HANDS-ON LEARNING:
THE INFLUENCE OF HAND GESTURES ON CHILDREN'S RECALL OF SCIENTIFIC INFORMATION

BY

LYNLEY ANN MCLAY

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

Victoria University of Wellington

2017

Abstract

When people speak they typically also gesture. Gesture and speech form an integrated communication system, with speech conveying information in a rule-bound and sequential manner (i.e. one word follows the other in accordance with grammatical rules) while gesture conveys information holistically in a visuospatial representation. These gestural hand movements not only aid the process of speaking, but also convey important information to the listener.

While observing gesture during learning can facilitate children's understanding and remembering of novel and isolated information (e.g. Cook, Duffy, & Fenn, 2013), observing gesture may also support children in recalling complex, discursive content. This thesis examined the role observing gesture may play in supporting children's learning and recall of narrative, scientific content. The 7- to 9-year-old children, who participated in this program of research, learnt about the solar system either with or without accompanying gestures. Children's recall was assessed via interviews, both at short delays (one day) and long delays (two weeks or seven months after learning). It was hypothesised that gesture would improve children's recall by grounding the abstract scientific ideas in a physical representation, disambiguating novel terms, and providing an additional representation for children to process, store, and retrieve.

In Study 1, the influence of observing gesture in supporting children's learning and recall was examined in combination with adult initiated *wh*-questions. The study was also conducted in the presence of visual aids. Results indicated that observing gesture only had a limited effect on children's recall in Study 1 (both independently and in combination with *wh*-questions), so Study 2 examined the role of observing gesture in the absence of additional visual and verbal supports. Children's recall was assessed both the next day and seven months later. Study 3 then manipulated both the gesture children observed at learning and the gesture children performed during recall the next day (i.e. instructed, allowed or restricted from gesturing). Finally, Study 4 examined children's recall of spatial terms across the three studies.

The overarching results revealed that children who observed gesture during learning tended to report more spatial terms, but did not show improved recall of the facts and concepts taught. When children observed gesture they did, however, produce a greater rate of representational gestures during recall. In particular, children who observed gesture were

more likely to mimic the gestures they had observed, and in doing so improve their verbal recall both within the same interview and across interviews. The instruction to produce gesture did not appear to be effective in augmenting the influence of children's gesture production, but restricting children from gesturing was found to hinder recall.

Observing gesture was only indirectly effective in supporting children's recall. One possible explanation for this findings may be that children found it difficult to integrate the gestural and verbal information into a cohesive message. Perhaps it was only when children produced gesture that they were able to non-verbally access the encoded gestural content and convert it into speech. While children's own gesture appears to be the driving force in improving children's learning and recall, adults must be aware of the way they move their hands during educational lessons, as these gestures likely set the stage for how children themselves will gesture.

Acknowledgements

Just as it takes a village to raise a child, it also takes a community to support a PhD student in completing her thesis. Throughout this journey, I have been so lucky to receive encouragement and reassurance from so many different people.

I would like to thank my supervisor Deirdre Brown, without whom this thesis would not exist. Thank you for your support, scaffolding, and mentoring. Thank you for sparking my passion in research, and for keeping that fire going. I would also like to thank you my secondary supervisor, Karen Salmon, for her feedback and for keeping my ideas contained and manageable. Thanks also must go to the school principals, office staff, teachers, and children who supported and participated in this research. Without you, there would be no data and nothing to talk about. Within the lab, thank you to those who helped interview, transcribe, and reliability code the data (especially those who coded gestures). Special thanks go to my lab mate, Missy. Thank you for sharing the desk beside me, listening to me, procrastinating with me, and just being an amazing support.

On a personal note, thank you to the friends and family who have listened when I needed to complain, provided junk food when things were not going well, and just generally been amazing and awesome. Thank you to Ellen, for writing with me and for keeping me motivated and sane. Thank you also to Alex, and Lauren, who helped proofread and who picked me up when times were tough.

Lastly, but certainly not least, thank you to my Mum and Dad. I would not have made it to this point without you. Thank you for listening to me, and reassuring me when I was convinced I could not do it, and thank you for tirelessly proofreading my thesis. It is because of you both that I had the emotional strength, financial means, and general grit and determination to finish this thesis.

Table of Contents

Table of Tables

Table of Figures

Chapter 1

What is gesture?

Gestures are a natural part of conversational speech. During face-to-face interactions, and even when conversational partners are not visible (i.e. when talking on the phone or when the listener is blind), people use gestures as they talk (Alibali, Heath, & Myers, 2001; Iverson & Goldin-Meadow, 2001). Yet, gestures are not something individuals are taught to produce. Despite never having seen gesture performed, people with congenital blindness will spontaneously gesture during reasoning tasks (Iverson & Goldin-Meadow, 1998). Beyond simply accompanying speech, gestures can also visually convey additional information to the listener (Beattie & Shovelton, 1999).

The aim of this thesis was to examine the role gesture may play in enhancing 7- to 9-year-old children's learning and recall of a complex and oral scientific lesson. The primary focus was to examine whether observing gesture during the lesson would improve children's recall of the information taught. In answering this question, however, two subsidiary questions were also of interest. First, whether children's recall could be further enhanced when verbal scaffolding was used in conjunction with observed gesture. Second, whether children's own gesture production could also support their learning and recall of the scientific lesson.

In the current chapter, the term "gesture" is defined, and the role gesture may play in communicating information is explored. Given that this thesis examines *children's* learning and recall, Chapter 2 includes an overview of children's development, in terms of their ability to understand and produce speech and gesture and remember information. Chapter 3 outlines a particular style of talk found to scaffold children's learning and memory, called elaborative talk. Then, Chapter 4 discusses the benefits of gesture, both observed by children and performed by children, in supporting their learning and recall. Finally, Chapter 5 ties together research examining gestural and verbal scaffolding, with the aim of exploring how these elements of communication may support children's understanding and recall of scientific concepts.

Defining gesture and gesture types

Gestures can be broadly defined as visible actions used to provide information (Kendon, 2004). The focus of this thesis was more specific, however. Like Goldin-

Meadow (2003), the hand and arm movements that accompany speech were of interest. *Body language* more generally, physical actions upon objects, or self-focused *adaptors* (unintentional and often habitual hand movements, such as rubbing your leg or putting your hair behind your ears), were not investigated.

The communicative role gesture plays within an utterance varies along a continuum, called "Kendon's continuum" (McNeill, 2005). At one end of the continuum, speech plays the principal role in providing information and gestures serve as an accompaniment. At the other end, gestures take on a language-like form and begin to follow rules (see Figure 1.1).

Communicative burden placed on speech			Communicative burden placed on gesture
Gesticulation	Emblems	Pantomime	Sign Language
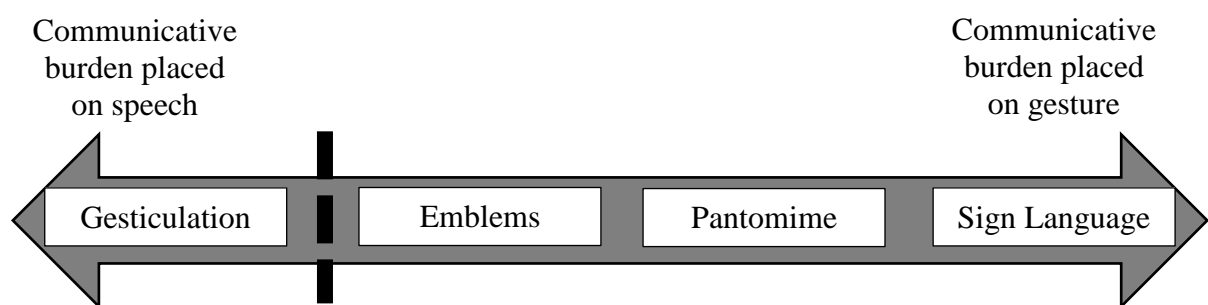
Figure 1.1. Kendon's continuum (McNeill, 2005)

According to McNeill (2005), *gesticulations* fall at the left-most end of the continuum. These gestures can only be interpreted when accompanied by speech. Take, for example, an arched gesture describing the trajectory of a dive. This same gesture could also convey any number of actions, or simply represent the shape of an object; it is only when the gesture is combined with speech (e.g. "she dived into the pool") that its meaning is understood. Gesticulations are also global, in that the individual parts of the gesture can only be interpreted by understanding the gesture as a whole (McNeill, 2005). In the example above, one must understand both that the hand represents the diver and the motion represents the trajectory. Such gestures also present the entire meaning in a single gesture, compared to the individual words required to describe the same content verbally. Similar to gesticulations are the *speech-linked gestures*. These gestures do not rely as heavily on language in order to be interpreted and accompany fill-the-blank speech, for example "Put that here" or "do it like this".

