find loading effects is entirely in-line with Goswami's relational primacy perspective. However, conclusions from non-significant relationships cannot be drawn, especially given the potential explanations for the failure listed in (i) above. Further research is therefore necessary to lend support to the concept that increased complexity does not represent an additional cognitive load.

3.2.6 Conclusion. Evidence provided thus far suggests that children may have enough WMC to be able to handle the RPA, a point which is in accordance with the view of Goswami (1992) and Richland et al. (2004, 2006, 2010), the latter of whom suggests that older children with greater WMC may be able to successfully manipulate similarity-constraints above younger children with lower WMC.

If this assumption is true then this experiment failed to show that older children could be induced into making errors in AR on par with younger age ranges, by lowering the available WMC with which to manipulate relational similarity. Yet while this potentially may be indicative of a lack of

cognitive weight associated with AR, it is more likely that the methodology used was insufficient for effects to be observed and conclusions be drawn.

Although it is accepted that this study was not successful in any of its aims, future studies should take into account the difficulties faced here, specifically the age ranges used, the dual tasks administered and the methodologies followed in the experimental use of those tasks (i.e. how the dual tasks are chosen). Regarding this latter point, this experiment suggests that because of the low power inherently associated with the RPA, more powerful testing be conducted with either larger participant pools or smaller pools using within groups designs (despite the dangers of practice effects in repeat-conditions).

The next experiment (Experiment 6) would now change focus, returning once more to investigate the association between analogical competency and cognitive faculties associated with Baddeley's (2000) model of WM. Although thus far in chapter 3 it has been suggested that relational responding in AR tasks might not tax WMC as much as might have previously be assumed- earlier findings from chapter 2 have already indicated that executive WMC may still strongly mediate this form of response. Experiment 6 would therefore reconsider this issue by exploring the role of individual EFs in AR (as indicated in chapter 2): its aim being to see if certain executive faculties may be associated with relational responding and whether this could account for the WM and IQ data reported in Experiments 2 and 3.

3.3 Experiment 6

Scene based analogies such as those used by Markman and Gentner (1993) and Richland et al. (2004, 2006) have shown that although WM may be necessary for the resolution of analogical problems, the ability to recognize the similarity-constraint may not be sufficient to explain differences in relational responding when a child is faced with conditions which offer alternative solutions to the task (Morrison et al., 2010; Richland et al., 2004, 2006, 2010, Thibaut, et al, 2010).

Richland and her colleagues prescribed these effects to the complexity-constraint hypothesis.   However, the theory also implicated executive functioning in mediating the relationship between maturation and the impact of distraction (Morrison et al., 2010)[54].

Executive functions (EFs) are processes involved in the active control of information in WM, the primary role of which appears not to be the representation of information but the division and selection of it (Baddeley, 2007). EFs are most commonly associated with precepts such as cognitive flexibility, multitasking, novel strategy development, the manipulation of information for meaning and, as discussed in Experiment 4, the ability to suppress non-relevant stimulus–response associations (Gilbert & Burgess, 2007).

---

[54] As Baddeley (2007) states, the concept that inhibition is an important component of WMC is not a new one. Inhibition having been conceptualized as a capacity since the work of Hasher and Zacks (1999), which showed age related changes in an attentional system could be strongly interpreted as a declining limited capacity system.

In the past decade, emergent evidence for an EF role in AR has arisen through neuroimaging studies (Cho et al, 2009; Christoff, Prabhakaran, Dorfman, Zhao, Kroger, Holyoak, & Gabrieli, 2001; Crone, Wendelken, Leijenhorts, Homomichl, Christoff & Bunge, 2009; Dumonetheil, Houlton, Christoff & Blakemore, 2010; Krawczyk et al, 2010b), and more recently Chuderska and Chuderski (2009), who, in an experiment involving several measures of EF and figural visually based analogies in an undergraduate population, found that AR was significantly correlated with goal monitoring (planning) and inhibition.

As stated in Experiment 4, despite recent interest in the field, little direct evidence has been shown for EF's involvement in AR in younger children, particularly in scene based analogies - the RPA being used to describe both the complexity-constraint effect and the role of inhibition in successful reasoning (Morrison et al., 2010).   This is an important gap in the literature; as shown in chapters 1-3, several possible interpretations of WMC's involvement in the RPA exist, and each has important implications on how we perceive a child's ability to develop analogical skill.   Specifically the question is raised as to whether EFs facilitate our understanding of an analogical problem via the formation of rule-sets, or whether they mediate performance by constraining parallel processing.

The current experiment was designed to further evaluate executive involvement in the RPA given the emphasis in Experiment 4 on prepotent

processes. This will utilize Baddeley's four (1995, 1996, 2000) and

Miyake's three (Miyake et al., 2000) component executive processes (see

chapter as well as the thesis' concept of processing (see chapter 2) to

construct a battery of six EFs which were considered to have theoretical or

previously shown involvement in AR.

The measures[55] (see below) selected were the Tower of Hanoi task or

'TOH' (planning, inhibition and control), STOP-IT (inhibition), Random

Number generation or 'RNG' (directed processing), a Local global Task or

'LGT' (task switching- this task was a replacement for the Wisconsin Card

sorting task which was found to be too difficult for participants at this

age[56]), the FAS task (processing efficiency, semantic recall) and a speed of

processing task or 'SOP' (processing efficiency).

Whilst the role of inhibition (Experiment 4) and SOP (Experiments 2 and 3)

has already been discussed, the other measures of EF described here may

also be considered to potentially contribute towards analogical success.

Directed processing and planning measures were chosen because in the RPA

a child is required to develop and maintain an operational schema which

must include, at least in part, the rule that featural objects must be

---

[55] Whilst none of the above functions are necessarily intended to be viewed in isolation
(speed of processing and inhibitory skill, for example, potentially being mediators of all the
measures chosen) it was considered that the experiment would help determine whether a
'blanket' faculty covering most cognitive faculties, or specific cognitive functions(s) helped
contribute towards analogical success.
[56] This represented a major change in paradigm for experiment 6, the WCST also being a
measure of rule learning and interpretation rather than just switch-cost.

dismissed; the child directing their responses accordingly in goal directed behaviour. The importance of planning, maintenance-of-aim (focus) and control are well known when considering the overall objective of the reasoning problem (Holding, 1989; Baddeley, 1996).

Task switching was incorporated because a switch cost between modes of thinking has previously been associated with analogical thought in the classical analogy format (Churderska & Chuderski, 2009; De Neys, 2006); and because switching between strategies may also be associated with mental fluidity (that is, how easily multiple relations and rules may be sequentially attended to in WM). In the RPA it is also highly relevant because of the necessity to direct thinking from any pre-held concepts of similarity (i.e. what looks the same) to novel task specific strategies (i.e. what does the same).   However, it should be noted that the RPA does not require the participant to continuously alternate between two opposing strategies (reverting from one to another), instead requiring them to maintain a *consistent* one. Planning and inhibition were therefore anticipated as possibly (although this by no means certain) being more explanatory of analogical success than task-switching.

A further executive function considered relevant was the Episodic-buffer's ability to associate (bind) data and retrieve relevant information from LTM. Whilst this remained difficult to practically test due to the uncertainty as to whether binding may effectively be shown to be an effortful, resource demanding process (which it appears it is not, Allen et al., 2006; Baddeley,

2007); functional aspects of it, such as retrieval speed and the strength of associations between semantic chunks may be measured indirectly through free-recall or verbal fluency measures. The strength of individual bindings are believed to contribute to the ability to think relationally and come to an acceptable conclusion (Hummel & Holyoak, 1993).

This experiment is designed to establish whether specific executive functions of WM (i.e. those defined by Baddeley, see p.65) are associated with relational performance. It is hypothesised that WMC constrains AR through the system-limiting distribution of attentional resources to faculties involved in the capacity to focus, divide or switch attention, or to retrieve information from LTM. In order to investigate this, the current experiment will adopt the same methodology as previous used in Experiments 2 and 3; specifically a 2x2 repeated measures design using the (2 level) factors of complexity and distraction (utilizing the relational scores as the dependent variable) and an experimental covariate in the form of six measures of EF.

3.3.1 Participants. Twenty-nine children aged between 10 and 11 (Year 5) were recruited from two primary schools in the North-East of England not previously enrolled in Experiments 1-5. Participants were recruited in accordance with the criteria previously described in the previous studies.

3.3.2 Materials and Design. In addition to a paper version of the RPA, six measures of EF were chosen (see below) whilst the Block design

243

was again chosen as a measure of non-verbal IQ. A simplified version of the Wisconsin card sorting task was also planned, but it proved too difficult for the children to learn given the scope/time frame of the study and this sub task was abandoned.

Tower of Hanoi task. Based on the original task design by Shallice (1982), the tower of Hanoi task is well known to be associated with the functions of the prefrontal cortex and executive processes therein (Cardoso & Parks, 1997). It has been described as a function of planning and strategy (Simon, 1975), and was chosen because of its association with inhibition (Goel & Grafman, 1995), updating, goal management, and conflict resolution (Wager et al., 2000).

For this version the child was presented with four disks of decreasing size placed on the far left of one of three pegs (largest disk on the bottom, smallest on the top). The child was asked to get all the disks from the peg on the left to the peg on the right. Specific rules were then given to complete this task, namely that any disk could be moved to any peg so long as only one disk was moved at a time and larger disks were not placed on top of smaller ones. The score was derived from the number of moves it took for each child to complete the puzzle. Reliability for the TOH task varies (Bishop et al., 2001) however, .72 has been quoted on a similar format of task by Gnys and Willis (1991).

Random number generation. The Random number generation task (RNG) was the original measure of executive functioning used by Baddeley (1966; 1987) to form the concept of executive attention as a resource. RNG is believed to measure the ability to suppress automatic responses in favour of directed processing.

In accordance with the original design (Baddeley, 1966; 1996) children are asked to vocalize as many numbers as they can in the space of two minutes, at a rate of one every two seconds. To aid them in this, a computerized metronome was introduced; giving a low background beat every 2000ms.

Although they could say the same numbers as many times as they liked (so long as they were between one and nine) they could not say the same number immediately after itself or say numbers adjacent to one another on the number line (e.g. if 3 was chosen they could not say 2 or 4 immediately afterwards). To help facilitate this, children were told to be as "random" as possible, the concept of which was aided by analogy that they were pulling random numbers from a bag whilst blindfolded. An additional rule was that they could not use "easy" patterns to "cheat" such as alternating between numbers or going up in two's or threes (in which case they were warned and asked to choose another, different number).

As the RNG trial was considered a reasonably complicated task, a minimum of five minutes training was given for each participant in this condition. During this time a number line was provided and the children were shown

what was meant by "cheating" with number patterns. Once they were able to generate numbers fluently they progressed onto generating numbers without it, before progressing to the main task. No reliability data for this task was available.

The task was measured by the amount of numbers successfully generated within two minutes. Answers which broke the rules did not count. For patterns the first number of each identified pattern was allowed whilst the others in the string were disallowed until the pattern had been abandoned.

STOP-IT. A computerized measurement of response inhibition in the prefrontal cortex (i.e. stopping a response that is otherwise automatic) the STOP-IT task (Verbruggen, Logan & Stevens, 2008) is based on the original Stop-Go paradigm of Logan (1994) where two forms of trial are presented to the participant: "GO" trials are where the child is expected to respond with a button press, and "STOP" trials are where a signal (is presented after the "GO" stimulus and the child is required to refrain from pressing the button.

The STOP-IT software was loaded onto a laptop. For this version of the paradigm children are asked to press a button on the left (be specific, which key?) if they saw a square and a button on the right (as above) if they saw a circle (the "GO" signal), but not to press anything if they heard the STOP signal (in this case an auditory cue). Throughout the practice trials the children were primed to answer as quickly as possible in order to create a

prepotent response.   Success was judged by a trial on which a STOP signal

was presented and the subject successfully managed to suppress their

response to press the key. Failure was judged by the inability to suppress

this response.

STOP-IT consisted of 32 practice trials, followed by 3 experimental blocks

of 64 trials each, with a delay of 250ms between the trial and the stop signal.

In accordance with Verbruggen et al. (2008), a number of measures exist for

STOP-IT, the majority based around the presence of the STOP signal.

- Mean probability of responding on stop-signal trials (PRS):

- Mean stop-signal delay (S-D)

- Mean stop-signal reaction time (SSRT)

- Mean reaction time on signal-respond trials (SR-RT)

- Mean reaction time on go-signal trials (NS-RT)

- Mean percentage of correct responses on no-signal trials (NS-HIT)

STOP-IT was chosen for this experiment as a measurement of the ability to

override automatic mental processes within WM (Macleod, 2007),

specifically the suppression of irrelevant information, highly likely

alternatives, and/or recently examined stimuli (Nigg, 2000; Rafal & Henik,

1994).   However, there is substantial debate as to how inhibition measures

may be interpreted and whether they represent the ability to dismiss non-

relevant thought (either task relevant or not-task relevant) or whether they

are measures of the efficiency of executive decision making processes

(Macleod, 2007). No reliability measures were available for the STOP-IT task.)

Verbal fluency: FAS/animal-naming task. Two well established measures of the function of the CE are the FAS/Animal naming tasks (Borkowski, Benton, & Spreen, 1967). These are measures of "directed processes" within the prefrontal cortex (Perret, 1974), specifically verbal fluency (Shelton et al., 2010) and search and retrieval processes from LTM (Baddeley, 1996).

For this task, participants had 90 seconds (per category) to generate as many words as possible from categories which included words starting with the letter F, then A, then S, followed by animal names. The total score was based on the correct number of appropriate responses minus the number of incorrect responses (i.e. repetitions or non-appropriate responses not fitting the category). No reliability measures for the FAS or animal naming tasks was available.

Local-global task. A measure of the executive ability to shift attention (Baddeley, 2000) the local global task (LGT), requires participants to switch between specific operations, specifically one information set to another.

Children were given 3 trials, each of which consisted of a sheet of paper (trials one, two and three) containing 52 geometric shapes known as Navon figures (Wager et al, 2000) which are themselves made up of smaller geometric shapes. Navon figures are either congruent (such as a triangle

made up of smaller triangles) or incongruent (such as a square made up of identical circles).   For each trial participants were given 90 seconds to count the number of sides each figure was composed of: for trial one these were the sides of the larger Navon figure (one for a circle, three for a triangle or four for a square). For trial two (an identical sheet to trail one) they were asked to ignore the larger shapes and focus on the number of sides of one of the smaller shapes within the figure (i.e. if a circle was made up of triangles, the answer would be three). For trial three (a new randomised order which was standardized between participants) they were asked to alternate between the two strategies learned in the earlier trials; the final score being the number of figures correctly identified in trial three. No reliability data was available for this subtask.