Further along the continuum are *emblems*. These gestures have a culturally understood meaning, which can be recognised without accompanying speech. These

typically include gestures like two thumbs up or the ok sign. *Pantomimes* do not accompany speech, and instead act out a story, like in a game of charades. These gestures do not follow set conventions, like in sign language, but each gesture tends to convey a specific and constrained meaning. For example, conveying the dive in a pantomime may require holding your hands together by your chest to represent yourself as the diver, and then producing the diving trajectory with your hands. At the far end of the continuum, gesture takes on a language-like form, as in *sign language* (Goldin-Meadow, 2005). Similar to speech, each gesture conveys a particular meaning, and is presented in a standardised manner. On Kendon's continuum (Figure 1.1), this thesis draws an imaginary line between gesticulations and emblems, with the focus of the thesis being on gesticulations, including speech-linked gestures.

The gesticulations and speech-linked gestures can be divided into four dimensions: Iconic, metaphoric, deictic, and beat. Iconic gestures convey concrete objects or actions and depict sematic information in an imagistic way (Goldin-Meadow, 2003; McNeill, 2005). For example, iconic gestures may include the arched motion described earlier, or producing a 'C' shape with your hand to indicate holding a cup. Metaphoric gestures represent abstract content. These may include holding one idea in one hand and a competing idea in the other hand (McNeill, 2005), dividing space into 'good' or 'bad,' or using a sweeping motion to represent the passing of time (Goldin-Meadow, 2003). Deictic gestures occur when a hand, body part, or held object is use to point (McNeill, 2005). Such pointing gestures can vary from concrete, such as saying "I gave it to her yesterday" while motioning towards a person, to more abstract, for example saying "I gave it to her" while pointing to a chair a person was previously sitting in (Goldin-Meadow, 2003). Finally, beat gestures are hand movements that move in time with speech. While these convey little sematic information to the listener, McNeill (2005) describes them as the equivalent of "using a yellow highlighter on written text" to signal the elements of speech deemed important (p. 40).

The problem with dividing gestures into categories, however, is that there is often substantial overlap in the way gestures are categorised. For example, the metaphoric gesture which divided space into 'good' and 'bad' also included deictic elements (McNeill, 2005). To deal with this issue, many studies group gestures as either representational or non-representational, as the distinction is often more concrete.

While non-representational gestures are typically defined as being beat gestures, there is variation in the definition of representational gesture. Some studies group only

iconic and metaphoric gestures as representational (Alibali & Nathan, 2012) and others also include pointing or deictic gestures (Cutica & Bucciarelli, 2011; Hostetter & Alibali, 2008). As both representational and pointing gestures have been found to benefit children's memory and recall (Cameron & Xu, 2011), this thesis follows the trend of grouping iconic, metaphoric and deictic gestures together as representational. Compared to beats, these representational gestures are the most important for aiding both the listener and speaker's understanding and recall (Cook, Yip, & Goldin-Meadow, 2012; Cutica & Bucciarelli, 2011; Feyereisen, 2006; So, Chen-Hui, & Wei-Shan, 2012).

How do individuals decide if a movement is representing information, or if it is just a hand movement produced for the sake of moving one's hands? Novack, Wakefield, and Goldin-Meadow (2016) found that a movement was more likely to be considered as representational, if the properties of the movement resembled the action needed to make changes in the physical environment. Additionally, the listener used cues, such as accompanying speech, to interpret whether the movement was meant to be communicative. This thesis focuses on the use of representational gestures that accompany speech and are utilised to communicate information to a conversational partner.

Speech and gesture form an integrated communication system

Speech and gesture appear to form an integrated communication system. Theories explaining the development of language assert that early humans communicated with their hands before verbal language took over (Corballis, 2002), and the brain areas still overlap (Dick, Goldin-Meadow, Solodkin, & Small, 2012).

Gesture and speech are consistently found to be temporally linked, with gesture slightly preceding the verbal output (Chu & Hagoort, 2014; McNeill, 2005; Morrel-Samuels & Krauss, 1992). When speech is disrupted, similar issues are also observed in the gestures produced. Those with lesions in Brocca's area show not only laboured speech but also disjointed gestural sequences (McNeill, 2005). Similar speech-gesture deficits are also found in stutterers (Mayberry, Jaques, & DeDe, 1998). This synchrony between speech and gesture makes sense, as if an interlocutor is to integrate the information contained within gesture and speech to form a cohesive message, the relationship between the content expressed across the two modalities has to be clear (De Ruiter, 2007).

Gesture and speech are also semantically linked (De Ruiter, Bangerter, & Dings, 2012; McNeill, 1992; So, Kita, & Goldin-Meadow, 2009). Those with lesions in Wernicke's area produce both speech and gestures lacking substantive information

(McNeill, 2005). Similarly, healthy individuals also show a strong link between the gestural and verbal content produced. In the Turkish and Japanese languages, unlike English, there is no direct equivalent for the verb "to swing." When describing a cartoon in which a cat swings on a rope, the Turkish and Japanese participants were more likely to use the verbs "go" or "fly," and were also more likely to produce forward motion gestures rather than the arch typically produced by English speakers (Kita & Özyürek, 2003). This gestural difference is unlikely to have resulted from observational learning, as it also occurs in individuals blind from birth (Özçalışkan, Lucero, & Goldin-Meadow, 2016). Gesture and speech appear to be synchronised in terms of what they convey and when they convey it.

Speech-gesture production theories

While researchers typically agree that speech and gestures are linked, *when*, *how* and *why* they are linked together is debated in the literature. There are two competing hypotheses. The Free Imagery Hypothesis (Krauss, Chen, & Chawla, 1996; Krauss, Chen, & Gotfexnum, 2000), suggests that both speech and representational gestures express information represented in working memory. Gestural sequences are believed to be created before speech is planned, in order to aid lexical access. The Interface Hypothesis (Kita & Özyürek, 2003; Kita et al., 2007), however, proposes that plans for gesture and speech production interact. Gesture production is believed to be informed by linguistic properties, and is created to help chunk rich spatial information into manageable units ready to be verbalised (Alibali, Kita, & Young, 2000).

The majority of research supports the Interface Hypothesis and the assertion that plans for gesture and speech production interact. For example, manipulating adults' speech, so that manner and path were described within a single clause, resulted in gesture also representing the information simultaneously. When manner and path were split across two clauses, however, gesture depicted information separately (Kita et al., 2007). Similarly, disrupting the timing of the visual feedback participants received about their gesture production (through virtual reality glasses) delayed the onset of speech, while disrupting speech resulted in participants prolonging their gestures as speech caught up (Chu & Hagoort, 2014). Altering the content or timing of either speech or gesture affected the other modality, even once action plans were in place (i.e. gesturing had already begun). These findings support the Interface Hypothesis and the assertion that speech and gesture plans continually influence each other.

Kita and colleagues (Kita & Özyürek, 2003; Kita et al., 2007), describe exactly how speech and gesture interact. They propose that when individuals have communicative intention, a preverbal message is created, which includes both what information should be expressed (content), and how to express it (grammatical structures). This then activates both the action generator (leading to gesture), and the message generator (leading to speech). The gestures that are produced represent the visuospatial information stored in memory. The form and content of the gesture, however, is determined by: 1) the communicative intention, 2) the possible action sequences which could effectively convey the content, and 3) the information from the message generator about lexical items and grammatical forms (Kita & Özyürek, 2003). Individuals organise their gesture to efficiently convey visuospatial information and to facilitate the process of verbalising the motoric-visuospatial content (Kita et al., 2007).

Differing roles of speech and gesture

While speech and gesture form a single communication system that is temporally and semantically linked, gesture and speech are co-expressive; both modalities convey important and non-redundant information (McNeill, 2005). The visuospatial nature of gesture allows for information to be presented in a global and integrated manner that the sequential and rule-bound verbal information often cannot capture (McNeill, 2005).

The similarities between the content conveyed in gesture and that conveyed in speech vary along a continuum (Goldin-Meadow, 2003). At one end are the gesture-speech 'matches' where the visual and verbal modalities work together to jointly convey a single concept (Goldin-Meadow, 2003). For example, a child may say, "I ran home over the hill" while producing an arched motion to represent the hill. In this matching example, however, the conversational partner still likely received new information about the size and shape of the hill that was not present in speech.

At the other end of the continuum are gesture-speech 'mismatches' where gesture adds novel information not conveyed in speech. For example, saying "I ran home" while producing the same arched gesture. In this case, the information about the hill is only present in gesture. The two modalities convey different information but they both contribute to the message. Gesture will rarely convey contradictory information, such as a swimming gesture while saying "I ran home," unless the individual is in a transitional state of understanding (this will be discussed further in Chapter 4, see Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Garber & Goldin-Meadow, 2002; Pine, Lufkin, & Messer, 2004). Gesture and speech are linked, but gesture typically adds important

visuospatial information that is not present in speech, and may aid the comprehension of the message as a whole.

Conclusions

Gesture and speech form an integrated system when communicating information. While the content conveyed in speech has typically been the focus of investigations examining communication, gestures also add important information within interactions. This thesis focuses on the naturalistic, representational gestures individuals produce and observe, and the influence of these hand movements in improving children's learning and recall of information taught.

Chapter 2

Child development, memory, and communication

This chapter outlines children's developing ability to understand and produce gesture and speech. Children's ability to remember information is also reviewed across the lifespan. While the majority of the research falls within the information-processing framework, sociocultural influences upon children's development are also discussed at the end of the chapter.

Information processing and the development of language, gesture, and memory

Within the developmental literature, children's cognitive development is often understood within the information-processing framework. This framework was developed in the 1970s, subsequent to the creation of the computer. It views the cognitive system as a machine, encoding, processing, and storing information so it can later be retrieved (McShane, 1991). From the developmental information-processing perspective, infants are born with innate capacities to attend, remember, speak in a native language etc. Think of this like buying a laptop with only the basic hardware. Across development, the software or skills, such as self-monitoring and strategy use, come online and are continually updated so the basic capacities of the computer can be used in more advanced ways (McShane, 1991). From this framework, development is seen to be continuous and following a single developmental path (Berk, 2009). Using research subsumed under this framework, the development of children's memory and communication skills (i.e. language and gesture) were discussed. As this thesis examined the influence of observing and producing gesture on 7- to 9-year-old children's learning and recall, particular attention was paid to this age group.

Up to two years old

Up to two years of age, there is substantial growth in the use of gesture and language. At around 8 to 12 months, children begin to recognise words and start to use preverbal gestures to signify objects or actions: for example, lifting their hand to their mouth to signify eating (Bates & Dick, 2002). While initially many of these gestures may serve a functional purpose, such as asking for food, children also start pointing to call the attention of others to things they find interesting (Leung & Rheingold, 1981). At around the age of 18 to 20 months old, children begin to produce two word combinations (e.g.

"open drawer"). A few weeks prior to this, however, children will produce the same content in a word-gesture combination (e.g. "open" plus point to drawer; Bates & Dick, 2002; McNeill, 2005). A similar progression also occurs in deaf children, with gesture-sign combinations preceding two sign sequences (Capirci, Montanari, & Volterra, 1998).

While initially gesture is the preferred method of communication, preceding language in all respects, once children start to develop speech the need for gestural symbols declines (Bates & Dick, 2002; Iverson, Capirci, & Caselli, 1994). Children instead begin to recognise speech as the primary method of communication. For example, Namy and Waxman (1998) taught children (18 to 26 months old) novel words and novel gestures to name familiar categories. They found that while the younger children learnt both the novel category names and gestures, the older children only learnt the verbal names, potentially because they knew that those were of primary importance to communication. With the development of speech, gesture use is reorganised to serve as an accompaniment, rather than the primary method of communication (McNeill, 1992).

In terms of memory development at this age, effective remembering is limited. While children may be able to recall information across a delay, their recall is inflexible and tied to specific contextual cues. In a deferred imitation task, children as young as six months old could recall information across delay, but only children 18 months of age could recall the information when contextual cues were slightly altered (Hayne, Boniface, & Barr, 2000). Without the use of language to encode information verbally, the memories remain fragile and specific (Morris & Baker-Ward, 2007).

Three to four years old

At this age, children's ability to communicate continues to develop rapidly. Children are producing and comprehending longer sentences and are now able to link sentences together to create very basic narratives (Berk, 2009). In terms of children's gesture, at age three their gesture production is becoming more adult-like, with their use of iconic and beat gestures increasing while pointing gestures decrease in frequency (Mayberry & Nicoladis, 2000). Children's ability to read gestures is also improving. At this age, children now comprehend gestures as well as they comprehend actions (Novack, Goldin-Meadow, & Woodward, 2015), and also show improvements in their ability to integrate gesture and speech together to clarify implied meaning (Kelly, 2001). Again, children of this age typically place greater importance on the verbal, rather than the gestural message (Sekine, Sowden, & Kita, 2015), but by age four children recognise and

comprehend the importance of the information presented in both speech *and* gesture (Graham, 1999).

Once children are proficiently producing speech and gesture, they provide themselves with a means of representing information in memory (see Simcock & Hayne, 2003). The development of language can then enhance the stability and generalizability of action and event representations (Hayne, 2004). As such, children's ability to recall and initiate talk about one-time past events seems to improve (Fivush, Gray, & Fromhoff, 1987; Hayne, 2004), as well as their ability to accurately recognise target items (Brown & Campione, 1972). Children still find using memory strategies challenging, however. In particular, children often show mediation strategy deficits, where they either fail to produce memory strategies or, even when provided with a strategy, do not show memory recall benefits (Miller, 1990; Schneider & Sodian, 1997).

Five to six years old

At age five, children enter the school system. They continue to develop their ability to narrate their experiences in more detailed ways (Hudson & Shapiro, 1991) and rely less on adult scaffolding while describing their experiences (Haden, Haine, & Fivush, 1997). With speech becoming the primary medium for communicating information, gesture is typically used as an accompaniment to enrich the information that is conveyed, and to test out budding ideas before they are verbalised (Goldin-Meadow, 2015). At age five, and in some cases as early as age three, children are integrating observed gesture and speech with similar proficiency to adults (Sekine et al., 2015), and can even combine novel gestural and verbal information together to aid their recall of a story as a whole (Macoun & Sweller, 2016).

Children's use of memory strategies is also improving. Baker-Ward, Ornstein, and Holden (1984) found that when children 4½ - 6½ years of age played with toys, the children who were specifically asked to recall the toys went on to engage in naming and visual scanning more than those who were not given memory instructions. However, these strategies only improved recall in the older, 6-year-old children. This indicates a utilization deficit in the younger children, as although they spontaneously produced a memory strategy they failed to benefit from it (Miller, 1990). Such utilization deficits are particularly common in children up to age seven, and decrease with age up to adolescence (Bjorklund, Miller, Coyle, & Slawinski, 1997).

11

Seven to ten years old

Middle childhood is a complex developmental stage in terms of children's gesture production and comprehension. At these ages, U-shaped developmental curves appear to emerge, so-called because children's initial competency regresses before returning later in development, often in a more advanced way. This is a phenomenon the Information Processing Framework has difficulty explaining.

At 7 – 8 years of age, children's proficiency at performing and reading gestures diverges, resulting in two U-shaped developmental curves unfolding in parallel. Children are at the top of the U-curve in terms of their ability to read and understand mismatching gestures accompanying speech (Kelly & Church, 1997), and can do so with the same accuracy as adults (Church, Kelly, & Lynch, 2000). Comparatively, children's gesture production has regressed. While, earlier in development, children's gestures were similar in complexity to those produced by adults, during this developmental period children appear to segment their gestures before later integrating them in more complex ways (McNeill, 2005).

Between the ages of nine and ten, children's ability to "read" gestures regresses. For example, when child participants watched other children solving conservation problems, the 7- to 8-year-old children could recognise both what the children had said and how they had gestured, regardless of whether they were tested verbally or non-verbally. In comparison, 9- to 10-year-olds were only able to recognise the verbal information they had heard when tested verbally, and the gestural information observed when tested non-verbally (Church et al., 2000). The authors proposed that the 7- to 8-year-old children may process gesture and speech as "an unanalysed whole." Comparatively, the 9- to 10-year-old children were beginning to process gesture and speech separately, before automatically combining the modalities again in adulthood (p. 169). In terms of gesture production, however, by age ten children are traversing up the other side of the U-curve, typically producing representational gestures at a similar rate to adults (Colletta, Pellenq, & Guidetti, 2010).

Across middle childhood, the ability to read and produce gestures is fragile. Children appear to pull gesture apart in order to analyse the components so that gesture, as a whole, is better understood and used more effectively. During this period, however, the additional load placed on children as they "read" and produce gestures may decrease the effectiveness of gesture in supporting learning and recall. For example, 8-year-old children typically produced few spontaneous gestures while recalling a spatial layout, but when

they were encouraged to gesture, they did so, and their verbal recall improved (Sauter, Uttal, Alman, Goldin-Meadow, & Levine, 2012).

In terms of children's ability to remember and recall information, children are beginning to understand the importance of chunking and creating retrieval cues. Eight- to 10-year-old children have been found to organise to-be-recalled information into groups (Bjorklund, Ornstein, & Haig, 1977; Nelson, 1974), and were found to place stars on cups in order to aid their recall of which cup hid candy (Ritter, 1978). Children seem to consider remembering as an active process, and appear to put strategies in place in order to aid encoding and retrieval.

11 years old +

From 11 to 12 years of age, and through adulthood, language, gesture, and memory strategies continue to develop but at a much slower pace. Individuals continue to advance their skills in providing shorter and more complex narratives when summarising events (Colletta et al., 2010), and from approximately 11 years of age children are using memory strategies more consistently and efficiently (Berk, 2009; Bjorklund et al., 1997; Ornstein & Trabasso, 1974).