Speed of processing task (SOP).   The local global task is divided into three trials, the first of which requires the child to complete a speed of processing task which can then be used to compare against the second trial if the task is scored by the difference between trial 1 and 2. Since this experiment used the alternative form of marking for the LGT (that of the third trial which consisted of the ability to swap between two modes of thought) the first trial was used as a measure of speed of processing.

3.3.3 Procedure. This experiment took place over the course of three sessions lasting an average of 30 minutes each.   For the first session, RPA conditions R2D and R1ND were administered followed by the Block design task. The second session consisted of the animal naming, RNG,

FAS/Animal naming, Local/Global and Tower of Hanoi tasks. The third

session consisted of the RPA conditions R1D and R2ND and the

computerized STOP-IT task. For each executive functioning task which did

not have a practice session included, participants were given five minutes of

training task prior to testing, except for the RNG, which was given an

unlimited practice time at the end of which the child was expected to be able

to fluidly articulate random numbers.   The range for this was between five

and ten minutes.

     3.3.4 Results.   See Tables 24 and 25 for the mean scores for the

experimental measures/RPA in Experiment 6. During testing it was found

that children inexplicably underperformed at the FAS task (mean response

was 19 although several scored under 10, which is exceptionally poor),

stating that it was too hard to complete, despite giving enough answers to

show that they were able to complete the task.   Post-testing it was decided

that the Vikki and Hoist (1994) methodology might have been a more

appropriate methodology to adopt for the FAS measure (where children are

measured not on the number of items they produce, as was used here, but on

the time taken to produce a set number of responses). However, no timings

for the FAS had been undertaken and this was impossible: meaning that the

analysis was confined to the amount of words produced.

Table 23. Mean relational scores (by RPA condition)

Relational Score	Mean	SD
All Conditions	63.34	6.24
Complex (R2ND and R2D)	30.66	3.89
Simple (R1ND and R1D)	32.7	3.28
Distracter (R2D and R1D)	30.52	4.3
No-Distracter (R2ND and R1ND)	32.83	2.7

Table 24. Mean experimental measure scores.

Measure	Mean	SD
Block Design Task	42.79	11.19
FAS Task	20.17	7.43
Random Number Generation (RNG)	19.55	4.48
Local Global Task	14.57	7.12
Speed of Processing (LGT task A)	30.035	2.32
Tower of Hanoi	42.55	24.12
PRS (STOP-IT)	31.2	6.66
SSD (STOP-IT)	140.6	121.75
SSRT (STOP-IT)	144.6	80.58
SR-RT (STOP-IT)	450.7	85.91
NS-RT (STOP-IT)	476.9	95.59
NS-HIT (STOP-IT)	73.4	5.68

Correlation analyses were then conducted between relational responding and each subtask (Table 25).

For this analysis, the TOH task (r=.438, p<0.05) and SOP tasks *(r=.408, p<0.05)* were found to be significantly positively correlated with the relational score, whilst the correlation between the block-design task (r=.359, p=0.056) and local global task (r=.333, p=0.078) and the relational score was found to be approaching significance. The FAS (r=.016, p>0.05) and RNG (r= -.091, p>0.05) were not significantly correlated with relational responding.

For the STOP-IT task each measure was assessed individually, the PRLS (r= .148, p>0.05), SSD (r=-.205, p<0.05), SSRT (r=.087, p>0.05), NSRT (r= -.185, p<0.05), and NS-HIT (r=.239, p>0.05) were all not significantly correlated to relational responding. However, the SS-RT (the signal-respond reaction time) was found to be approaching significance (r=-.365, p=0.051).

Signal respond trials are described as trials in which a stop signal occurred but the subject failed to withhold a response (De Jong, Coles, Logan & Gratton, 1990), in the case of STOP-IT this is the mean reaction time for such trials (Verbruggen et al., 2008). Such data may best be viewed from the 'first out of the gate' model of inhibition (Logan & Cowan, 1984) which suggests that a degree of competition exists in STOP-GO paradigms whereby attended control is pitted against unattended heuristic processes;

the first process to be activated being the one that is acted upon. In this experiment increased relational responding was correlated with participants who were slower at eliciting non-suppressed prepotent responses, meaning that out of all of those who responded incorrectly, those who were not as quick to go on a no-go trial were more likely to have a higher relational score.

Since the mean reaction time on go-signal trials (NS-RT) was not negatively correlated with the relational score it is unlikely that this represents participants who were prudently waiting longer to see if a stop signal was elicited, suggesting instead that some form of processing conflict was being reflected in the SS-RT delay. This can be interpreted in a number of ways, but two explanations are considered the most plausible: that these are either individuals whose executive control processes are, despite the error, more effective than their counterparts, in which case this may conceivably be a measure of inhibition, despite the proportion of incorrect no-go responses (PRS) being uncorrelated. Or that this is a measure of processing speed as these individuals have been quick enough to recognize that there is a conflict between controlled and prepotent responses but have been unable to successfully resolve it.

Table 25.   Correlational matrix for experimental measures and relational responding

	Block	FAS	RNG	LGT	SOP	TOH	STOP (PRS)	STOP (SSD)	STOP (SS-RT)	STOP (SR-RT)	STOP (NS-RT)	STOP (NH-HIT)
Relational	0.359†	.016	-0.091	.333	.408*	.438*	.148	-.205	-.087	-.365†	-.185	-.239
Block		.282	-0.09	.364†	.368*	.133	.095	-.009	-.293	-.136	-.263	-.022
FAS			.229	.408*	.447*	.223	-0.001	.152	-.517**	-.083	-.235	-.241
RNG				.164	-0.001	.028	.248	-.105	.112	-.064	-.021	-.003
LGT					.705**	.333	.166	.137	-.416*	-.156	-.166	-015
SOP						.353	.195	.146	-.499**	-.181	-.228	-.204
TOH							.045	.025	.001	.123	.026	.007

† Correlation is approaching significance between the 0.05-0.06 level

** Correlation is significant at the 0.01 level

* Correlation is significant at the 0.05 level

Repeated-measures analyses of covariance (ANCOVA) were conducted

using the mean centred approach (Delaney & Maxwell, 1988) adopted

earlier, the ANCOVAS being carried out for the TOH and SOP tasks

respectively in order to see if these measures could provide a complexity

interaction (missing from Experiments 2 and 3) which may be interpreted as

a function of WMC.

TOH. For the TOH covariate analysis, the main effects of

TOH ($F_{[1,27]}$ = 6.40; $p<0.05$, η_p^2 =.19), complexity ($F_{[1,27]}$ = 10.81; $p<0.05$, η_p^2

=.29) and distraction ($F_{[1,27]}$ = 11.43; $p<0.05$, η_p^2 =.30) were identified.

Interactions were also found between complexity x TOH ($F_{[1,27]}$ = 5.14;

$p<0.05$, η_p^2 =.16) and complexity x distraction ($F_{[1,27]}$ = 44.11; $p<0.001$, η_p^2

=.62). No interactions were reported for distraction x TOH ($F_{[1,27]}$ = 0.04;

$p>0.05$, η_p^2 =.001) or complexity x distraction x TOH ($F_{[1,27]}$ = 2.44; $p>0.05$,

η_p^2 =.08).

Figure 19. Interaction between complexity and the TOH task. The median

split for High/Low performers in the TOH task was introduced at 34 moves.

Paired sample t-tests reported a significant difference between condition R1

and R2 in low TOH performers ($t(12)$= 3.196, $p<0.05$, η^2=.46) but not

between high TOH performers ($t(15)$= 1.444, $p>0.05$, η^2 =.12).

An independent sample t-test with the grouping variable of TOH

performance found no significant difference between high and low TOH

performers in condition R1 ($t(27)$= -0.485, $p>0.05$, η^2=.01) or condition R2

($t(27)$= 1.659, $p>0.05$, η^2 =.09).

As can be seen the difference between simple and complex conditions was

less for participants who were better (high) performers at the TOH (i.e. took

the least number of moves to complete the task)
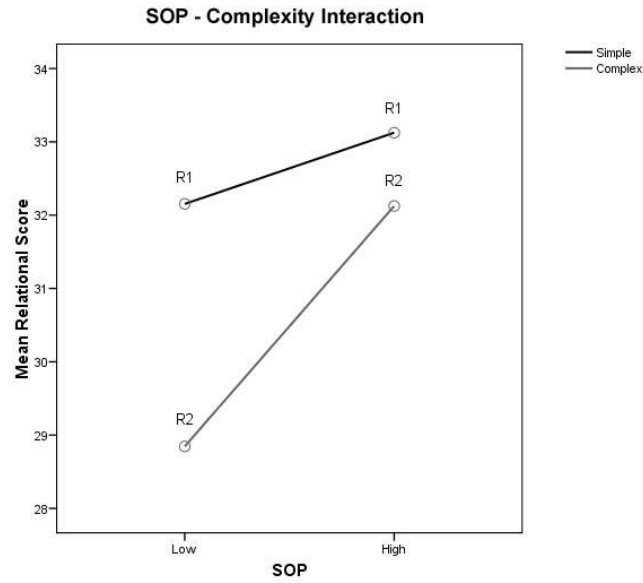
Figure 20. Interaction between complexity and the SOP measure. The median split for High/Low performers in the SOP measure was introduced at the 29 level.

SOP. For the SOP task, a significant effect of SOP ($F_{[1,27]} = 5.38$; $p<0.05$, $\eta_p^2 =.17$), complexity ($F_{[1,27]} = 10.17$; $p<0.05$, $\eta_p^2 =.27$), and distraction ($F_{[1,27]} = 11.46$; $p<0.05$, $\eta_p^2 =.30$) was reported, as well as significant interactions between complexity x distraction ($F_{[1,27]} = 44.46$; $p<0.001$, $\eta_p^2 =.62$). No interactions were observed between complexity x SOP ($F_{[1,27]} = 3.24$; $p>0.05$, $\eta_p^2 =.11$), distraction x SOP ($F_{[1,27]} = .10$; $p>0.05$, $\eta_p^2 =.004$) or complexity x distraction x SOP ($F_{[1,27]} = 2.67$; $p>0.05$, $\eta_p^2 =.09$) were reported.
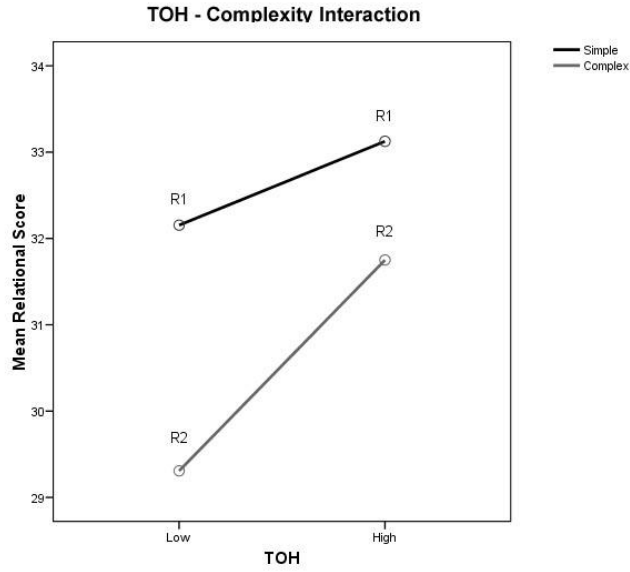
Paired sample T-tests reported a significant difference between condition R1 and R2 in low SOP performers (t(11)= 2.818, $p<0.05$, $\eta^2=.42$) but not between high SOP performers (t(16)= 1.570, $p>0.05$, $\eta^2=.13$).

An independent sample t-test with the grouping variable of SOP performance showed a significant difference between high and low SOP performers in condition R2 (t(27)= -2.122, $p<0.05$, $\eta^2=.14$) but not condition R1 (t(27)= -0.714, $p>0.05$, $\eta^2=.02$).

3.3.5 Discussion.   These findings support the view that executive functions are involved in analogical thinking. Although it has not been demonstrated that specific processes are associated with relational responding (such as task-switching, or inhibition); Experiment 6 has never-the-less demonstrated that a more general approach to cognition (i.e. one which takes into account the speed and efficiency of the systems utilized, as well as planning and the capability to construct strategies beneficial to the outcome of the task[57]) may account for analogical success in the RPA. It has also provided important evidence that executive ability may mediate relational responding under conditions of increased complexity (the complexity-constraint hypothesis).

---

[57] Flexible representations of goals and intentions being a salient component of our concept of executive functioning (Gilbert & Burgess, 2007) in that they represent non-routine processes where the ability to recognize and adjust sub-optimal performance in reasoning.

This conclusion was reached from the observation that both the SOP and TOH tasks were correlated with high relational scoring, whilst participants efficient in TOH or SOP appeared less affected by higher levels of complexity: their relational scores in complex condition being more comparable to the proportion of responses elicited in simple conditions. Whilst it is understood that conclusions ascertained from a correlational study are risky in that they may be the result of extraneous variables, this finding is in-line with the earlier research from Experiments 2-6 which reported a moderate-strong contribution of SOP/Verbal Fluency towards successful AR in the RPA. It also supports the work of Chuderska and Chuderski (2009) who described planning and strategy as a key attribute in AR[58].

While the potential role of SOP in AR has been described in previous experiments[59], the role of the TOH task can be viewed from two perspectives. Either as a general ability to develop strategies and learn rules; or as a 'mixed bag' of specific processes, such as task-switching, inhibition of ineffective behaviour and the focusing of attention. In either case the two main theories of AR, relational-primacy and complexity-constraint, interpret the contribution of this measure differently.

---

[58] Although this study utilized visually based classical analogies and not scene-based analogies.

[59] In the context of EFs the explanation of SOP's involvement is likely to fall into two categories, representing either more efficient bindings and/or the increased efficiency of the executive selection processes

Relational primacy theory (Goswami, 1992) presents a hypothesis which is based on the creation of schemas (non-automatic actions or strategies that have become automatic) to facilitate the resolution of the task: describing task-understanding as a performance factor for analogical success. Although no connection is ever drawn by Goswami directly connecting her theory to Norman and Shallice's (1986) conceptualization of the CE, it is suggested here that it is self-evident how the formation of new and the adaptation of old schema, as described in the SAS model (see chapter 1), may be seen as being synonymous with Goswami's view of rule-building. The processes associated with the TOH task (planning and directed goal driven behaviour) are argued to reflect part of the ability to interpret experimental demand.