By adulthood, gesture, language, and thought become intertwined. Observed gesture is easily converted into speech, to the extent that individuals often cannot recall what modality information was presented to them in (Kelly, Barr, Church, & Lynch, 1999). Similarly, altering a task so that an individual's earlier gesture production no longer provides accurate task information will affect subsequent performance on the task (Beilock & Goldin-Meadow, 2010).

Summary

Basic communication and memory abilities exist from birth and, with maturation, additional skills come online. The development of speech and gesture, for example, appears to enable children to encode and store information in a more organised fashion, improving the ease of retrieval. While it is important to understand trends across development, the information-processing framework fails to account for the social and cultural factors surrounding the child. These factors must be considered, as they influence individual differences.

The role of factors external to the child

Social and cultural influences are often deemed to be of limited importance in theories that explain development through maturation (Ratner, 1984). Environmental

influences can play a significant role in the development of language, gesture, and memory, however. Gesture production, for example, varies significantly across cultures. Just as the stereotype suggests, Italians gesture significantly more than Americans and their gestures are typically richer and more elaborate (Kendon, 2004). Similarly, Italian children (10 – 24 months) also produce a larger variety of representational gestures than their American counterparts (Iverson, Capirci, Volterra, & Goldin-Meadow, 2008). Even differences in cognition are observed across cultures. When attending to an image, those from Western societies typically focus on a single salient object (e.g. a fish) while those from Asian cultures will usually attend to both the object and its surroundings (e.g. the fish and the river it is in; Nisbett & Miyamoto, 2005). Children's social and cultural milieu can shape how children develop their communication and memory skills.

When examining factors influencing human development, information-processing and sociocultural approaches need to be interwoven. Vygotsky's sociocultural theory, like the information-processing framework, views individuals as having basic capacities for language and cognition, but suggests these are activated through contact with the environment (Berk, 2009). Through support from adults, or more competent peers, children are able to complete advanced tasks and use their skills in more complex ways. Scaffolding targets children at their Zone of Proximal Development (ZDP), the point just beyond children's actual developmental level where the child is only able to function with assistance (Vygotsky, 1978, 1986). Once support occurs at the social level (interpsychological) it can be internalised by the child (intrapsychological) to guide their behaviour and cognition (Vygotsky, 1978). To provide a concrete example, take the development of the pointing gesture. When a child fails to grasp an object, those around the child interpret this motion as an indirect request. With support and adult scaffolding, children then come to internalise the pointing gesture as a way of referring to objects and directing attention (Vygotsky, 1978).

Across development, children become more proficient at communicating information, both verbally and through their gesture. As a result, children's ability to recall and report their experiences also improves. How children develop these skills, however, may be influenced by the support they receive from their sociocultural environment. This thesis examines the role adults may play in supporting 7- to 9-year-old children's ability to learn and remember complex content, the likes of which they may encounter during a science lesson. While this thesis primarily focuses on the way adults' gesture production may scaffold children's learning and recall, it was also of interest to investigate how

14

adults' style of talk may support children's learning. In Chapter 3, the influence of a specific style of adult-talk, used to scaffold children's encoding and recall of events, is examined. Chapter 4 then reviews the role of both adult and child gesture in influencing children's learning and recall. Lastly, Chapter 5 reviews how these two scaffolding techniques may support children in learning and remembering complex scientific information, and outlines the specific aims of the thesis.

Chapter 3

Adult talk and memory

The way adults talk with children can scaffold how children internalise skills for encoding, organising, storing, and retrieving memories (Göncü & Gauvain, 2012; Nelson & Fivush, 2004; Ratner, 1984). For example, when teachers use mnemonic language, 6-year-old children use more memory strategies to improve their recall (Coffman, Ornstein, McCall, & Curran, 2008). Similarly, the more mothers organise to-be-recalled information while talking with their 3-year-old children, the better children's independent strategy use and recall of items at four years of age (Güler, Larkina, Kleinknecht, & Bauer, 2010). Consistent with Vygotsky's sociocultural theory of development (Vygotsky, 1978, 1986), teachers and parents play an important role in scaffolding children's cognitive development. Beyond influencing memory strategies, however, the way adults talk to children can also improve children's ability to encode and recall *event* information.

This thesis examined the role of a particular style of talk, namely elaborative talk, in supporting 7- to 9-year-old children's learning and recall of complex and discursive educational content. Specifically, this thesis examined the influence of adult posed *wh-*questions during a science lesson. *Wh-*questions have been noted to be a key ingredient within elaborative talk, and their effectiveness as a form of scaffolding, was examined both in the presence and absence of gestural content.

Elaborative Talk

The influence of adult-talk on children's encoding and recall has primarily been examined within the autobiographical memory literature. Two main styles of talk have been identified. First, elaborative talk involves adults extending the conversation, making associations, and posing questions. Second, a less elaborative style is comprised of repeated questioning and minimal elicitation of child participation within the conversation (Farrant & Reese, 2000). An elaborative style of talk is not simply a measure of verbosity, however, as mothers who use more *wh-*questions as they reminisce with their children, are no more talkative than mothers who use a less elaborative style (Haden, Ornstein, Rudek, & Cameron, 2009).

Across studies, there are variations in exactly how elaborative talk has been defined. While some studies emphasise the importance of adults providing additional

information, including statement elaborations in their definition of elaborative talk (Haden et al., 2009), others emphasise the importance of the child's participation in the conversation, by answering questions. This second view is split further; with studies differing in the types of questions they consider part of elaborative talk. Some studies include closed yes-no question elaborations in their definition of elaborative talk, such as "do you remember when…?" (Farrant & Reese, 2000; Fivush, Haden, & Reese, 2006; Haden et al., 2009). Other studies, however, only include open-ended *wh*-questions (i.e., who, what, where, when, why, how) within their definition, for example "what is this?" or "what shall we do next?" (Boland, Haden, & Ornstein, 2003; Cleveland & Reese, 2005; Hedrick, San Souci, Haden, & Ornstein, 2009). Children's participation, via open-ended *wh*-questions, appears to be a key ingredient in elaborative talk (Farrant & Reese, 2000).

Sensitive to the needs of the child

Adults who use elaborative talk typically engage children in a developmentally sensitive way. In infancy, parents tend to provide most of the information and ask yes/no questions, with open-ended *wh*-questions becoming more prevalent as children reach the end of preschool (Farrant & Reese, 2000; Fivush et al., 2006; Haden et al., 2009). This change in questioning style across development may also explain the differing definitions of elaborative talk across various studies (i.e. whether or not yes-no elaborative questions are included in the definition).

While adults may adjust their use of elaborative talk strategies, they are fairly stable in whether or not they use an elaborative style (Farrant & Reese, 2000; Reese, Haden, & Fivush, 1993). Elaborative talk can be taught, however, and has been found to be effective for enhancing children's recall of events, regardless of whether it is used spontaneously (e.g. Farrant & Reese, 2000), or following instruction (e.g. Boland et al., 2003).

Reminiscing

The role of elaborative talk in improving children's ability to report personally experienced events was first examined within parent–child dyads, during reminiscing conversations. In these studies, the parent (typically the mother) and the child talked about a past event that the child had experienced. The parent talked either as she naturally would, or was trained to use an elaborative style of talk. When mothers talked naturally during past-event conversations, the more *wh*-questions they asked during reminiscing the more new event information the children (18- to 30-months) recalled (Haden et al., 2009). Similarly, when mothers were trained to use elaborative talk, their 3- to 5-year-old children

also recalled more information than children of mothers who were not trained to elaboratively reminisce (Reese & Newcombe, 2007).

Beyond the immediate benefits within the reminiscing interaction, children also seem to internalise this elaborative style and use it on their own. For example, 3- and 5-year-old children, who had previously been engaged in elaborative reminiscing, went on to individually recall more detailed narratives compared to children who were previously engaged in less-elaborative reminiscing (Cleveland & Reese, 2005; Reese & Newcombe, 2007). Even younger children (2- to 3-years-old) seem to internalise their mothers' elaborative style of reminiscing, and exhibit more memory elaborations independently eight months later (Farrant & Reese, 2000). The skills children internalise also seem to be transferable to different memory tasks. Ratner (1984) found that when 2- to 3-year-old children were asked questions (either *wh-* or yes/no) while talking about past events they subsequently showed better performance on standard tests of long term memory (Ratner, 1984). It seems that through adult–child conversations, particularly elaborative ones, children are supported in internalising how to remember and recall information.

Talk during an event

Beyond influencing children's retrieval of past events, elaborative talk can also aid the encoding and organisation of episodic events as they unfold. Studies examining the influence of during-event talk on children's recall, typically engage children in an event and either manipulate or measure the elaborative talk that occurs. Children's recall of the event is then assessed following a delay. Elaborative talk during an event has had variable effects on the amount of information 3- to 6-year-old children recall, with some studies finding improvements (Hedrick, Haden, & Ornstein, 2009; McGuigan & Salmon, 2004), and others not (Boland et al., 2003). Studies consistently find, however, that the detail of children's recall improves (Boland et al., 2003; Hedrick, Haden, et al., 2009). Additionally, the amount and accuracy of the information children recall *non-verbally* is also enhanced by elaborative talk during an event (McGuigan & Salmon, 2006). This suggests that the children's memory representation might be shaped by the talk they are engaged in.

Recall of semantic information. The influence of elaborative talk during an event has also been examined when the experience teaches semantic content. Semantic memory is proposed to be a memory system distinct from, but overlapping with, episodic memory (Tulving, 1972). When participating in and recalling educational experiences, children are likely to engage both semantic and episodic memory.

During educational experiences, such as visits to museums, elaborative talk has been found to improve children's recall. When children, 4 – 8 years old, visited a building exhibit with their parents, children whose parents had been trained in building strong structures and elaborative talk, engaged in more joint-talk and talked more about engineering the next day than children of either untrained parents, or parents only trained in elaborative talk. The combination of building and elaborative talk training also resulted in children recalling more exhibit information two weeks later (Benjamin, Haden, & Wilkerson, 2010). In a similar study, Jant, Haden, Uttal, and Babcock (2014) gave a portion of the participating parents cards which provided example *wh*-questions to use with their 2- to 6-years-old children throughout museum exhibits. The other parent–child dyads interacted as normal. The next day and two weeks later the children and their parents reminisced about the museum exhibit. When the dyads had used cards during the exhibit, parents showed less of a reduction in their use of *wh*-questions across the two interview time points and, compared to the 1-day delay, children were more responsive to questions and recalled more information two weeks later.

More experienced children can also scaffold their peers' understanding of semantic information. King and Rosenshine (1993) found that when 10- to 11-year-old children were trained to use elaborative *wh*-questions during science lessons, children better understood the lesson immediately, and six days later. The children engaged with *wh*-questions were also better at making inferences about the lesson compared to those who were not trained to use the question stems.

Together, these studies indicated that elaborative talk might be effective for aiding children's recall of learning experiences. Both children's episodic recall and their understanding and recall of semantic information appeared to improve when they were engaged in elaborative talk. This thesis examined the role adults may play in enhancing children's recall of semantic information. Specifically, it explored the influence of posing *wh*-questions to children during a complex scientific lesson, on children's recall of the lesson the next day.

Reciprocal influences

Elaborative talk can scaffold how children learn and remember, but there are also reciprocal influences within an interaction. Beyond simply posing the *wh*-question, the child's response is also important. When mothers asked *wh*-questions throughout an event, the elements that the mother asked about were typically better recalled when the child answered the question than when they did not (Haden, Ornstein, Eckerman, & Didow,

2001). Studies also show that when children (3 – 6 years old) were more responsive to their mothers' *wh*-questions during an event they recalled more event features and more details about those features, than children who were less responsive (Hedrick, San Souci, et al., 2009). While *wh*-questions appear to be the active ingredient in elaborative talk, children's own responses to the questions are also important.

Why is Elaborative Talk effective?

Regardless whether the content is episodic or semantic, elaborative talk appears to be effective for guiding children's attention, framing the important information, and enhancing the depth of encoding and ease of retrieval. Each of these processes is discussed in turn.

Attention and organisation. When children are engaged in elaborative talk during an event, adults are able to highlight salient events through the questions they ask (Boland et al., 2003; Ornstein, Haden, & Hedrick, 2004). Such questions likely direct children's attention in order to ensure that the important content is encoded and stored. For example, Butler and Markman (2012) asked four-year-old children to watch toy animals interact. Prior to observing the animals, however, some children were asked, "Which animal makes lion laugh." Children, who received the framing questions, were significantly more able to pick the target animal than children who were not asked questions. This finding suggests that the *wh*-question appropriately directed children's attention to the relevant information and away from distractions.

Once the important information has been identified, follow-up *wh*-questions can then scaffold how the incoming information is organised and stored in memory (Boland et al., 2003). Questions like, "what did you notice happening next?" for instance, can provide children with a temporal order to the events they store.

Encoding and retrieval. When a child responds to *wh*-questions their ability to encode and retrieve the important information may also be enhanced. First, the Information Processing Theory proposed by Craik and Lockhart (1972) suggests that information is best encoded when individuals attend to meaning and make connections. When children answer questions they are creating labels, descriptions, and explanations, and likely encoding a richer representation of the event (Hedrick, San Souci, et al., 2009; Ornstein et al., 2004). Secondly, the joint-talk, resulting from asking *wh*-questions, may provide children with an opportunity to practice generating and retrieving important information. This may then make it easier to recall the information again in the future (Karpicke & Blunt, 2011). Finally, when children answer questions, adults are able to

assess what the child does and does not understand and provide any necessary corrective feedback in order to ensure that the memory is accurately stored (Ornstein et al., 2004).

Conclusions

The way adults talk to children, and specifically whether or not they use an elaborative style of talk, may significantly influence the amount and detail with which children recall their experiences or the educational content they are taught. The effectiveness of elaborative talk appeared to be driven by the use of *wh*-questions. Such questions may serve to highlight the important information, enhance the depth of encoding, and provide children with practice at retrieving the important content. Adults may also assess the children's current state of knowledge and provide corrective feedback where necessary to ensure the memory trace is accurately stored. This thesis examined the use of *wh*-questions and the potential role they may play in supporting children to learn and correctly recall complex scientific information.

Chapter 4

Gesture and memory

As described in Chapter 1, gesture is a ubiquitous part of communication, but why do we do it? Does it benefit the speaker, or do individuals produce gesture in order to aid the listener in comprehending what they are saying? The most likely answer is – both. People will gesture more when others are present, clearly showing gestures are performed to be communicative (Alibali et al., 2001; Krauss, Dushay, Chen, & Rauscher, 1995). Equally, however, congenitally blind individuals will also produce gesture when conversing with others who are blind, suggesting gesture must, in some way, benefit the speaker (Iverson & Goldin-Meadow, 1998, 2001). The role gesture plays in communication is multifaceted, both enhancing listener comprehension and the process of speaking itself.

This thesis examined the role of gesture, both observed and performed, in supporting 7- to 9-year-old children's recall of a discursive and complex scientific lesson. The research aims for this thesis were threefold: 1) to explore the role adult gesture may play in supporting children's learning and recall, 2) to examine how adult gesture may shape children's own gesture production, and 3) to assess whether children's own gesture production could improve their learning and verbal recall. The literature relating to each of these aims is discussed in turn, along with the limitations to the extant studies and mechanisms by which gesture may improve children's learning and recall. As gesture research with children tends to be more piecemeal than that conducted with adults, the adult literature is also integrated within the chapter.

Observing gesture

During interactions, adults seem to intuitively recognise that gesture may scaffold children's understanding and recall of information. When individuals know a listener has gaps in their knowledge they produce more gestures (Holler & Stevens, 2007), and this is especially likely when the listener is a child. For example, Kang, Tversky, and Black (2015) found that when adults were asked to explain a scientific process to novice children, they gestured more, and produced larger virtual diagrams in their gestures than when they were explaining the same content to expert adults. Similarly, when task demands are difficult (Grimminger, Rohlfing, & Stenneken, 2010), or when children have

cognitive deficits (Grimminger et al., 2010; Iverson, Longobardi, Spampinato, & Caselli, 2006), mothers will also produce more gestures, likely as a form of scaffolding. Research findings suggest that adults are sensitive to the needs of children and alter their gesture production accordingly. This thesis investigated whether such gestural support actually benefits children's learning and recall.

Observing gesture and comprehension

When gesture is observed during communication, studies have found that children's comprehension improves. Infants have been shown to be more responsive to their mother's verbal requests in the presence of gesture (Hodapp, Goldfield, & Boyatzis, 1984), and 5-year-olds more accurately interpreted phonetically ambiguous speech when adults pointed to the relevant object (Thompson & Massaro, 1986). When describing how to put two objects together, mothers of children with language delays gestured more than mothers of typically developing children and, as a result, mitigated any difficulties the children may have shown on the task (Grimminger et al., 2010).

Even when the content contained in speech and gesture differ, studies show that children are able to combine information across the modalities to improve their understanding of the message as a whole. For example, 3- to 5-year-old children were able to combine across the speech content (e.g. "don't forget, it's raining") and gestural content (e.g. point to rain coat) to improve their interpretation of an implied request (Kelly, 2001; Kelly et al., 1999). When gesture accompanies speech, findings indicate that children are better able to accurately interpret and respond to the information being conveyed.