Like Gentner (1983) Goswami's (1992) relational primacy theory potentially explains the TOH's mediatory effect of complexity through mapping, which allegedly only occurs when a problem is ambiguous: in this case through an increased number of relational-error objects which would represent increased levels of task ambiguity.

 It could be that a better understanding of the rules represents the ability to determine a better-fit from the options available, however this still requires an assessment of the qualities that the base and potential targets possess (mapping) and unfortunately Goswami does not describe how mapping may take place in accordance with her views on cognitive loading.

Halford's complexity-constraint theory (1992, 1993, 1998), as well as (computational) spreading-neural-activation-models such as LISA (Hummel & Holyoak, 1997, 2003) are more forthcoming in terms of how the TOH may be applicable to mediating complexity. As with relational primacy theory, these also highlight the role of the CE in the AR, the TOH task most likely representing the ability to ascertain the more optimal paths for task success, or the ability to inhibit incorrect modes of thought. The main difference being that Halford suggests that children with greater WMC are better able to represent more complex relationships during mapping and are therefore more likely to be better relational-thinkers..

Both theories may therefore claim to predict the complexity data presented in Experiment 1, as well as the TOH's/SOP's mediation of complexity reported here. Goswami, through the difficulty in initially representing the problem and how it may be resolved- a greater understanding reducing ambiguity among similar (relational-error) objects. Halford through greater levels of WMC and the number of relations accessed simultaneously.

While given the age ranges tested it has been difficult to establish in Experiments 1-6 that increased complexity may or may-not load WMC, the question must still be asked as to whether ability at the TOH may be representative of WMC resources, and therefore interpretable within the boundaries of complexity-constraint theory?

Despite a relationship no firm relationship being established between WMC and the TOH task, it is possible that elements of the TOH measure may be considered representative of WMC. In the past VSWM and executive WMC have been strongly associated with the TOH task, as well and the ability to plan (Handley, Capon, Copp & Harper, 2002; Numminen, Lehto & Ruoppila, 2001; Welsh, Satterlee-Cartmell, & Stein, 1999), findings which are in-line with the Mr.X effect reported in Experiments 2 and 3.

Reassuringly, for the complexity-constraint perspective Halford (1992, 1993, 1998) also believes that WMC constraints may be represented by SOP, which mediates the number of relationships that may be processed in parallel and the efficacy of sequential processing. Increased processing speed having recently been linked to the ability to complete variations of the TOH task (Sorel, & Pennequin, 2008), something which is supported by the SOP findings from Experiments 2, 3 and 6, despite no significant association being observed between SOP and TOH here.

It was therefore considered possible that the interactions between the TOH and SOP tasks and complexity might be interpreted as being representative, at least in part, of the thus-far unidentified WMC constraint effect on AR postulated by Halford (1992, 1993, 1998).

Inhibition and task switching. It was predicted that inhibitory skill or switch cost would be directly associated with RPA performance, either of

which would provide evidence for the selective control of attention within WM. However, this was not the case.

Despite this, it should be noted that the possible lack of prepotent-responses (which the STOP IT task is designed to measure) had already been predicted in Experiment 4. Whilst the risk of Type I or II errors is always great when interpreting null findings, this could indicate that the selection of relational objects in the RPA is not as automatic as has been previously suggested (Morrison et al., 2010) and may be more controlled. In which case featural responding may be less about prepotent distraction (as suggested by Richland et al., 2004, 2006) than about a failure to fully comprehend the goal of the task (Goswami, 1992), and correctly inhibit irrelevant conclusions. Future research should readdress this issue.

3.3.6 Conclusion. Executive control has been proposed as an explanation for the pattern of relational responding reported in Experiment 1, both through the functionality of skills such as planning and goal maintenance, as well as computational aspects such as processing efficiency and the speed of EFs.

Although the role of inhibitory skill has been strongly indicated in previous research (see Experiment 4) the current research remains inconclusive as to whether this may be applicable to the paradigm being used here (i.e. the RPA). It is suggested that one direction for future research explores this possible role, specifically determining whether featural relationships

represent a heuristic (non-analytical) form of response, or whether, as suggested in Experiments 4 and 6, inhibition is more actively involved. The implications of this experiment are further discussed in the thesis conclusion (chapter 4).

Chapter Four

4.0 Discussion

      4.0.1 Summary of Experiments 1-6. Experiment 1 reported a

pattern of data similar to Richland et al. (2004, 2006) suggesting that

relational responding was arbitrated by conditions of increased complexity

and/or distraction.   Experiment 2 compared psychometric measurements of

IQ and WM as predictors of AR. Results suggested that although WMC

(from the Mr.X measure) appeared to be significantly contributing to AR,

IQ appeared to be the greatest contributor to relational responding.

Experiment 3 extended the research on WM's perceived involvement using

additional measures which measured storage and processing roles. The

results suggested that the constraint effects that had been demonstrated in

the previous experiment were more likely to be related to processing than

storage aspects of the WM model.   Verbal fluency was also suggested as a

contributing factor to both AR and WM performance, increased efficiency

being indicative of increased relational responding.

Although Experiment 5 failed to show significant WMC effects,

Experiments 4 and 6 provided further evidence to support the notion that

differences in the processing capacities of WM could account for individual

differences in task performance. Experiment 4 suggested that although

inhibitory skill might be involved in AR, it was unlikely to involve

prepotent forms of response. Furthermore, Experiment 6 showed that

executive planning, processing speed and efficiency were strongly associated with AR.

4.0.2 Relational primacy and complexity constraint theory. By highlighting the role fluid intelligence and executive processes may play in the resolution of analogical scene-based problems where the similarity-constraint is recognized and understood by the participant prior to testing, this thesis has demonstrated that the relational primacy theory (Goswami, 1992) may not be an adequate explanation of AR.

Relational primacy. Despite understanding the analogies used in the task, as evidenced by both the high success rate in Experiment 2-6 and by Richland et al.'s claims that the similarity-constraint is capable of being represented (as claimed by Richland et al. 2004, 2006, 2010); children between the ages of 5 and 10 consistently elicited a pattern of relational responses that were indicative of competing response-stimuli influencing the mapping (comparison) stage of the reasoning process. Thus an extreme knowledge approach, i.e. one that proposes that all a child needs to successfully reason analogically is to recognize the underlying analogy, is insufficient to explain the data presented in Experiments 1-6.

This conclusion is partly due to the fact that the relational primacy theory was never fully developed to take into account the ambiguity of relational problems. Formulated in the early 1990's, Goswami's approach was advanced almost entirely as a reaction to the traditional structuralist

argument that young children were unable to reason analogically. As such it was highly successful, providing almost insurmountable evidence that even the youngest ages can and do reason relationally, so long as they understood the task and what was required of them.   However, resultant criticism surrounding the theories inability to describe how distracting featural stimuli may affect mapping, or what defines ambiguity, has meant that it has been increasingly left behind in an age where information processing theories seek to provide increasingly accurate computational accounts of the reasoning process.

In defence of relational primacy one of the greatest problems in the literature has been to assume that the theory does not take information processing into account. In fact it does, asserting that mapping is only necessary when the analogy is unclear; individual differences in analogical success being explained by performance factors mediating the ability to form abstract schemas appropriate for the resolution of the problem.   Such a conceptualization is based on the work of Norman and Shallice (1988) and Piaget et al (1977) and it is likely that such functions relate to a child's general reasoning ability, fluid intelligence and executive WMC, all of which play a role in determining relevance and the division and selection of appropriate (relevant) responses.

If this is supported then children will construct an initial schema for the RPA during training with individual differences determining how well the overall goals of the task are understood from the rules provided. This

schema will then be developed across the task to accommodate the individual base scene before being challenged in the target scene by competing forms of response (objects which share a high number of relations with the base query object increasing ambiguity).

From this evidence it is suggested that if relational primacy theory is developed to better incorporate the latter-stages[60] of reasoning (mapping) in circumstances where the similarity-constraint may be fully understood; then such a perspective may, in the future, prove to be a valuable contributor to the field, beyond its already influential work in providing evidence for young-reasoners (chapter 1).

Complexity-constraint theory. Given the above it must therefore be considered whether complexity-constraint theory, which predicts that higher WMC equates to the degree of complexity that can be processed in parallel (Andrews & Halford, 2002; Halford, 1992; 1993; 1998; Halford et al., 2002) is more explanatory of the mapping process than relational primacy.

In this thesis a correlation between WMC and relational responding has been indicated in Experiments 2,3,4,5 and 6. This provided evidence which could be construed as supporting both relational primacy (through task understanding) and complexity-constraint (through capacity limits) theories.

---

[60] I.e. by moving away from a stance which suggests that AR will always be successful when the similarity-constraint is fully understood; incorporating the possibility that errors may be made when candidate relations compete.

However, important evidence for the latter was missing, specifically crucial WMC/complexity interactions.

It has since been argued that SOP, verbal fluency, and the efficiency of executive systems may represent Halford's concept of WMC (Halford, 1992, 1993, 1998) by placing limitations on reasoning. These constraints may account for the IQ-complexity interactions reported in Experiments 2, 3 and 6, in that those participants who are more efficient processors are more likely to choose relational forms of response in conditions where complexity is higher.

Early in the thesis it was argued that the children tested in this series of experiments may have passed the ability threshold to process ternary relations (Halford, 1993, 1998, Halford et al., 2004). This ability to maintain and process ternary relations (the developmental milestone being described by Ratterman and Gentner (1998) and Halford (1998) as taking place around the age of 5) could explain the lack of significant WMC-complexity interactions in Experiments 2 and 3; as well as a lack of significant individual contributions to relational responding from WM measures other than Mr.X. children can represent the relationships involved in "chasing being chased" *ergo* increased WMC may not be as beneficial to older children.

However, Halford himself states that only 50% of children may be able to process such problems at the age of five and that this ability grows during

childhood (Halford, 1993, 1998), meaning that it was by no means certain that individual differences in WMC were constraining the problem.

It is therefore proposed that the data presented thus far supports processing models such as those described by Halford et al, which allow the possibility of processing constraints being present in the latter stages of the mapping process. It is suggested that the CE plays a directorial role whereby the control of attention results in the selection or inhibition of relevant meanings in regards to the goal. As suggested earlier increased efficiency (i.e. verbal fluency and processing speed) appears to facilitate the reasoning process, allowing the CE to establish and reject bindings faster and more accurately (as in Hummel & Holyoak's (1997) LISA model)[61], as well as possibly placing a limitation on the speed in which relations may be processed in parallel.

Summary: Relational primacy and complexity-constraint theories encompass different focuses on early and late stage processing in analogical problem solving. As has been shown (chapter 1) the two theories possess a number of individual standpoints that make them unique, however besides from Goswami's statement that knowledge of the similarity-constraint is all that is needed to reason successfully, the chief difference between the two may be reduced to Goswami's (1992) belief that additional complexity should not represent an additional cognitive load for the reasoner.

---

[61] Interestingly high verbal fluency measures may also indicate stronger, more-established connections allow more efficient chunks of data to be recalled, reducing the processing requirements.

As previously noted (chapter 1), the term 'complexity' is perceived differently by different theorists, however the RPA (which follows Gentner's (1983) and Halford's (Halford, 1993, 1998) views of complexity) relies on the concept that the more descriptive axes exist, the more relationships are needed to be maintained in parallel and the greater the processing power required to select the correct response.

Despite this, recent studies investigating the processing weight of conceptual 'chunks' (containing multiple dimensions of bound arguments) are surprising in that they suggest that bound information and the conceptualisation of groups of relationships may not be as resource demanding as might have been imagined (Baddeley, 2007; Baddeley & Hitch, 2006). The WM system having been shown to be capable of holding reasonably complex bindings (Allen et al, 2007; Vogel, Woodman & Luck, 2001): meaning that increased complexity may not necessarily equate to a greater processing demand.

In short, the practical value of considering complexity as being demanding of WMC resources may be considered limited if the demands of the task do not tax the overall processing capacity of the child. Equally, if the child has passed a developmental milestone which means they have sufficient cognitive resources to handle effects of distraction and/or complexity- then interactions may be so small that their precise measurement, as well as their value as an indicator of analogical success, will require highly detailed computational modelling. It remains to be shown whether extraneous task

demands on measures of analogical reasoning such as the RPA, may be able to reduce WMC to a level where performance is significantly reduced.

4.1 Conclusion

Overall, the present study has helped define the role of WM and fluid intelligence in analogical thought, offering support for theories which focus on the active processing of relations maintained in conscious thought.

Although it is suggested that the relational primacy theory be expanded upon to increase late-stage reasoning processes in order to address the concerns of modern structuralists, these results do not clearly distinguish whether the relational primacy or complexity-constraint theory may be seen as applicable to the current study. The present findings failing to address the WM requirements imposed on reasoning through an increased number of contextual bindings in relational objects (the core element which separates the two theories). Further research will be required to ascertain the true loadings of relational objects in the RPA, wherein the proportion of descriptive relations is manipulated against interference from dual-tasks,

It is still not clear what role the slave-systems of WM play in the analogical process when the demand of the (scene based) problem is within the boundaries of the constraint theoretically imposed by WMC on AR; although the current research suggests that the VSSP and PL may play a secondary role to higher-order executive decision making processes

associated with fluid-intelligence, the central executive and/or the episodic buffer.

Such a theory is consistent with the idea of selective attention and executive control outlined by Baddeley (1996[b]; 2000) and Norman and Shallice (1986). For these theories, active attentional processes (EFs) attend to meaningful information (i.e. believed relevant to overall goals) held in WM. Such a perspective is in accordance with one of the central observations of this study, in that cognitive abilities such as the CE and fluid-intelligence, as well as the overall efficiency of these systems, have been associated with success in scene based analogies-providing important implications for both the modelling of the reasoning process, as well as the perceived role of WMC.

It is speculated that the mediatory effect of Mr.X reported in Experiments 2 and 3 may fall into one of these facilitatory categories- the task possibly being more demanding than other WM measures in terms of its requirements for faster, more efficient processes- but this by no means certain.