Observing gesture and problem solving

When children observe gesture at encoding, noticeable improvements in their understanding and problem solving have also been found. When children were taught to solve mathematical equivalence problems (e.g. $4 + 7 + 8 =$ ___ $+ 4$) by making "one side equal to the other side," the presence of sweeping or pointing hand gestures, signaling problem solving strategies, improved 7- to 10-year-old children's written problem solving performance (Goldin-Meadow, Kim, & Singer, 1999; Singer & Goldin-Meadow, 2005). Observing gesture during mathematical equivalence lessons has also been found to enhance the children's ability to transfer their learning to new problems and aid the consolidation of their understanding across a delay (Cook et al., 2013). Similarly, pointing and tracing gestures have also improved four-year-old children's ability to solve symmetry problems, compared to when gesture was not observed (Valenzeno, Alibali, & Klatzky, 2003).

Beyond the benefits of deictic (i.e. pointing) gesture, iconic gestures (i.e. gestures representing concrete actions or objects) can also improve children's problem solving. For example, 5- to 7-year-old children showed improved performance on conservation tasks when the lesson was accompanied by gestures which highlighted the size and shape of the task objects (Church, Ayman-Nolley, & Mahootian, 2004; Ping & Goldin-Meadow, 2008). Observed gesture, whether it is deictic or iconic, seems to improve children's ability to understand and appropriately solve problems.

Observing gesture and recall

As previously outlined, the focus of the thesis was to examine whether observing gesture during learning could improve children's verbal recall of discursive, scientific information. While prior research has found that observing gesture improves 4- to 9-year-old children's recall of isolated words and sentences (So et al., 2012; Thompson, Driscoll, & Markson, 1998) and 11- to 13-year-old children's non-verbal recall of a directional path (Wermeskerken, Fijan, Eielts, & Pouw, 2016), it is only recently that research has examined children's verbal recall of narrative information.

Two studies carried out by Naomi Sweller examined the influence of observing gesture in enhancing 3- to 5-year old children's verbal recall of a narrative story. Austin and Sweller (2014) told children about the path taken by a Lego man and the actions he conducted along the way. Children heard a verbal description either alone, accompanied by beat gestures, or accompanied by representational and beat gestures. When children were asked to verbally describe the path, children who had observed gesture (regardless of type) recalled more information compared to those who did not observe gesture, and those who saw representational gestures recalled more spatial content. A similar study by Macoun and Sweller (2016) also found that children were better able to recall a narrative story, if they observed either iconic or deictic gestures during encoding, compared to either no gesture or beat gestures. The participating children were even able to verbally recall information that they had only observed in gesture. This finding indicated that the children had integrated information across the modalities into a cohesive understanding. For both isolated and semantically integrated content, the presence of gesture at encoding has been found to improve children's verbal recall compared to when gesture is absent.

Justification for the current studies

Across the studies conducted to date, there were indications that when children observed gesture at encoding their learning and recall improved. One major gap in the literature reviewed, however, was that research did not examine whether observing a

variety of representational gestures during a complex, narrative lesson also improved children's learning and verbal recall.

Research has shown that observing gesture can improve children's understanding of concepts *or* improve narrative recall, but no study has examined both. In terms of children's understanding of concepts, observing gesture at learning improved children's understanding and ability to solve problems (e.g. Cook et al., 2013; Ping & Goldin-Meadow, 2008). The limitation of these studies, however, was that the learning experiences children were engaged in did not necessarily reflect what or how children learn within everyday educational settings. The concepts were isolated and children were typically taught the same verbal phrase repeated with the same gestural content (e.g. "I want to make one side equal to the other side" accompanied by the same sweeping motion).

When children were required to recall *narrative* content, however, the studies were limited by the simplicity of the information children were required to recall. For example, children recalled a story or route, which included information they were familiar with (Austin & Sweller, 2014; Macoun & Sweller, 2016). Within these studies, gesture was also held for longer than it naturally would be produced within a normal interaction (4 seconds per gesture; Macoun & Sweller, 2016). In order to understand the role of gesture within classrooms settings, research needs to examine whether observing gesture can improve children's learning and verbal recall of complex and narrative educational content.

Studies with adults have examined the influence of observing gesture on the participants' understanding and recall of discursive scientific concepts. The findings have been mixed, however. Stieff, Lira, and Scopelitis (2016) found that those who observed gesture during a lesson on molecules did not perform any better at drawing correct molecules than those who read the text. Conversely, adults' understanding of mitosis was improved by observing gesture, but their total verbal recall of the lesson did not differ across the gesture observed and gesture not observed conditions (Kang, Hallman, Son, & Black, 2013). Throughout both these studies, however, the adults observed gesture in conjunction with visual diagrams. This may have diluted the effectiveness of gesture, either due to participants' attention being split or because of the increased cognitive load required to integrate the information across the sources (Sweller, Ayres, & Kalyuga, 2011). Further research is needed to examine the influence of gesture without such limitations.

An additional gap found in the literature, was the limited understanding of how observed gesture may influence recall across longer delays. In the studies reviewed, children's learning was typically assessed immediately after encoding, in some cases following a short distractor task. One study by Cook et al. (2013), however, did include a 24 hour delay between learning how to solve mathematical equivalence problems, and the subsequent post-test. In their study, the beneficial effects of observing gesture were enhanced over time. Research is needed to examine the influence of observing gesture on recall across longer delays as, in educational settings, children are often required to recall information weeks and months after learning.

The aims of this thesis

Based on the limitations of the previous research, this thesis aimed to answer a number of questions. The primary aim was to examine whether observing a variety of naturalistic representational gesture during encoding could improve children's verbal recall of a complex, discursive learning experience.

The lesson was designed to optimise any potential benefits observing gesture may exert on children's learning and recall. First, the learning experience was intended to be challenging for 7- to 9-year-old children, as gesture is typically more effective in aiding comprehension when the information is complex (McNeil, Alibali, & Evans, 2000). Second, the lesson included spatial content, as gesture is especially effective at conveying spatial information like size, shape, orientation, and movement (Alibali, 2005). Finally, in designing the experiment, the presence of visual aids was considered. The first study was conducted with them, and the remaining studies without.

The gesture itself, however, was not altered in order to enhance potential benefits. Unlike previous studies, which repeated the same verbal content with the same gestural information, or held the gestures for long periods, the science lesson included naturalistic gesture, which fitted with the verbal content. This was done to more closely map onto the types of gestures that may be used within educational experiences at schools.

As well as examining the influence of observed gesture on children's verbal recall of the scientific lesson, this thesis also investigated whether observing gesture influenced children's reporting of spatial content. Studies have found that when children observe gesture at learning they are better at non-verbally recalling a spatial route (Wermeskerken et al., 2016) and can verbally recall more spatial terms (Austin & Sweller, 2014). This thesis examined children's reporting of spatial terms as they recalled the learning experience.

Finally, this thesis assessed children's verbal recall at both short (1-day) and long (2-week and 7-month) delays. The various delays were examined in order to understand the effects that observing gesture may exert on children's recall over time.

Performed gesture

The way children move their hands can also support their learning and verbal recall. In fact, studies have found that the act of producing gesture may be more beneficial for learning than watching gestures (Goldin-Meadow et al., 2012). This thesis examined the influence that adults' gesture could have on the gestures children themselves produced, as well as the role of children's own gesture production in supporting their learning and verbal recall. While a variety of studies have found that performing gesture at *encoding* improves both children and adults' recall of information (Cook, Yip, & Goldin-Meadow, 2010; Cutica & Bucciarelli, 2013; de Nooijer, van Gog, Paas, & Zwaan, 2013; So, Ching, Lim, Cheng, & Ip, 2014), this thesis was interested in the gesture performed at *retrieval*.

Reciprocal influences between adult and child gesture

While children can shape the input they receive from adults through the types of gestures they produce (Goldin-Meadow, Alibali, & Church, 1993; Goldin-Meadow & Singer, 2003), this thesis investigated the role adults' gestures may play in shaping children's gesture production. Within an interaction, children will often automatically "read" the gestures they are observing and reproduce them (Cook & Goldin-Meadow, 2006), with similar effects also found in adult populations (Chui, 2014; Cook & Tanenhaus, 2009). Over time, young children (14- 58 months) appear to internalise the 'gestural style' of the adults they converse with, producing more gestures if their parents also gesture prolifically (Namy, Acredolo, & Goodwyn, 2000; Rowe & Goldin-Meadow, 2009). While gestural mimicry is typically common, it does not always occur (see Austin & Sweller, 2014; Ping & Goldin-Meadow, 2008; Valenzeno et al., 2003).

Performing gesture during verbal recall

When children gesture while recalling information, the amount of accurate content they recall typically improves. When participants have observed gesture at learning, 4- to 5-year-old participants who were instructed to gesture at retrieval recalled more story details than children restricted from gesturing (Cameron & Xu, 2011). Similarly, 9- to 11-year-old children also recalled more verbs when asked to mimic the gesture they had seen at encoding, compared to when no gestural instructions were given (de Nooijer et al., 2013).

The benefits of producing gesture are also apparent when gesture has not been previously observed. Six- to 7-year-old children, who were instructed to gesture while recalling a staged experience about a visit to a pirate, recalled more verbal information two weeks later than those who were either permitted to gesture, or restricted from gesturing (Stevanoni & Salmon, 2005). Additionally, when children (8- to 10-years-old) were asked to recall a spatial layout, the more they gestured, the more spatial information they reported in speech (Sauter et al., 2012). Together, the research findings indicate that, regardless of whether gesture has been previously observed, children's gesture production at recall can aid their retrieval of information.