It is also conceivable that Mr.X is a more sensitive measurement of WMC: potentially being more demanding than other WM tasks and therefore more likely to discriminate smaller differences in a child's WMC. Further research is recommended, particularly in determining whether the

contribution of Mr.X to relational responding is predominantly due to a global WMC resource, or a visuo-spatial processing capacity.

Certainly the concept that visual storage and processing limits constrain visual analogical problems is enticing, however given the current findings it is considered more probable that Mr.X possesses unique functions also associated with the resolution of scene based analogies, rather than the ability to maintain *and* manipulate information.

Regarding functionality, although processing speed and verbal fluency have been consistently shown to be associated with analogical success, it is not argued that these factors are the only predictors of AR. Despite both almost certainly representing the speed at which executive processes can operate (facilitating complex thought and reducing the loading of the VSSP and PL by reducing the time information has to be maintained) as well as the strength of bindings separating semantic concepts (Hummel & Holyoak, 1997, 2003), efficiency components are alone unlikely to represent the decision making process underlying processing. There is supporting evidence to suggest that while important to the overall functioning of the CE and integral to reasoning (Baddeley, 2008), speed alone is insufficient to explain most deductive processes (Bayliss, Jarrold, Baddeley, Gunn, & Leigh, 2005; Colom et al., 2004; Kyllonen, & Stephens, 1990).

However, such factors are likely to contribute towards the concept of the hypothesized constraints imposed by a maturationally developing capacity

as described by Halford (1992, 1993, 1998): Halford specifically indicating that developmental increases in processing speed may be one of the factors involved in this constraint.

Despite this prediction, Richland et al. (2004, 2006, 2010) and Morrison et al. (2010) have recently hypothesised that the capacity limitations may also be defined by inhibitory skill. Yet the current research has been unable to find evidence to support such a suggestion, it being argued that the childhood development of processing efficiency may potentially equally account for Richland et al.'s (2004, 2006) observation that the presence of a distracter object reduces relational responding. Cognitive factors beyond inhibition possibly underlying the decision making process (such as those determining relevance) rather than the dismissal of prepotent responses.

Although a traditional storage *and* processing approach to WMC is arguably not the most appropriate framework to explain relational responding, other methodologies have been more productive: particularly those centring on executive functionality of the WM model.

For example, viewing the WM concept of capacity as defined by Halford (1998) as a set of constraints imposed by faculties which aid decision making and planning/strategy development, has allowed us to accommodate both Goswami's early (1992) "facilitation gradient" hypothesis as well as Gentner's concept of late-stage processing. Thus, the passive representation of similarity (when analogies have been successfully represented in short-

term memory) has been side-lined in favour of more active executive decision making once representation has taken place.

It is hoped that future research will be able to further this insight, by more clearly defining the roles component processes of the CE may or may not play in AR.

4.1.1 Critical evaluation of the RPA. Although correlations between relational responding and WMC were obtained, no interactions between WM measures and complexity/distraction were ever reported in Experiments 2 and 3. This lack of data may be interpreted by both relational primacy and complexity-constraint theories which suggest that children (of the ages tested here) possess the cognitive ability comprehend the demands of the task (Goswami, 1992) or represent the binary or ternary relation used (Halford, 1992, 1993, 1998). However, it does not explain why complexity effects have been consistently reported in age ranges above 10-11 years using other analogical paradigms such as the people-piece-analogies[62] (Cho et al., 2007; Morrison et al., 2001; Viskontas et al., 2004), nor does it explain the observation from this series of experiments that the

---

[62] As discussed in chapter 1, the concept of complexity in the PPA is hypothetically the same as the RPA, in that the number of relationships attended to defines complexity. However, unlike the RPA the PPA is a classical paradigm, and an increased number of descriptive relationships in the mapping of terms is viewed by Morrison et al (2001) to make the analogy easier. A line is therefore drawn between actual complexity of the problem (how many similar dimensions the terms share) and the number of items being processed (the number of relationships required to be attended); complex analogies being defined by those with low actual complexity, and high attended complexity (requiring the participant to select the correct relationship from a number of candidates according to experimental demand. This raises an important question as to whether the two views of complexity are compatible.

featural score in the RPA is reduced under conditions of complexity, and not increased as predicted by Waltz et al. (2000).

One possible explanation for this is that complexity, as presented in the RPA, may not load WMC enough for WMC-complexity interactions to be observable; a claim similar to the one made by Morrison, Holyoak and Truong (2001) regarding the original Markman and Gentner (1993) scene-based cross-mapping paradigm.

Establishing that processing speed and efficiency may be a possible candidates for the maturational constraint effect indicated by Richland et al. (2004, 2006) and Halford (1998) and not inhibition (see chapter 1), leaves the important question unanswered as to whether the RPA was indeed measuring what was originally intended. For this reason a critical evaluation of the RPA is necessary.

Repeated measures design. When utilizing correlational methodologies, such as those in Experiments 2, 3 and 6 which associated WMC with AR, it is important to consider alternative ways in which the primary variable may interact with the experiment and experimental environment. In the RPA one possible WM factor not associated with the primary manipulating of relational objects is remembering the relational object you chose last. In the RPA there are 20 objects with 4 conditions, all of which use the same images in more or less the same locations. This means that if a child does the complete RPA in one sitting they will encounter each problem four

times.   If a child does this over the course of one or more sessions, those who have higher WMC might perform better, and crucially negate any complexity factor in favour of previous response recall.   The RPA might then become a test of remembering previous responses and not complexity. The alternative is to conduct the RPA as a between groups paradigm, which has already been shown to significantly reduce the statistical power of the task (especially in featural analyses) or to leave ample time between sessions.   An easier method might be to introduce more problems than the current 20 using the between conditions design.

Despite these recommendations it should be noted that a visuo-spatial STS (in this case representing the ability to remember location and type of objects) was not reported in Experiment 3.   However, the ability to hold these objects in mind against interference may have been indicated through the Mr.X effect (a visuo-spatial WMC measure).   It is suggested that performance comparisons between repeated and between measures designs be conducted to determine any possible differences in relational or featural scoring.
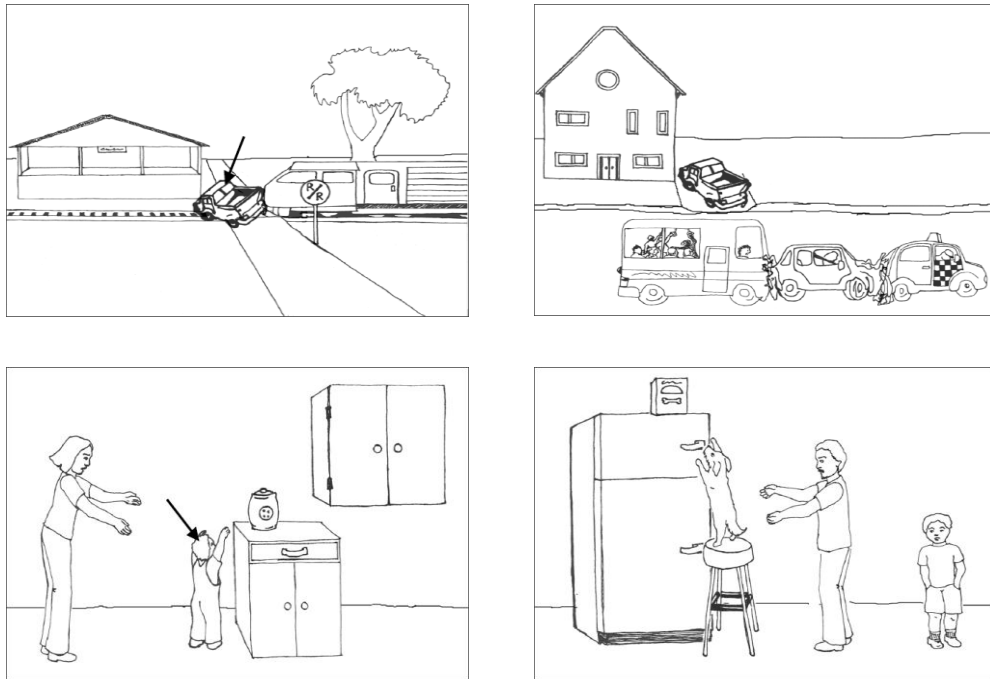
Figure 21. Same or Different? Base (left) and Target (right) scenes showing an almost identical featural object (truck) on the top and a non-identical 'transformed' featural object (boy) on the bottom. It is anticipated participants might have a harder timer relating the boy as the same boy in the base scene.

Reduced distraction. It is uncertain what role distraction is playing in the RPA. Many featural distracters represent conceptual binding, that is so say that they are not identical to the base query item and have been 'transformed' into more relational forms of response (i.e. "cat walking" rather than "cat sitting"). Complicating the desired distraction effect, it is not always clear that the item in the target is the same as the item in the base (see Figure 21). A proportion of featural objects are identical. Oberauer (2005) suggested that base objects that are held in mind may require

279

inhibition if they are displayed in the target scene, so it is logical to assume that the inhibitory demand of a problem may be being increased by making it more like the target object. Further research is needed in order to deduce the inhibitory demands of certain types of distracter item.

Reduced complexity. According to Halford (1992, 1993, 1998) as well as Gentner's taxonomy of relations (Gentner et al., 1993), a ternary relation is defined by the presence of an interacting relationship consisting of three arguments. In the RPA this appears as two extra objects on either side of the relation one. Unfortunately because an end object in the ternary relationship can only ever have one relational argument (chased *or* chasing) this means that the response is always the middle of the relationship. There is some concern that chunking the term "being chased by *and* chasing" may be reduced into "middle", reducing the WMC load for AR. This load may be further reduced if the child works out that the solution is always the middle response, thus making complex conditions easier. Hence, WMC may be correlated with the ability to deduce the task demands from a few simple trials, without the need for further processing.
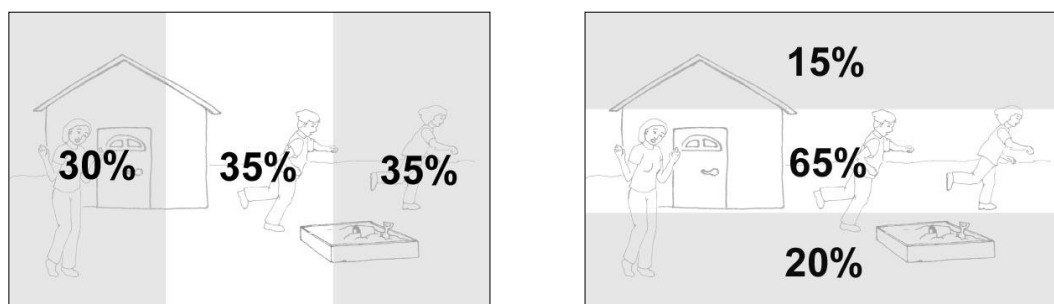
Figure 22. Location of relational objects in the RPA as a percentage.

Signposting. One possible reason for the reduced variance in the complex conditions and a high proportion of relational responses is that relational responses were signposted. Already discussed in detail in chapter 2, eye tracking experiments have shown that meaningful interactions on a page fixate attention more than isolated objects (Henderson et al, 1999). As relational objects attract longer cycles of attention in the RPA (Gordon & Moser, in press), this makes it more likely that a relational response may be selected. A further issue is that complex relationships are always presented in 'threes' (compared to 'twos') taking up a large proportion of space on the page, thus increasing the likeliness that they are the desirable form of response (see potential distracters below).

Potential distractors. Although it was claimed that the size and location of the 5 objects in each scene was controlled (Richland et al., 2004), some objects are clearly disproportionately represented, no consideration appearing to have been given to object size, appearance or location (see below). Whilst a larger, unusual or more central relational object (for

example) may or may not necessarily be the correct relational solution to the problem the combination of any or all of the factors mentioned here may have made certain objects more or less noticeable (i.e. a distracter) resulting in some problems being unintentionally easier or harder.

Control of location.   Although the RPA has a good left-right distribution of targets, the majority of relational responses are presented in the middle of the page (see Figure 22 above), potentially priming a child where to look. During testing, children frequently responded, with "it's always in the centre," (referring to the item location not the relational objects location in the ternary pattern).   Obviously, having the relational responses signposted in this manner (see complexity) may reduce task difficulty. Conversely, moving the featural object to a similar location within the target scene as it inhabited in the base may increase its distraction effect.

Control of object size.   Some objects in the RPA are observably larger than others, which may lead them to be recognised more quickly.   Eye-tracking experiments have previously shown that larger objects are more likely to fix attention, and are easier to locate at high speeds (Djamasbi et al., 2011).

Figure 23. Example of a complex image from the RPA with poorly defined, over-complex visual boundaries

Control of appearance. Eye-tracking experiments by Henderson and Ferreira (2004) showed that items of interest were more likely to attract attention. Some objects in the RPA task stood out, many were animals, leading to children to comment on the number of rabbits and dogs in the paradigm (the dog was often the solution or the base query item), and some were unorthodox (a man with no hair feeding a bird was commented upon frequently). Any task dealing with potential effects of inhibition would need to reduce the background semantic 'noise' of such stimuli to make the effects of experimental stimuli (featural responses) stand out.

Drawing style. Whilst it may appear trivial to criticise a design because it is hand drawn or over-complex (see Figure 23), both in fact make recognition and selection difficult whilst also making it harder for the researcher to see what is precisely being pointed at (i.e. is the brush being pointed to, or the clown-doll?). It is important in today's age of computer aided design that such factors are presented with clear precise designs, or in the case of computerized tasks: more forced choice (i.e. highlighting the entire object

283

and not just the sleeve) in order to remove the possibility of guesses or ambiguous responses which are concealed by the finger (i.e. pointing at the clowns brush rather than the clown). An alternative way to alleviate this is for the child to verbally state what they pointed at after selection (this may load WM further so a control study using the original methodology may have to be used)

4.2 Directions for future research

Manipulation of the similarity constraint. Quasi analogies have been around for some time (Levinson & Carpenter, 1974). However at the time of writing this paradigm has not been applied to scene based problems. By design the RPA does not account for the recognition process underlying the similarity-constraint (relational primacy theory), but highlights the query object in the base scene with an arrow and relying on the child's ability to maintain the instruction "See? they are the same but look different." As Richland herself recognizes (Richland et al., 2006) the word "the same" can be considered ambiguous.