The studies described, however, typically only found benefits when children were explicitly instructed to gesture. In particular, Sauter et al. (2012) found that the 8-year-old children in their study *spontaneously* conveyed significantly less spatial information in both their speech and their gesture compared to the adults and older children. It was only when the 8-year-olds were *instructed* to gesture that their reporting of the spatial layout rivalled that of the older children and adults. Children may benefit from gesture when they perform it, but in middle-childhood, children may only reap the benefits of producing gesture if they are specifically instructed to move their hands.

Performing gesture during problem solving

As was previously stated, this thesis examined the effect children's gesture production during retrieval may have on their *verbal recall*. Examining children's gesture during problem solving, however, offers important insights into the role gesture may play in children's development and understanding of ideas.

When children (5 – 9 years old) produce gesture prior to instruction they typically show improved performance on a variety of tasks, including moral reasoning problems (Beaudoin-Ryan & Goldin-Meadow, 2014), Piagetian conservation tasks (Church, 1999), and using weights to balance a beam (Pine et al., 2004). The children's gesture production, and, in particular, the creation of gesture-speech mismatches, appeared to provide the children with an opportunity to test out their burgeoning ideas. If children subsequently received verbal instruction from adults, the ideas previously tested in their gesture production could be more easily verbalised (e.g. Broaders et al., 2007).

Regardless of the type of task, the information contained within the children's gestures is of upmost importance for problem solving. When 8- to 10-year-old children were instructed to produce either correct, partially correct, or no gesture prior to a lesson on mathematical equivalence, the more correct the children's gesture, the better the

children's post-test performance (Brooks & Goldin-Meadow, 2016; Goldin-Meadow, Cook, & Mitchell, 2009). It has been suggested that when children produce informative gestures, regardless of whether or not the hand movements are spontaneous or instructed, children appear to pick up on the information contained in their hand movements and add new problem solving strategies to their repertories (Broaders et al., 2007; Goldin-Meadow et al., 2009).

The benefits of gesture production also appear to be robust across a delay. When 8- to 10-year-old children were instructed to mimic an instructor's gesture, while learning how to solve mathematical equivalence problems, children were more likely to retain their learning across a four-week delay compared to children who only mimicked the instructor's speech during learning (Cook, Mitchell, & Goldin-Meadow, 2008). The researchers suggested that gesture might create a stronger working memory trace than verbal information.

Together, findings indicate that when children gesture in accurate ways, regardless of whether they do so spontaneously or following instruction, children utilise the gestural content to propel their understanding and maintain it across delays.

Justification for the current study

While the literature to date has made an important first step, in terms of understanding the role of children's own gesture production in scaffolding learning and recall, there are a number of gaps within the literature.

First, in accord with the limitations described during the observing gesture section, gesture production studies typically have not examined children's learning and verbal recall of complex information. If children were required to understand novel concepts, such as mathematical equivalence, the influence of producing gesture at *encoding* was typically assessed non-verbally with a written, problem solving post-test. If children were required to verbally recall information, the content was typically simple and familiar, such as a story or an experience. For example, when Sauter et al. (2012) asked children to recall the locations of soft toys, the children knew all the locations when tested via a recognition paradigm; the problem was verbally recalling the information. Further research is needed to examine the potential benefits that performing gesture may exert when children are required to recall a difficult and unfamiliar educational lesson.

Second, the influence of children's gesture production was typically examined in isolation from the gestures observed at learning. Studies either have all participating children observe gesture at encoding (e.g. Cameron & Xu, 2011) or no gesture was

observed within the study (e.g. Stevanoni & Salmon, 2005). Only two studies have manipulated both the gesture observed at encoding and examined children's gesture production. Cook and Goldin-Meadow (2006) manipulated both the gesture 9- to 10-year-old children observed and performed, but all during encoding. They found that children mimicked the instructor's gestures during learning and, in doing so, improved their problem solving performance. The second study, conducted by Austin and Sweller (2014) was carried out with younger children (3 – 4 years old). The authors manipulated whether or not the children saw gesture as they heard about the actions carried out by a Lego man travelling through a scene. They found that children who observed gesture recalled more information than those who did not but that seeing gesture did not influence the number of gestures children produced. Further research is needed to examine the effect that observing gesture during learning may have on the gestures children produce at retrieval, as well as the effect children's gesture production may have on their verbal recall.

Finally, the effect of children's gesture production on verbal recall has rarely been examined across long delays. One study examined the influence of children's gesture production when recalling a staged event which occurred two weeks earlier (Stevanoni & Salmon, 2005) and the other examined the effect of children's gesture production during *encoding* on children's recall four weeks later (Cook et al., 2008). Further research is needed to examine the influence of children's gesture production as they recall information they learnt weeks and months prior.

Aims of the thesis

Given the limitations of the previous research and the gaps in the literature, this thesis explored two main questions. First, it was of interest to examine whether children, who observed a variety of representational gestures accompanying interrelated facts and concepts, would mimic the gestures they had seen. Second, it was assessed whether children's own gestures could scaffold and enhance their verbal recall of a complex scientific lesson, as well as their reporting of spatial information. Children's recall was assessed at both short (1-day) and long (2-week and 7-month) delays. The interactions between observing gesture at learning and producing gesture at recall were also examined when children's gesture production was manipulated.

<center>

The function of gesture production

</center>

While theories explaining the function of gesture as a communicative tool are sparse, researchers have examined the role gesture production plays for the speaker. The

<center>31</center>

two main theories are the Lexical Retrieval Hypothesis and the Information Packaging Hypothesis. Each is discussed in turn.

The Lexical Retrieval Hypothesis suggests that gesture improves the production and flow of speech by priming the search for spatial and semantic items from the lexicon (see Krauss et al., 1996; Rauscher, Krauss, & Chen, 1996). By seeing or producing gesture, individuals have a motoric-visuospatial way of searching for items as well as a verbal one, increasing the likelihood that the appropriate spatial word will be found. Support for this theory comes from the finding that speech describing spatial content is more fluent when individuals are able to gesture versus when gesture is restricted (Rauscher et al., 1996).

The Information Processing Hypothesis suggests that gesturing aids thinking, rather than just speech. By gesturing, individuals can explore potential ideas and examine spatial relationships by creating a holistic representation of visuospatial information. Representing the information in gesture also enables individuals to chunk the visuospatial information into manageable units ready for verbalisation (Alibali et al., 2000; Kita, 2000).

Studies comparing these hypotheses typically manipulate the difficulty of the task while keeping the lexical items required during verbalisation consistent across the conditions. If the Lexical Retrieval Hypothesis is supported then gesture production should be the same across the conditions, but if gesture is important for organisation the spatial content, as suggested by the Information Packaging Hypothesis, the difficulty of conceptualisation should influence gestural output. When 5-year-old children were asked to describe conservation objects or explain how to solve the conservation task, they produced more gesture during the explaining condition (Alibali et al., 2000). Similarly, when adults had to describe dot patterns connected into shapes, versus dots alone, the conceptualisation load required to verbalise the dots alone resulted in participants producing more gesture (Hostetter, Alibali, & Kita, 2007b). While gesture may aid lexical access, this does not appear to be the primary role gesture plays in aiding the speaker. Instead, producing gesture seems to aid the organisation of complex information and the preparation of visuospatial content for verbalisation (Kita, 2000).

An additional theory put forward, is that producing gesture acts as a form of embodied cognition (Alibali & Nathan, 2012; Cook, 2011; Cook & Tanenhaus, 2009; Hostetter & Alibali, 2008; Vasc & Ionescu, 2013). This perspective moves away from the traditional idea of the mind alone being of solitary importance for cognition, and is concerned with the role the body plays in engaging with and understanding the

environment (Wilson, 2002). Gestures are similar to actions, but at the same time are also representations of thought (Vasc & Ionescu, 2013). As such, gesturing offers a unique opportunity to externalise and potentially alter thinking. In fact, it has been found that the way participants gesture can affect their later thinking and problem solving (Beilock & Goldin-Meadow, 2010), regardless of whether or not they can see their gestures (Cooperrider, Wakefield, & Goldin-Meadow, 2015).

Explaining the idea of embodied cognition, the Gesture as Simulated Action (GSA) framework proposes that language and imagery both activate sensorimotor features (Hostetter & Alibali, 2008). In producing language, real-world perceptual and motor information has to be converted into symbols. Similarly, a mental image is also a representation of perceptual and motor information. When such sensory-motor information is activated, while creating speech or mental imagery, motor sequences also become activated and may be realised in simulated action (i.e. gesture production). Whether a gesture is produced or not will depend on the strength with which the simulated action is activated, the threshold for gesture production, and whether or not speaking is occurring (with gesture being more likely to occur with speech). This framework explains how gestures can embody an individual's thinking.