One way to reduce task ambiguity is to give the child a verbal analogy problem before the presentation of the base. i.e. "what is being chased and chasing?" When the child has identified the correct item, the child would then be directed to find the same similarity in the target scene. Theoretically (see Richland et al., 2004, 2006) this would not make the task any easier, as the same distractions/levels of complexity would still exist.

Richland et al. (2006) argues that children doing the RPA always knew the similarity-constraint. However, it is predicted that that by making the similarity-constraint explicit in each scene (i.e. removing the relational primacy aspect altogether) performance might be increased.

Increase potential distraction load in the RPA. Following the guidelines suggested above, it should be possible to construct a new version of the RPA. This may help us to better comprehend the overall load the current task places on the reasoning processes and allow individual variables, such as increased size, location, detail, type etc., to be more accurately manipulated in order to determine their potential demands on processing resources.

One possible example of how further research may be conducted is to compare the inhibitory demands and RT's associated with transformed featural (i.e. a cat sitting in the base scene compared to a cat walking/different type of cat in the target) compared to non-transformed (using the same, identical image in base and target scenes) featural objects. It is hypothesized that objects held in mind which are closer in appearance to the image of the base featural object might place greater emphasis on the inhibitory system than objects which share a semantic concept, but are visually dissimilar. The reasoning being that that there would be greater level of interference for more similar objects simply because the featural object is concurrent, whereas a non-visible relationship has to be retrieved via a developed strategy.

Understanding how similarity may affect inhibitory processes would be of significant interest for future research.

Eye tracking.   Gordon and Moser's (2007) eye-tracking paradigm has provided a useful platform from which to understand the value of individual responses in the RPA; particularly how the comparison process works. However, the paradigm has limited explanatory value for the current research in that it was both centred around an older (undergraduate) population and the whole analogical process (saccades and the movement from target to base rather than focus on individual relational objects) rather than which objects act as a fixation point. A comparison between such points in older and younger children, highlighting any changes which may be present might help us understand the reasoning differences in maturation reported in this thesis, whilst also allowing further insight into the appeal of distracter and other forms of objects mentioned above.

Comparison of analogical tasks.   One of the biggest failings in AR has been the lack of best practice in the field, particularly in terminology and understanding of the theoretical 'glue' that binds different analogical tasks together.   It is suggested that in order to test the concept of analogical skill, WMC/STS comparisons between a range of widely used analogical tasks should be conducted to establish if similar operations are involved in different forms of task (such as the People Piece Analogies, or forms of classical or problem analogies), and whether analogical ability is transferrable between measures.

Repeat dual task measures. It is suggested that the dual task methodology of Experiment 5 be repeated, with different forms of visuo-spatial, verbal and executive interference (i.e. instead of RNG, which children found too hard, verbal fluency tasks such as "name as many animals as possible"). Ascertaining which dual-task might best represent the executive process or processes shared by AR (if any) would provide important insight into the functions involved in the resolution of the RPA.

Test older and younger children. Richland et al. (2004, 2006) examined a wider age range in her original experiments. Although children younger than 5 are difficult to test effectively due to issues surrounding experimental understanding, simpler WM and executive measures might still be employed to determine a child's ability. Possibly answering questions as to whether individual differences in WMC may constrain relational responding (as predicted by Halford, 1998), or whether the same factors exist as has been observed in their older peers (i.e. fluid intelligence mediates relational responding).

Use additional WMC measures. Expanding the range of executive tasks (to include Executive VSSP and PL measures outside of the AWMA) would increase our understanding of how the modules of WMC interact with aspects of the RPA such as complexity or distraction. As mentioned, the Mr.X effect reported since Experiment 2, combined with the lack of other WM data from other measures suggests that Mr.X possesses certain

attributes necessary for the solution of scene based problems. It may be that other WM measures from other batteries are able to better establish the involvement of visuo-spatial or phonological capacities (executive or STS). In the current research the potential involvement of other WMC measures (other than Mr.X) was indicated in Experiment 3, however the relationship between the other WMC measure (listening recall task) and relational responding failed to achieve significance.

In order to further test the hypothesis that Mr.X is either particularly sensitive to WMC due to its high executive and storage demands, and/or that it requires specific executive processes in common with that of the RPA: the addition of further WM measures is recommended, specifically ones that are equally (or more) demanding to children of the age ranges tested.   Such research could also provide additional evidence for whether the capacity demonstrated by Mr.X is visuo-spatial specific or more global in nature (as has been hypothesized here).

References

Ackerman, P. Beier, M. & Boyle, M. (2005). Working memory and intelligence: the same or different constructs? *Psychological Bulletin*, *131*(1), 30–60.

Alexander, P. Willson, V. White, C. & Fuqua J. (1987). Analogical reasoning in young children. *Journal of Educational Psychology*, *79*(4), 401-408.

Allen, R. Baddeley, A., & Hitch, G. (2006). Is the binding of visual features in working memory resource demanding? *Journal of Experimental Psychology, 135* (2), 298-313.

Allen, R., Hitch, G., & Baddeley, A. (2009). Cross-modal binding and working memory. *Visual Cognition, 17*(1), 83-102.

Alloway, T. (2008). Working memory, but not IQ, predicts subsequent learning difficulties in children with learning difficulties. *European Journal of Psychological Assessment, 25* (2), 92-98.

Alloway, T., (2007). *Automated working memory assessment.* London: Pearson Assessment.

Alloway, T. & Gathercole, S. (2006). How does working memory work in the classroom? *Educational Research and Reviews, 1* (4), 314-319.

Alloway, T & Alloway, R. (2008). *Working memory: is it the new iq?* [Online] Available from Nature Proceedings: http://precedings.nature.com/documents/2343/version/1/files/npre20082343 -1.pdf [Accessed March 7th, 2011]

Alloway, T., Gathercole, S., Willis, C. & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in early childhood. *Journal of Experimental Child Psychology*, *87*(2), 85-106.

Alloway, T., Gathercole, S., Adams, A., Willis, C., Eaglen, R., & Lamont, E. (2005). Working memory and other cognitive skills as predictors of progress towards early learning goals at school entry. *British Journal of Developmental Psychology*, *23* (3), 417-426.

Alloway, T. Gathercole, S. Kirkwood, H & Elliott, J. (2008). Evaluating the validity of the automated working memory assessment. *Educational Psychology, 28* (7), 725–734.

Alloway, A. & Passolunghi, M. (2010). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*, 133-137.

Andrews, G., & Halford, G. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, *45* (2), 153–219.

Arbuckle, J. (2007). *Amos 7.0 build 1140*. Amos Development Corporation.

Ashworth, J. (2008). "Medieval theories of analogy", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)* Zalta, E. (ed.) [Online] Available at http://plato.stanford.edu/archives/fall2008/entries/analogy-medieval/ [Accessed Jan 10th 2009]

Awh, E., Barton, B. & Vogel, E. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18* (7), 622-8.

Baddeley, A., Bressi, S., Della Sala, S., Logie, R. & Spinnler, H. (1991). The decline of working memory in alzheimer's disease: a longitudinal study. *Brain*, *114*, 2521–2542.

Baddeley, A. (1995). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (1966). The capacity for generating information by randomisation. *Quarterly Journal of Experimental Psychology*, *18*(2), 119-129.

Baddeley, A. (1996b). Exploring the central executive. *The Quarterly*

Journal of Experimental Psychology,

http://www.informaworld.com/smpp/title~content=t713683590~db=jour~ta

b=issueslist~branches=49 - v49*49*(1), 5-28.

Baddeley, A. (2000). The episodic buffer: a new component of working

memory? *Trends in Cognitive Science, 4*, 417-423.

Baddeley, A. (2007). *Working memory, thought and action.* Oxford: Oxford

University Press.

Baddeley, A. & Hitch, G. (1974). Working Memory. In Bower (Ed), *recent*

advances in learning and motivation. Vol. VIII (7-90) New York: Academic

Press.

Badre D, Poldrack R, Pare-Blagoev E, Insler R, Wagner A. (2005).

Dissociable controlled retrieval and generalized selection mechanisms in

ventrolateral prefrontal cortex. *Neuron 47* (6), 907-918.

Barrouillet, P. (1996). Transitive inferences from set‐inclusion relations

and working memory. *Journal of Experimental Psychology: Learning,*

Memory, and Cognition 22, 1408–1422.

Bayliss, D., Jarrold, C., Baddeley, A., Gunn, D., & Leigh, E. (2005). Mapping the developmental constraints on working memory span peformance. *Developmental Psychology,* 41, 579-597.

Beier, M. & Ackerman, P. (2005). Working memory and intelligence: different constructs. Reply to Oberauer et al. (2005) and Kane et al. (2005). *Psychological Bulletin*, *131*(1), 72–75.

Bishop, D., Aamodt-Leaper, G., Creswell, C., McGurk, R., & Skuse, D. (2001). Individual differences in cognitive planning on the Tower of Hanoi task: Neuro psychological maturity or measurement error*? Journal of Child Psychology and Psychiatry*, *42*, 551–556.

Borkowski, J. Benton, A. & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia*, *5*(2), 135-140.

Bryan, J., Luszcz, M., Crawford, J. (1997). Verbal knowledge and speed of information processing as mediators of age differences in verbal fluency performance among older adults. *Psychology and Aging, 12*(3), 473-478.

Brewin, C. R., & A. Beaton. (2002). Thought suppression, intelligence, and working memory capacity. *Behaviour Research and Therapy*, *40*, 923–930.

293

Brown, A. (1989). Analogical learning and transfer: What develops? In

Vosniadou, S & Ortony, A. (Eds.) *Similarity and Analogical Reasoning* ,

369-412. Cambridge, MA: Cambridge university press.

Brown, A. & Kane, M. (1988). Preschool children can learn to transfer:

learning to learn and learning from example. *Cognitive psychology, 20*(4),

493-523.

Brown, A., Brockmole, R., Gow, A. & Deary, I. (2012). Processing speed

and visuo-spatial executive function predict visual working memory ability

on older adults. *Experimental Aging Research, 38 (1),* 1-19.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of

children's mathematics ability: inhibition, task switching, and working

memory. *Developmental Neuropsychology, 19*, 273–293.

Bungel, S., Wendelken, C., Badre, D. & Wagner, A. (2004). Analogical

reasoning and prefrontal cortex: evidence for separable retrieval and

integration mechanisms. *Cerebral Cortex, 15* (3), 239-249.

Campbell, R. (2000). Reflection abstraction in context. In Campbell, R,

(ed). *Studies in reflecting abstraction.* Hove: Psychology Press

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the raven progressive matrices test. *Psychological Review, 97,* 404–431.

Case, R., Kurland, D. & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span, *Journal of Experimental Child Psychology, 33,* 386-404.

Cattell, R. (1963). Theory of fluid and crystal intelligence: a critical experiment. *Journal of Educational Psychology, 54*(1), 1-22.

Chen, Z. (1999). Schema induction in children's analogical problem solving. *Journal of Educational Psychology, 91*(4), 703-715.

Chen Z.& Daehler, M. (1989). Positive and negative transfer in analogical problem solving. *Cognitive Development, 4,* 327-344.

Chiappe, P., Siegel, L., & Hasher, L. (2002). Working memory, inhibition and reading skill. In Shohov, S. (Ed.) *Advances in Psychology Research*, *vol. 9* (30-51). Huntington, NY: Nova science.

Cho, S., Holyoak, K., & Cannon, T. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory and Cognition*, *5*(6), 1445-1455.

Cho, S., Moody, T.D., Fernandino, L., Mumford, J.A., Poldrack, R.A., Cannon, T.D., Knowlton, B.J., Holyoak, K.J., (2010). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. *Cerebral Cortex,* 20, 524–533.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J., Holyoak, K., & Gabrieli, J. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimaging*, *14*, 1136–1149.

Chuah, Y., & Maybery, M. (1999).Verbal and visuo spatial short-term memory: common sources of developmental change? *Journal of Experimental Child Psychology. 73*(1), 5-44.

Chuderska, A., & Chuderski, A. (2009). Executive control in analogical reasoning: Beyond interference resolution. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (1758-1763). Austin, TX: Cognitive Science Society

Clement, C. & Gentner, D. (1988). Systematicity as a selection constraint in analogical mapping. *Proceedings of the tenth annual conference of the cognitive science society*, 412-418.

Cohen, J. (1977) *Statistical Power Analysis for the Behavioural Sciences*, New York: Academic Press.

Col, R., France, B., Taylor, I. (2005). The role of models/and analogies in science education: implications from research. *International Journal of Science Education, 27*(2), 183-198.

Collins, A. & Burstein, M. (1989). Afterword: a framework for a theory of comparison and mapping. In Vosniadou, S. & Ortony, A. (eds.) *Similarity and analogical reasoning* (546-565) Cambridge: Cambridge University Press.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa & Kyllonen, P (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, *32*, 277-296.

Colom, R., Rebollo, I., Abad, F., & Shih, P. C. (2006). Simple span tasks, complex span tasks. and cognitive abilities: a re-analysis of key studies. *Memory and Cognition*, *34*, 158−171.

Colom, R., Jung, R. & Haier, R. (2007). General intelligence and memory span: evidence for a common neuroanatomic framework. *Cognitive Neuropsychology*, *24*(8), 867-878.

Conway, A. & Engle, R. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology, 123* (4), 354-373.

Crisafi, M. & Brown, A. (1986). Analogical transfer in very young children: combining two separately learned solutions to reach a goal. *Child Development*, *57* (4), 953-968.

Darling, S., Sala, S., Logie, R., Cantagallo, A. (2006). Neuropsychological evidence for separating components of visuo-spatial working memory. *Journal of Neuropsychology*, *253*, 176–180.

Darling, S., Sergio, S. & Logie, R. (2009). Dissociation between appearance and location within visuo-spatial working memory. *The Quarterly Journal of Experimental Psychology, 62 (3)*, 417-425.

Davidson, C., Amso, D., Anderson, C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4-13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037- 2078.

Delaney, D. & Maxwell, S. (1981). On using analysis of covariance in Repeated measures designs. *Multivariate Behavioural Research*, *16,* 105-123.

DeLoache, J., Miller, K., & Pierroutsaks, S. (1998). Reasoning and problem solving. In Damon, W., Kuhn, D. and Siegler, R. (Eds.) *Handbook of Child Psychology: vol. 2: Cognition, perception and language* (801–850). New York: Wiley.

Dempster, F. & Brainerd, C. (1995*). Interference and inhibition in cognition*. San Diego, CA: Academic press.

Dejong, G (1989). The role of explanation in analogy; or, the curse of an alluring name. In Vosniadou, S. & Ortony, A. (Eds.) *Similarity and analogical reasoning.* (346-365) Cambridge: Cambridge University Press.

Dejong, R., Coles, M., Logan, G. & Gratton, G. (1990). In search of the point of no return: the control of response processes. *Journal of experimental psychology: Human Perception and Performance. 16*(1), 164-182.

De Neys, W. (2006[a]). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*(5), 428-433.

De Neys, W. (2006[b]). Automatic-heuristic and executive-analytic processing in reasoning: chronometric and dual task considerations. *Quarterly Journal of Experimental Psychology*, *59*, 1070-1100.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: retrieval and inhibition of stored counterexamples. *Thinking & Reasoning*, *11*, 349-381.

De Neys, W. & Verschueren, N. (2006[b]). Working memory capacity and a notorious brain teaser: The case of the monty hall dilemma. *Experimental Psychology,* 53(2) 123-131.

De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: when our brains detect that we are biased. *Psychological Science*, *19*(5), 483-489.

De Neys, W., & Van Gelder, E. (2008). Logic and belief across the life span: the rise and fall of belief inhibition during syllogistic reasoning. *Developmental Science*, *12*, 123-130.

DeStefano, D. & LeFevre, J. (2003). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*(3), 353-386.

Diamond, A., Kirkham, N., Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a proponent response. *Developmental Psychology, 38*(3), 353-62.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.) *Principles of frontal lobe function* (466–503). London: Oxford University Press.

Djamasbi, S., Siegel, M., & Tullis, T. (2011). Visual hierarchy and viewing behavior: an eye tracking study. *In proceedings of HCI* (1). 331-340.

Donnelly, C. & McDaniel, M. (1993). Use of analogy in learning new scientific concepts. *Journal of Experimental Psychology: Learning, memory and cognition, 19*(4), 975-987.

Dumontheil, I., Houlton, R., Christoff, K. & Blackmore, S. (2010). Development of relational reasoning during adolescence. *Developmental science*, *13*(6), 15-24.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58(5) Whole no 270.

Engle, R (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19-23

Engle, R. (2010). The role of working-memory capacity in cognitive control. *Current Anthropology, 51* (1), 17-26.

Engle, R., Tuholski, S., Laughlin, J & Conway, A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331.

Engle, R., & Kane, M. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.) *Psychology of Learning and Motivation: Advances in research and theory,* 145–199.

Engle, R., Cantor, J. & Carullo, J. (1992). Individual differences in working memory and comprehension: a test of four hypothesis. *Journal of Experimental Psychology: Learning, memory and cognition, 18,* 972-992.

Evans, T. (1968). A program for the solution of a class of geometric analogy intelligence test questions. In Minsky, M. (ed) *Semantic information processing.* (271-353). Cambridge, MA: MIT press.

Falkenhainer, B., Forbus, K & Gentner, D (1989). The Structure mapping engine: algorithm and examples. *Artificial intelligence, 41*(1), 1-63.

Flanagan, D. & Harrison, P. (2005). *Contemporary intellectual assessment: Theories, tests, and issues.* New York: Guilford press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Flanagan, D. & Kaufman, A. (2004). *Essentials of WISC-iv assessment.* Hoboken: New Jersey, John wiley and Sons.

Forbus, K. (2001). Exploring analogy in the large. In Gentner, D., Holyoak, K., & Kokinov, B. (Eds.) *The analogical mind: perspectives from cognitive science* (23-58). Cambridge MA: MIT press.

Forbus, K., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*(2), 141-205.

Freid, M. (2003). The use of analogy in book VII of Apollonius' conica. *Science in Context, 16*(3), 349-365.

Friedman, N.P., & Miyake, A. (2004). The relations among inhibition and interference cognitive functions: a latent variable analysis. *Journal of Experimental Psychology*, *133*, 101–135.

French, R. (2002). The computational modeling of analogy making. *Trends in Cognitive Sciences, 6*(5), 200-205.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237-241.

Fry, A & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54*(1) 1–34.

Gallagher, J., & Wright, R. (1979). Piaget and the study of analogy: Structural analysis of items. In Magary. J. (Ed) *Piaget and the helping professions*. *Volume 8*, 114-119.

Gathercole, S., Pickering, S., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2) 177-190.

Gathercole, S., & Baddeley, A. (2003). *Working Memory and Language*. New York: Psychology Press.

Gathercole, S., Lamont, E., & Alloway, T. (2006). Working memory in the classroom. In Pickering, S. (Ed.). *Working memory and education*, 219 - 240. Oxford: Elsevier.

Gentile, J., Tedesco-Stratton, L., Davis, E., Lund, N. & Agunanne, B. (1977). Associative responding versus analogical reasoning by children. *Intelligence, 1*(4), 369-380.

Gentner, D. (1977a). If a tree had a knee, where would it be? children's performance on simple spatial metaphors. *Papers and reports on child language development*, *13*, 157-164.

Gentner, D. (1977b). Children's performance on a spatial analogies task. *Child Development*, *48*(3), 1034-1039.

Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cognitive Science, 7*(2), 155-170.

Gentner, D. (1986). Analogical inference and analogical access, in Preiditis, A . (Ed .) *Analogica: proceedings of the first workshop on analogical reasoning.* London: Pitman.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development*, *59* (1), 47-59.

Gentner, D. (1989). The mechanisms of analogical learning. In Vosniadou S, & Ortony, A. (eds.) *Similarity and analogical reasoning* (199-241) Cambridge: Cambridge University Press.

Gentner, D. & Ratterman (1991). Language and the career of similarity. In Gelman,S., & Byrnes, J. (Eds.) *Perspectives on thought and language: Interrelations in developmen*t, 225-277. Cambridge: Cambridge University Press.

Gentner, D., Rattermann, M., & Forbus, K. (1993). The roles of similarity in transfer: Separating retrievability and inferential soundness. *Cognitive Psychology*, *25*(4), 524-575.

Gentner, D. & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science,10*, 277–300.

Gentner, D., Holyoak, K., & Boicho, N. (2001). *The analogical mind: Perspectives from cognitive science* (199-253). Cambridge MA: MIT press.

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In Gentner, D., Holyoak, K., & Kokinov, B. (Eds.) *The analogical mind: perspectives from cognitive science* (199-253). Cambridge MA: MIT press.

Gick, M., & Holyoak, K. (1980). Analogical problem solving. *Cognitive Psychology, 12*(3), 306-355.

Gick, M., & Holyoak, K. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), l-38.

Gilbert, S., & Burgess, P. (2007). Executive function. *Current Biology, 18*(3), 110-114.

Gilhooly, K., Logie, R., Wetherick, N., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition*, *21*, 115-124.

Gilhooly, K.(1998). Working Memory, Strategies and Reasoning tasks. In. Logie, R. & Gilhooly, K. (Eds) *Working Memory and Thinking* (7-43). Hove: Psychology press.

Glynn, S. (2007). Methods and strategies: the teaching-with-analogies model. *Science and Children*, *44*(8), 52-55.

Glynn, S. (1995). Conceptual bridges: using analogies to explain scientific concepts. *The science teacher*, *62*(9), 25-27.

Glynn, S. & Takahashi, T. (1998). Learning from analogy-enhanced science text. *Journal of research in science teaching, 35*(10), 1129–1149.

Gnys, J. & Willis, W. (1997). Validation of executive function tasks with young children. *Developmental neuropsychology*, *7*(4), 487-501.

Goel V. & Grafman J (1995). Are the frontal lobes implicated in "planning" functions? Interpreting data from the tower of Hanoi. *Neuropsychologia*, *33*(5), 623-642.

Goldstein, G & Hersen M (2000). *Handbook of psychological assessment (3rd ed.)* Oxford: Pergamon.

Goldstone, R. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, memory, and cognition, 20*(1), 3-28.

Goldman, S., Pellegrino, J., Parseghian, P., & Sallis, R. (1982). Developmental and individual differences in verbal analogical reasoning. *Child Development*, *53*, 550-559.

Gordon, P, Moser, S. (2007). Insight into analogies: evidence from eye movements. *Visual Cognition*,*15*, 20-35.

Goswami, U. (1992). *Analogical reasoning in Children.* Hove: Erlbaum

Goswami, U (1994). Phonological skills, analogies, and reading Development. *Reading*, *28*(2), 32–37.

Goswami, U. (1996). Analogical reasoning and cognitive development. In Reese, H. (Ed.) *Advances in child development and behaviour* (92-135). San Diego, CA: Academic press.

Goswami, U. & Brown, A. (1989). Melting chocolate and melting snowmen: analogical reasoning and causal relations. *Cognition*, *35*(1), 69-95.

Halford, G. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development*, *35* (4), 193-217.

Halford, G. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.

Halford, G. (1998). Complexity of relations. In Logie R. & Gilhooly, K. (Eds.) *Working memory and thinking,* 139-157. Hove: Psychology press.

Halford, G., Maybery, M., O'Hare, A. & Grant, P. (1994). The development of memory and processing capacity. *Child Development*, *65*(5), 1338-1356.

Halford, G., Wilson, W., Guo, J., Gayler, R., Wiles, J., & Stewart, J. (1994). Connectionist implications for processing capacity limitations in analogies. In Holyoak K., &. Bamden, J. (Eds.) *Advances in connectionist and neural computation theory: vol. 2. Analogical connections* (363-415). Norwood, NJ: Ablex.

Halford, G., Andrews, G., Dalton, C., Boag, C., & Zielinski, T. (2002). Young children's performance on the balance scale: The influence of relational complexity. *Journal of Experimental Child Psychology*, *81*(4), 383–416.

Hallyn, F. (2000). *Metaphor and analogy in the sciences.* Netherlands: Kluwer Academic Publishers.

Hamilton, C., Coates, R.., & Heffernan, T. (2003). What develops in visuo-spatial working memory development? *European Journal of Cognitive Psychology*, *15*(1), 43-69.

Handley, S., Capon., A, Copp C, Harper C. (2002). Conditional reasoning and the Tower of Hanoi: the role of spatial and verbal working memory. *British Journal of Psychology*. 93(*4*), 501-18.

Hanley, J., Young, A., & Pearson, N (1991). Impairment of the visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology*, *43*, 101- 125.

Harcourt Assessment, Inc. 2008. Miller analogies test. [Online] http://harcourtassessment.com/haiweb/Cultures/en-US/Harcourt/Community/PostSecondary/Products/MAT/mathome.htm [accessed 17th April 2008]

Hartman A., Pickering. R. & Wilson B. (1992). Is there a central executive deficit after severe head injury? *Clinical Rehabilitation*, 6, 133–40.

Hasher, L., Zacks, T., & May, P. (1999). Inhibitory control, circadian barousal, and age. In Gopher, D. & Koriat, A. (Eds.) *Attention and performance 17: Cognitive regulation of performance: Interaction of theory and application* (653–675). Cambridge, MA: MIT press.

Hayes, D. & Tierney, R. (1982). Developing readers' knowledge through analogy. *Reading Research Quarterly*, *17*(2), 256-280.

Heinz-Martin, S., Oberauer, K., Werner W., Wittmann, W. & Schulze R. (2002). Working-memory capacity explains reasoning ability-and a little bit more. *Intelligence, 30*(3), 261–288.

Henderson, J., Weeks, P. & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental psychology: Human perception and performance*, *25*(1), 210-228.

Henderson, J. & Ferreira, F. (2004). Scene perception for psycholinguists. In Henderson J. & Ferreira, F. (Eds.). *The interface of language, vision, and action: Eye movements and the visual world*. New York, NY: Psychology Press.

Henson, R. (2001). Neural working memory. In Andrade, J. (Ed.) *Working Memory in Perspective.* 151-74. Hove: Psychology Press.

Hofstadter, D. (2001). Epilogue: analogy as the core of cognition. In Gentner, D., Holyoak, K., & Kokinov, B. (Eds.) *The analogical mind: Perspectives from cognitive science* (499-538). Cambridge MA: MIT press.

Holyoak, K. (1984). Analogical thinking and human intelligence. In
Sternberg, R. (Ed.) *Advances in the psychology of human intelligence.
Volume 2* (199–230). Hillsdale, NJ: Erlbaum.

Holyoak K., &. Bamden, J. (2001). *Advances in connectionist and neural
computation theory: vol. 2. Analogical connections* (363-415). Norwood,
NJ: Ablex.

Holyoak, K., Junn, E. & Billman, D. (1984). Development of analogical
problem-solving. *Child Development*, *55*(6), 2042-2055.

Holyoak, K. & Thagard, P. (1989). Analogical mapping by constraint
satisfaction. *Cognitive Science*, *13*(3), 295-355.

Holyoak, K., Gentner, D. & Kokinov, B. (2001). Introduction: the place of
analogy in cognition. In Gentner, D., Holyoak, K., & Kokinov, B. (Eds.)
The analogical mind: Perspectives from cognitive science (1-19).
Cambridge MA: MIT press.

Hummel, J. & Holyoak, K. (1997). Distributed representations of structure:
a Theory of analogical access and mapping. *Psychological Review*, *104*(3),
427-466.

Hummel, J. & Holyoak, K. (2003). A symbolic-connectionist theory of relational inference and generalisation. Psychological Review, *11* (2), 220-264.

Idling, M. (1997). How analogies foster learning from science texts *Instructional Science, 25*(4), 233–253.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York, NY: Basic books.

Iroise, D. Houlton, R. Kalina C. & Blakemore, S. (2010). Development of relational reasoning during adolescence *Developmental Science*, *13* (6), 15–24.

James, M. & Scharmann, L. (2007). Using analogies to improve the teaching performance of preservice teachers. *Journal of Research in Science Teaching*, *44*(4), 565-585.

Jefferies, E., Lambon Ralph, M., & Baddeley, A. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language, 51*, 623-643.

Johnson-Laird, P. & Byrne, R. (1983). *Deduction.* Hove: Erlbaum.

Johnson-Laird, P. (2001). Mental models and deduction. *Trends in Cognitive Sciences*, (5) *10*, 434-442

Kane, M., Hambrick, D., & Conway, A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66–71.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199–225.

Kaufman, A. (1994). *Intelligent testing with the WISC-iii.* New York: Wiley.

Kaufman, A. & Lichtenberger E. (2005). *Assessing adolescent and adult intelligence*, New York: Wiley.

Kaufman, S., DeYoung, C., Gray, J., Brown, J. & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, *37*(4), 374-382.

Keane, M. & Brayshaw, M. (1988). The incremental analogical machine: a computational model of analogy. In D. Sleeman (Ed.) *Third european working session on machine learning*. (53-62) San Mateo: Kauffman.

Klauer, K. C., Stegmaier, R., & Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking & Reasoning*, *3*, 9-47.

Kokinov, B. & Petrov, A. (2001). Intergrating memory and reasoning in analogy-making: the ambr model. In Gentner, D., Holyoak, K., & Kokinov, B. (Eds.). *The analogical mind: Perspectives from cognitive science* (199-253). Cambridge MA,MIT press.

Krawczyk, D., Morrison, R., Viskontas, I., Holyoak, K., Chow, T., Mendez, M., Miller, B., & Knowlton, B. (2008). Distraction during relational reasoning: the role of the prefrontal cortex in interference control. *Neuropsychologia*, *46*(7), 2020-2032

Krawczyk, D., Hanten, G., Wilde, E., Li, X., Schnelle, K., Merkley, T., Vasquez, A., Cook, L., McClelland, M., Chapman, S. & Levin, H., (2010a). Deficits in analogical reasoning in adolescents with traumatic brain injury. *Human Neuroscience*, *62*, 1–13.

Krawczyk, D., McClelland, M., Donovan, C., Tillman, G. & Maguire, M. (2010b). An fMRI investigation of cognitive stages in reasoning by analogy. *Brain Research*, 1342, 63–73.

Kyllonen, P. & Christal, R. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, *14*(4), 389-433.

Kyllonen, P., & Stephens, D. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences, 2*(*2*), 129-160.

Levine, T. & Hullet, D (2002) Eta Squared, Partial Eta Squared, and Misreporting of Effect Size in Communication Research. *Human communication research*, *28(4),* 612-625.

Levinson, P. & Carpenter, R. (1974). An analysis of analogical reasoning in children. *Child Development*, *45*(3), 857-861.

Lezak, M, Howieson, D. & Loring, D. (2004). *Neuropsychological assessment* (4th edition). Oxford: Oxford University Press.

Liddell, H. & Scott, R (1996). *A greek-english lexicon*. Revised by Jones H. &. McKenzie, R. 9th Ed. Oxford: Clarendon Press.

Logan, D. (1994). On the ability to inhibit thought and action: a user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.) *Inhibitory processes in attention, memory, and language* (189–239). San Diego, CA: Academic press.

Logan, D., & Cowan, W. (1984). On the ability to inhibit thought and action: a theory of an act of control. *Psychological Review, 91*, 295–327.

Logie, R & Gilhooly (1998). *Working Memory and Thinking*. Hove:

Psychology press.

Logie, R., & Pearson, D (1997). The inner eye and the inner scribe of visuo-

spatial working memory: evidence from developmental fractionation.

European Journal of Cognitive Psychology, 9(3), 241-257

Logie, R. (2003). Spatial and visual working memory: a mental workspace.

In: Irwin D.E., Ross B.H. (Eds.) *The psychology of learning and motivation

(volume 42): Cognitive vision*. (37-78). New York: Academic press.

Lunzer, E. (1965). Problems of formal reasoning in test situations

Monographs of the society for research in child development, *European

research in cognitive development: report of the international conference on

cognitive development, 30*(2), 19-46.

Luo,Q., Perry, C., Peng D, Jin, Z., Xu, D. & Ding,G. (2003). The neural

substrate of analogical reasoning: an fMRI study. *Cognitive Brain

Research, 17*(3), 527–534.

Lustig, C., May, C. & Hasher, L. (2001). Working memory span and the

role of proactive interference. *Journal of Experimental Psychology, 130*(2),

199-207.

MacLeod, C. (2007). The concept of inhibition in cognition. In MacLeod, C & Gorfein, D. (Eds.) *Inhibition in Cognition*, 3–23.Washington, DC: American Psychological Association.

Markman, A., Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, *25*(4), 431-467.

Martinez, M. E. (2000). *Education as the cultivation of intelligence*. Mahwah, New Jersey: Erlbaum.

Miyake, A., Friedman, N., Rettinger, D., Shah, P., Arbor, A., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology, 130*(4), 521-540.

Miyake, A. Friedman, N., Emerson, M., Witzki, A. Howerter, A. & Wager, T. (2000). The unity and Diversity of Executive Functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis, *Cognitive Psychology*, *41*(1), 49-100.

Miller, W. (1947). *Miller Analogies test.* New York: Psychological Corp.

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Minsky, M. (1988). *The Society of Mind.* New York: Simon & Schuster.

Morrison, R., Holyoak, K., & Truong, B. (2001). Working memory modularity in analogical reasoning. *Proceedings of the twenty-fourth annual conference of the cognitive science society*, 663-668. Mahwah, NJ: Erlbaum.

Morrison, R., Doumas, L, & Richland, L. (2006). The development of analogical reasoning in children: A computational account. *Proceedings of the twenty-eighth annual conference of the cognitive science society*, 603-608. Mahwah, NJ: Erlbaum.

Morrison, R., Doumas, L. & Richland L. (2010). A computational account of children's analogical reasoning: balancing inhibitory control in working memory and relational representation. *Developmental Science, 14*(3), 516-29.

Morrison, R., Krawczyk, D., Holyoak, K., Hummel, J, Chow, T., Miller, B., & Knowlton, B. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience*, *16*, 260–271.

Mulholland, T., Pellegrino, J. & Glaser, R., (1980). Components of geometric analogy solution. *Cognitive Psychology, 12*(2), 252–284.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383.

Nigg, J. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition hypothesis. *Psychological Bulletin*, *126*(2), 220-246.

Norman, D. & Shallice, T. (1986*). Attention to action: willed and automatic control of behaviour.* In Davidson, R., Schwartz, G. & Shapiro, D. (eds) *Consciousness and self-regulation: Advances in research and theory.* (1-18) New York: Plenum press.

Novick, L. (1988). Analogical transfer, problem similarity, and expertise *Journal of Experimental Psychology; Learning, memory, and cognition, 14*(3), 510-520.

Numminen, H., Lehto J, Ruoppila. I (2001). Tower of Hanoi and working memory in adult persons with intellectual disability. *Research into Developmental Disabilities*, 22(*5*), 373-87.

Oberauer, K., Schulze, R., Wilhelm, O. & Süß, H (2005). Working memory and intelligence- their correlation and their relation: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*(1), 61–65.

Oberauer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition. *Journal of Experimental Psychology, 134*(3), 368–387.

Oppenheimer, R. (1956). Analogy in science, *American Psychologist, 11*(3), 127-135

Olden, D., Thompson, R., Premack, D. (2001). Can an ape reason analogically? cinprehension and production of analogical problems by Sarah, a chimpanzee (pan troglodytes) In Gentner, D., Holyoak, K., & Kokinov, B. (Eds.) *The analogical mind: perspectives from cognitive science* (471-497). Cambridge MA: MIT press.

Parks, R., & Cardoso J. (1997). Parallel distributed processing and executive functioning: tower of hanoi neural network model in healthy controls and left frontal lobe patients international. *Journal of Neuroscience,* http://www.informaworld.com/smpp/title~db=all~content=t713644851~tab =issueslist~branches=89 - v89*89*(3), 217 - 240

Pearson Assesments for general and clinical use. (2008). General ability measure for adults [Website]
http://www.pearsonassessments.com/tests/gama.htm [accessed 17th April 2008]

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia*, *12*(3), 323 -330.

Phillips, L. & Hamilton, C. Andrade, J. (ed) (2001). The working memory model in adult aging research. In Andrede, J. (ed) *Working memory in perspective,* 101-125. Hove: Psychology Press.

Piaget, J. (2001). *L'abstraction reflechissante*. Translated from French by Campbell, R. Hove: Psychology Press. (Originally published 1977)

Piaget, J., Montangero, J. & Billeter, J (1977). The formation of analogies. In Campbell, R (ed) *Studies in reflecting abstraction.* Hove: Psychology Press.

Piaget, J. & Inhelder, B. (1969). *The Psychology of the Child*. London: RKP.

Plato (1999). The allegory of the cave. In. Cottingham, J (Ed) *Western philosophy: an anthology*, 63-70. Oxford: Blackwell publishing.

Prifitera, A., Saklofske, D., & Weiss, L.G. (2005). *WISC-IV clinical use and interpretation: scientist-practitioner perspectives (practical resources for the mental health professional).* Burlington, MA: Elsevier.

Psychological Corporation. (1992). *The Wechsler individual achievement test, manual*. San Antonio, TX.

Ratterman, M.J., & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: children's performance on a casual mapping task. *Cognitive Development*, *13*, 453-478.

Rafal, R., & Henik, A. (1994). The neurology of inhibition: integrating controlled and automatic processes. In Dagenbach, D. & Carr,T (Eds.) *Inhibitory processes in attention, memory, and language* (1-51). San Diego, CA: Academic press.

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence.* London: Lewis,

Redick, T. S., Heitz, R. P., & Engle, R. W. (2007). Working memory capacity and inhibition: Cognitive and social consequences. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition.* (pp. 125-142). Washington, DC US: American Psychological Association.

Reitman, W. (1965). *Cognition and thought: an information processing approach.* New York, NY: John Wiley and Sons.

Richland, L., Morrison, R, & Holyoak, K. (2004). Working memory and inhibition as constraints on children's development of analogical reasoning.

In Forbus, K.,Gentner, D., & Regier, T. (Eds.) *Proceedings of the twenty-sixth annual conference of the cognitive science society,*1149-1154. Mahwah, NJ: Erlbaum.

Richland, L., Morrison R., Holyoak, K. (2006). Children's development of analogical reasoning: insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*(3), 249-273.

Richland, L. & McDonough, I. (2010). Learning by analogy: Discriminating between potential analogs. *Contemporary Educational Psychology*, 35, (1), 28-43.

Richland, L., Chan, T. Morrison, R, & Au, T. (2010). Young children's analogical reasoning across cultures: Similarities and differences. *Journal of Experimental Child Psychology*, *105*, 146–153.

Roberts., Robbins T., & Weiskrantz, L.(1998). *The prefrontal cortex. Executive and cognitive functions.* Oxford: Oxford University Press.

Rogers, D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology, 124*(2), 207-231.

Rowe, G., Hasher, L., & Turcotte, L. (2010). Interference, aging, and visuospatial working memory: the role of similarity. *Neuropsychology*, *24*(6), 804-807.

Rowe, G., Hasher, L., & Turcotte. (2008). Age differences in visuospatial working memory. *Psychology and Aging*, *23* (1), 79-81.

Rutherford, E. (1911). The scattering of a and b particles by matter and the structure of the atom. *Philosophical Magazine Series, 21*(6), 669-688.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Sattler, J., & Dumont, R. (2004). *Assessment of children: WISC-IV and WPPSI-III supplement*. California: Jerome M. Sattler.

Shallice, T. (1982). Specific impairments of planning. *Royal Society of London philosophical transactions series* B, *298*, 199–209.

Shelly, C. (2003). *Multiple analogies in science and philosophy.* Amsterdam: John Benjamins Publishing Co.

Simon, H. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7(2) 268–288.

Sinclair, J. (1991). *Colins English dictionary*. Glasgow: Harper Collins.

Singer-Freeman, K. (2005). Analogical reasoning in 2-year-olds: the development of access and relational inference. *Cognitive Development*, *20*(2), 214-234.

Smith, L. (1992). *Jean Piaget: critical assessments.* London: Routledge.

Smith, E. & Jonides, J. (1997). Working Memory: A view from neuroimaging. *Cognitive Psychology, 33*(1), 5-42.

Sorel, O., Pennequin, V, (2008). Aging of the Planning process: The role of executive functioning. *Brain and Cognition*, 66(*2*), 196-201.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition,* London: Macmillan.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Spector, A, & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89* (4), 669-679

Sternberg, R. (1977). *Intelligence, information processing, and analogical reasoning: the componential analysis of human abilities.* Hillsdale, NJ: Erlbaum.

Sternberg, R. (1977). Component processes in analogical reasoning. *Psychological Review, 84*(4), 353-378.

Sternberg, R. (1982). *Handbook of human intelligence*, Oxford: Oxford University Press.

Sternberg, R., & Gardner, M. (1982). A componential interpretation of the general factor in human intelligence. In Eysenck, H. (Ed.) *A model for intelligence* (231-254). New York: Springer-Verlag.

Sternberg, R. & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Development*, *51* (1), 27-38.

D, Struss & R, Knight (2002). *Principles of frontal lobe function.* New York: Oxford University Press.

Swanson, H. & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: both executive and phonological processes are important. *Journal of Experimental Child Psychology, 79*(3), 294–321.

Swanson, H. & Beebe-Frankenberger, M. (2004). The Relationship between working memory and mathematical problem solving in children at risk and

not at risk for serious math difficulties. *Journal of Educational Psychology,*
96(3), 471–491.

Tabachnick, B. & Fidell, L. (1997). *Using multivariate statistics* (5th Ed).
Saddle River, NJ: Pearson Education/ Allyn & Bacon.

Thibaut, J.-P., French, R. M., & Vezneva, M. (2010). The development of
analogy making in children: Cognitive load and executive functions. Journal
of Experimental Child Psychology, 106(*1*), 1-19.

Thibaut, J, French, R.M., Missault, A., Gérard, Y., Glady, Y. (in press)
Proceedings of the thirty-third annual meeting of the cognitive science
society.

Theodoridis, G. (2006) *Prometheus Bound* [Online]
Available at http://bacchicstage.com/Prometheus.htm [Accessed 24th
August 2009]

Thompson, J., Hamilton, C., Gray, J., Quinn, J., Mackin, P., Young, A. &
Ferrier, I. (2006). Executive and visuospatial sketchpad resources in
euthymic bipolar disorder: implications for visuospatial working memory
architecture. *Memory*, 14(4) 437−451.

Thompson, R., Boysen,T. & Oden, L. (1997). Language-naive chimpanzees
(pan troglodytes) judge relations between relations in a conceptual
matching-to-sample task. *Journal of Experimental Psychology*, *23*(1), 31-43

Toms, M., Morris, N., & Ward, D. (1993). Working memory and conditional reasoning. Quarterly journal of experimental psychology: *Human Experimental Psychology*, *46*, 679-699.

Towse, J. & Mclachlan, A. (1999). An exploration of random generation among children. *British Journal of Developmental Psychology, 17*, 363–380

Towse, J., & Neil, D. (1998). Analyzing human random generation behavior: A review of methods used and a computer program for describing performance. *Behavior Research Methods, Instruments, & Computers*, *30*(4), 583-591.

Tulsky, D & O'Brien, A. (2008). Information processing speed in clinical populations. In: *Information processing speed in clinical populations,* 1-28. DeLuca, John (Ed.); Kalmar, Jessica H. (Ed.) Philadelphia, PA: Taylor & Francis.

Urbanial C. & Plous, S. (1997). Research randomizer [online] available at http://www.randomizer.org/about.htm [accessed 20th Sept, 2009]

Verbruggen, F., Logan, G. & Stevens, M. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods, 40(2),* 479-483.

Vallar, G., & Baddeley, A. D. (1984a). Fractionation of working memory.

Neuropsychological evidence of a phonological short-term store. *Journal of*

Verbal learning and Verbal behaviour, *23*, 51–161.

Vilkki, J., & Hoist, P. (1994). Speed and flexibility on word fluency

tasks after focal brain lesions. *Neuropsychologia*, *32*, 1257- 1262.

Viskontas, I., Morrison, R., Holyoak, K., Hummel, J. & Knowlton, B.

(2004). Relational integration, inhibition, and analogical reasoning in older

adults. *Psychology and Aging, 19* (4), 581-591.

*Vitruvius. "*De Architectura*, Book IX, paragraphs 9–12, text in English and*

Latin". university of chicago book IX, paragraphs 9–12, text in english and

latin". [online] available at

http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Vitruvius/9*.html.

[Retrieved 2007-08-30].

Vogel, E., Woodman, G., & Luck, S. (2001). Storage of features,

conjunctions, and objects in visual working memory. *Journal of*

Experimental Psychology: Human Perception and Performance, 27, 92–

114.

Vosniadou, S., & Ortony (1989). *Similarity and analogical reasoning.*

Cambridge: Cambridge University Press.

Wager, D., Miyake, A., Freidman, N., Emerson, M., Witzki, A. & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal Lobe' tasks: a latent variable analysis. *Cognitive Psychology, 41*(1), 49–100

Waltz, J., Knowlton, B., Holyoak, K., Boone, K, Mishkin, F. & De Menezes Santos, M. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10* (2), 119-125.

Waltz, J., Lau, A., Grewal, S. & Holyoak, K. (2000). The role of working memory in analogical mapping. *Memory and Cognition*, *28* (7), 1205-1212.

Watt, S. (1997). *Introduction: the theory of forms. Plato: republic*. London: Wordsworth editions.

Wechsler, D. (1991). *Wechsler intelligence scale for children, 3rd. ed. (WISC-III)*. London: Pearson Assessment.

Wechsler, D. (2005). *Wechsler individual achievement test, 2nd ed*. London: Pearson Assessment.

Weiss, L., Saklofske, D., Prifitera, Holdnack, J. (2006). WISC-IV *advanced clinical interpretation: basics and beyond (practical resources for the mental health professional)*. London: Academic press.

Welsh, M., Satterlee-Cartmell, T., & Stein, M. (1999). Towers of Hanoi and London: Contribution of working memory and inhibition to performance. *Brain and Cognition*, *41*, 231–242.

Willner, A. (1964). An experimental analysis of analogical reasoning. *Psychological Reports*, *15*, 470-494.

Wharton, C., Grafman, J., Flitman, S., Hansen, E., Brauner, J. Marks, A. & Honda, M. (2000). Toward neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology, 40*(3), 173–197.

Whitely, S. (1976). Solving verbal analogies: some cognitive components of intelligence test items. *Journal of Educational Psychology*, *68*(2), 234-242.

Zelazo, P. & Frye, D. (1998). Cognitive complexity and control: the development of executive function. *Current Directions in Psychological Science*, *7*(1), 121–126.

Zelazo, P. & Müller, U. (2002). The balance beam in the balance: reflections on rules, relational complexity, and developmental processes. *Journal of Experimental Child Psychology*, *81* (4), 458–46.

Zelazo, P., Muller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, *68,* (3, Serial No. 274).

Appendix A: Ethical procedures and consent form

Contact phase. Schools were contacted using on cold-call basis, either by email or phone. If the school was interested in participating a meeting was arranged face to face with that schools head-teacher wherein the research was explained. Consent forms were then left at the school for review and the school given no less than 24 hours to review all the documents and discuss the research with staff, before a second interview was arranged. The second interview consisted of a question and answer session with the Principle Investigator, at the end of which the school could consent to take part or decide to withdraw from the process.

Consent phase. If the school agreed to take part, the consent procedure would begin. Experiments used a two tier consent system wherein the first tear (opt out) 'Head-teacher consent' was the preferred method of recruitment. The second tear (opt in) was made available if the head-teacher preferred. In both cases letters were sent out the parents/guardians of all children involved informing them of the participation of their child's school and giving them the option to opt in/out respectively. Parents were given at least a week to return their forms, although this recruitment period normally ran between 2 weeks and a month depending on how the school wished to proceed. The consent forms used follow at the end of Appendix A. This was adapted for parents/teachers and for each experiment performed, by altering the bold text.

Data collection phase. All children were individually asked if they wanted to take part prior to testing. The only personal information taken was the child's name, age, school and class which collectively made up what became the child's Personal Information or PI. For each child their PI was then randomly associated with a unique number known as the 'Personal Identification Code' or 'PIC' which had been randomly generated prior to the experiment. Thus, the only way to identify a participant by name either post-testing or during the analysis later was the PI (enabling ethical security and un-biased anonymous analysis).

Data protection: All PI was stored on a 128-bit encrypted, password protected Excel file. This was stored in a fingerprint accessed laptop keyed to the Principle Investigator only. Only one file was made per experiment and all passwords consisted of 7 randomly generated numbers suffixed with a letter (denoting which experiment it related to). The original hardcopy paper version of the PI was then destroyed after the recruitment phase.

Backup data of the encrypted PI file was done once a month, replacing a duplicate PI file at an unspecified location in a locked cabinet. Passwords for the PI files were stored in a separate undisclosed secure location with no reference as to what they were.

Hardcopy data collected (such as WISC scores) was stored in secure locations in a locked room, and used the digital PIC system.

Consent Form

Title of Project: 'Working Memory and Analogical Reasoning'

Principle Investigator: Adam Robson. BSc (Hons), MEd.

Information Sheet

Please ensure you have read and initialled each page.

1. PURPOSE OF THE STUDY:

Your School is being asked to participate in a research study. The purpose of this study is to investigate the relationship between children's memory and their analogical reasoning ability.

Analogical reasoning is the ability to find a comparison between two dissimilar objects. It facilitates learning by allowing us to understand difficult concepts, solve novel problems, discover new facts and make scientific discoveries.

A better understanding of the processes behind analogy could determine how we comprehend our children to think, and perhaps how they might be taught. As a result, analogical reasoning is thought to be crucial in a child's development.

Your School qualifies for this study because it contains any number of children aged between 5 and 12.

---

2. DESCRIPTION OF THE RESEARCH:

This study is being performed for research purposes. It is part of the PhD Psychology programme at Durham University, England. It is intended that approximately 50 participants be enrolled for this particular phase of the project, however it is not necessary or intended that all of these be recruited from your school.

Each child participating in this study will conduct two one-to-one sessions with the experimenter lasting approximately 20-30 minutes each, depending on the individual child. Each child will conduct only one session per day, with a gap of at least 1 day in between sessions. The specific days upon which each session takes place will be at the discretion of the individual teacher involved and/or head teacher.

Sessions will consist of 2 measurements of Analogical reasoning ability, and a short battery of tasks looking at executive functions of working memory, namely: a) Inhibition b) task switching c) attentional focus and d) memory refreshing/updating.

In total the estimated contact time for each child will therefore be 40-50 minutes on average, after which the Childs participation will be considered complete.

All tasks used have been specifically designed for children. No physical measurements will be taken, and the majority of the answers given will be of a verbal nature. Some pencil based tasks will be used.

3. COSTS:

Participation in this study is entirely voluntary; no costs or reimbursements to either the individual, Parent/guardian's or the school will be incurred or appropriated through participation.

4. POTENTIAL RISKS AND DISCOMFORTS:

This Study is considered Minimal Risk by the Principle Investigator: the definition for Minimal risk being that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological tasks.

Although this study is non-invasive the Principle Investigator is aware that the potential Risks for the children participating in the study may include task-stress and/or anxiety. If this occurs, testing will stop immediately and the teacher overseeing the Childs class will be consulted, possibly resulting in the child being withdrawn from the project. The likelihood and severity of such an event are considered low.

5. POTENTIAL BENEFITS

Although there is no direct benefit to either the participants or the schools involved it is hoped that this research with further the scientific understanding underlying children's analogical reasoning ability and its development.

6. ALTERNATIVES TO PARTICIPATION / VOLUNTARY PARTICIPATION:

Participation in this study is entirely voluntary and the only alternative is not to participate. If you decide not to participate, your decision will not affect the school or any of the individuals concerned.

In addition, any new information that develops during the course of the study which might affect your decision to continue to participate will be given to you immediately. A signed copy of this consent form will be offered to you for your records as the head of the school/institution (i.e. head teacher). A simple, written summary of the project will be offered/made available to you following the project, should you desire it.

7. INCLUSION CRITERIA

This study operates an equal rights policy. Children will not be excluded on the basis of gender, disability, or ethnicity. However, children participating in this study must be able to speak English well enough in order for them to

be able to understand and follow instruction. Children must also be able to hold (and use) a pencil.

The principle investigator reserves the right to exclude any data from the study which he believes has been impaired due to language ability and/or physical or mental difficulties.

8. CONFIDENTIALITY:

The identity of participants (including which classes they attend/attended), as well as the information obtained during the course of this study (the research record) will be kept confidential to the extent permitted by law. In addition to this, the participants identities will be kept confidential in any publication of the results of this study.   However the research record may be reviewed by government agencies, individuals who are involved in or authorized to supervise or audit the research, the Ethics Advisory sub committee at Durham University, and the Principle Investigator himself.

9. TERMINATION OF PARTICIPATION:

As the contact person for this study you may discontinue the participation of your school/institution or any individual or class therein. A child may also be withdrawn from the study by the parent/guardian, the relevant class teacher, or by the child him/herself. Withdrawal may take place at any time without penalty or loss of benefits to which the person or institution terminating participation is otherwise entitled. Withdrawal means the right

to insist that all data previously provided by the child, school or class, be removed from the dataset.

The principle investigator reserves the right to withdraw participation of the school, class or individual at any time.

10. DISCLOSURE OF FINANCIAL INTERESTS:

None.

11. CONTACT PERSON(s):

If you have any questions about this research, or want to discuss any possible study related injuries please contact the Principle Investigator Adam Robson at:

Telephone number: Mob:

Email Address:

Mail Address: Department of Psychology.

Ebsworth building

University of Durham, Queen's Campus.

University Boulevard

Thornaby, Stockton-on-Tees. TS17 6BH

Alternatively if you still you still have questions regarding the study and do not wish to speak to the Principle investigator, then you may discuss them with the person supervising the research: **Dr. John Adams** by telephone: **0191 334 0108** or email: **j.w.adams@dur.ac.uk**

12. SIGNATURE:

The project has been given advisory approval by the Ethics Advisory Sub-Committee at Durham University.

By signing the line below I have read and understood the above information, initialling each page (including this one) as is required, and acknowledge that I have had an opportunity to ask questions regarding the study.

Teachers Signature:

NAME: _____ TITLE _____ DATE: _____

Researchers Signature:

NAME: _____ TITLE _____ DATE: _____

Appendix B: Standardized instructions for the RPA (Richland 2006)

"Are you ready? We are going to play the picture game. Let me show you how it works. On every page there are two pictures like this. There is a certain pattern in the top picture, and the same pattern happens in the bottom picture, but it looks different. Let me show you what I mean on this page. See up in the top picture, there is a bigger boy and a smaller boy. This is the bigger boy, and this is the smaller boy **[the experimenter pointed to each object as it was described]**. Now in the bottom picture, there is a bigger bear and a smaller bear **[the experimenter pointed]**. See, the same pattern happens in both, but it looks different. Now, in this game, first you have to figure out what the pattern is that happens in both pictures. Okay? Then I am going to point to one thing in the top picture, and your job is to tell me what is in the same part of the pattern in the bottom picture. So, on this first page, if we have a smaller boy and a bigger boy, and a smaller and a bigger bear, if I point to the smaller boy, which one is like the smaller boy in the bottom picture? Which one is in the same part of the pattern in the bottom picture? [the experimenter pointed to each object as it was described]."

If the child responded correctly, the experimenter gave feedback and then moved to the next sample problem. If the child responded incorrectly, the experimenter gave feedback and then repeated the description of the relational objects in the top and bottom pictures. The experimenter then

343

asked the question again. If the child again gave an incorrect answer, the experimenter pointed out the correct answer (the smaller bear) and moved to the next sample problem.

"Now sometimes the pattern will have two parts, like the one you just saw with the bigger boy and the smaller boy, and sometimes the pattern will have three parts. Let me show you what I mean. In this top picture, there is a mom reading to a girl, who is reading to a teddy bear [the experimenter pointed to each object]. Then in the bottom picture, there is a dad reading to a boy, who is reading to a doll. See, the pattern is the same in both pictures, but it looks different. Now, if I point to this girl, you can see that she has someone reading to her and she is reading to someone else. She has two things happening to her. Now, if I point to this girl, who is like her in the bottom picture? What is in the same part of the pattern? What is in the same part of the pattern in the bottom picture? [the experimenter pointed to each object as it was described]"

If the child answered correctly, the experimenter responded with ''Good job, perfect because this boy is the only one that both has someone reading to him and is reading to someone else. Great, let's do some more.'' If the child answered incorrectly, the experimenter gave feedback and then repeated the instructions above beginning with the description of the pattern. If the child's answer was still incorrect, the experimenter continued with this cycle

a third time and then gave the answer and went on to the experimental problems.

On each page, the experimenter pointed to the object with the arrow in the top picture and asked, ''What is like the [Insert appropriate]in the bottom picture?''