How might gesture aid cognition?

When individuals observe and produce gesture, it typically improves their memory and recall of information as well as their understanding and problem solving. Exactly how gesture benefits learning and memory is not clear, but there are a number of proposed hypotheses which tie into the theories described above.

Gesture and attention/engagement

One hypothesis is that gesture exerts its influence through the direction of attention. Many experiments have found support for this, showing that pointing was more effective for establishing joint-attention with infants than just looking (Deák, Flom, & Pick, 2000). During lessons, children look away less when gesture is present (Valenzeno et al., 2003) and are more likely to focus on tasks for longer (Wang, Bernas, & Eberhard, 2001). The presence of gesture has been found to highlight perceptual features relevant to the task at hand (Alibali & Kita, 2010), and has been proposed to make the accompanying verbal information more salient (Goldin-Meadow & Wagner, 2005). The influence on attention cannot explain the entire effect of gesture, however.

Gesture and cognitive load

Gesturing is also proposed to increase the cognitive resources available in working memory. Working memory is a limited and short term store that holds verbal and visual information in separate subsystems, called the phonological loop and the visuospatial sketchpad (Baddeley, 2003a). The phonological loops holds verbal and auditory information, providing names for information and enabling verbal rehearsal (Baddeley, 2003a). The visuospatial sketchpad, as its name suggests, holds visual and spatial information as a unified representation (Baddeley, 2003a). Each of these stores can only hold a limited amount of content but additional information can be held in one store without significantly affecting the other (Sweller et al., 2011).

Gesture is consistently found to reduce cognitive load. Individuals seem to implicitly realise this and produce more gestures when task demands are difficult (Melinger & Kita, 2007; Morsella & Krauss, 2004; Suppes, Tzeng, & Galguera, 2015), or when they have a lower working memory capacity (Chu, Meyer, Foulkes, & Kita, 2014; Marstaller & Burianová, 2013). Studies examining the influence of gesture on working memory typically ask participants to explain difficult problems while simultaneously holding irrelevant information (e.g. consonants or words) in mind. When allowed to gesture, both 7- to 9-year-old children (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Ping & Goldin-Meadow, 2010), and adults (Cook et al., 2012), recalled more of the working memory items than when gesture was restricted. The type of gesture produced appears to be important, however, as while representational gestures aided recall of the working memory items, beat gestures did not (Cook et al., 2012).

Freeing up cognitive resources, whether by observing or performing gesture, appears to improve children's memory and learning. When children have additional resources they can attend to more of the information they are being taught and put information together in new ways (see Kang et al., 2013). They also have additional resources available in order to retrieve more information at recall (Ornstein et al., 2004). How gesture might free up cognitive resources could differ depending on whether gesture is performed versus observed.

Performed gesture. When children perform gesture, they may be able to reduce cognitive load by externalising information into their hands. Just like writing something down, or physically manipulating Tetris blocks, producing gesture provides a concrete representation outside the mind (Pouw, de Nooijer, van Gog, Zwaan, & Paas, 2014). As gestural representations are easily manipulated, they can highlight new relationships,

spatial locations, or physical features, which would have placed substantial demand on working memory in order to internally represent, or compute (Chu & Kita, 2011; Novack & Goldin-Meadow, 2015; Pouw et al., 2014). With time and practice at a task, the cognitive demand placed on an individual lessens (Kalyuga, Ayres, Chandler, & Sweller, 2003). As such, the need to externalise information into the hands also reduces and individuals rely less on the external representation created in gesture (Chu & Kita, 2011; Hatano, Miyake, & Binks, 1977; Hatano & Osawa, 1983).

   Observed gesture. Like when individuals perform gesture, observing gesture also appears to reduce cognitive load. The difference is that, instead of offloading and embodying the information in the hands, observing gesture may allow content to be processed and held in an alternate working memory subsystem. When speech and gesture co-occur, verbal information is likely held in the phonological loop while gestural content is held in the visuospatial sketchpad. If gesture and speech present similar content, information can be spread across the working memory systems, reducing the load placed on any one system (Sweller et al., 2011). When novel information is presented in gesture, which is not found in speech, the additional information may be held in the visuospatial sketchpad without overloading the phonological loop. This is evidenced by the finding that children show improved problem solving performance when taught one strategy in speech and another in gesture, compared to when they are taught a single strategy across modalities or two different strategies in the same modality (Singer & Goldin-Meadow, 2005).

   There is debate, however, over whether the gestural modality or representational format of gesture is important for reducing cognitive load. The gestural modality refers to *how* gesture is presented (i.e. motor movements), while the representational format references *what* is presented (i.e. the visuospatial depiction of information). If the gestural modality is important, then gesture should not reduce cognitive load when produced by signers, as both language and gesture are presented in the same modality. Deaf signers, do benefit from producing sign-gesture mismatches, however, just like normal-hearing speakers (Goldin-Meadow, Shield, Lenzen, Herzig, & Padden, 2012). This finding suggests that it is the holistic and visuospatial nature of gesture (i.e. the representational format) which is important for reducing cognitive load, rather than the modality per se. Gesture appears to be processed differently from language, regardless of whether language is visual or verbal (Goldin-Meadow, 2015; Paivio, 1971).

Gesture conveys additional content

When individuals can encode information in a number of different ways, the depth and strength of encoding may be enhanced and the memory may be more easily retrieved (Paivio, 1971, 1991). Gesture provides a unique opportunity to present content in a number of different representational formats. Not only is gesture visual, but it also contains a motor component when performed. Even observed gesture may create an additional motor memory trace through the activation of mirror neurons (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). The presence of gesture accompanying speech, means individuals can store and retrieve memories via a visual, verbal, or motor trace. This likely enhances the depth of encoding, and the ease of retrieval.

Observed gesture. The benefit of observing additional information presented in gesture, likely differs depending on whether the gesture is complimentary (gesture-speech match), or supplementary (gesture-speech mismatch). When gesture is complimentary, children have an additional visuospatial representation of information, which they can encode and store along with speech. The associated gesture can also clarify ambiguous or confusing verbal information (Kelly, 2001; Thompson & Massaro, 1986). When gesture is supplementary, children and adults appear to combine the information across the modalities (Church, Garber, & Rogalski, 2007; Kelly et al., 1999; Sekine et al., 2015). This seems to be a fairly automatic process, as the brain is activated in a similar way when processing matching or mismatching gesture and speech (Kelly, Kravitz, & Hopkins, 2004). Supplementary gesture not only clarifies and constrain the meaning of speech but, by creating a unified message across the gestural and verbal information, can also enhance the strength of the memory trace as a whole (Kelly et al., 1999).

Performed gesture. When children perform gesture, they are likely provided with an additional representation to process, store, and retrieve. Additionally, children's gesture may also advance their thinking. The representational format of gesture may allow children to test out new ideas, and the spatial information they produce in their hands could then be used to organise and chunk the information they wish to verbalise (Kita, 2000; Kita et al., 2007; Wagner, Nusbaum, & Goldin-Meadow, 2004).

One may argue that the benefits performing gesture affords could also be achieved through performing action. Yet, action typically does not exert the same benefits as gesturing. When participants gestured a route during rehearsal they showed better recall of the route than those who physically drew it (So et al., 2014). Those who gestured while learning how to solve mathematical equivalence problems were also better able to transfer

36

their knowledge than those who physically manipulated the numbers (Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014). When participants gesture, their understanding is not bound to specific objects. This may enable children to develop a flexible and transferable representation (Novack et al., 2014). Children may also represent only the important information in gesture, leaving out superfluous details that might add confusion or overload working memory (So et al., 2014).

Summary

Gesture appears to be beneficial beyond its ability to direct attention. When individuals observe or perform gesture they may reduce cognitive load and free up resources to focus on processing, encoding, storing, and retrieving additional information. Performing gesture appears to offload information into the hands, presenting information in a way that supports children in testing out new ideas and organising speech. When individuals observe gesture, they may be able to spread the information across the working memory systems, reducing working memory load, and enhancing the depth of encoding. The influence gesture exerts on children's learning and recall of information appears complex and multifaceted.

Conclusions

When children observe or produce gesture, recall and comprehension is likely to improve. There is still work to be done in examining the effectiveness of gesture in real-world educational environments, however. This thesis explored: 1) the effectiveness of observing gesture during a comprehensive learning event on children's subsequent verbal recall, 2) whether the gesture children observe influences their own gesture production during recall, and 3) whether children's own gesture production could support their recall of a comprehensive learning experience, both as they gesture in the moment and across a delay.

Chapter 5

Science education

As highlighted throughout the introduction, this thesis examined the role that both gesture and the use of *wh*-questions (during instruction) may play in supporting children's learning and recall of a complex scientific lesson. In Vygotsky's (1986) view, children are particularly likely to need scaffolding when learning about scientific concepts. Science concepts may be difficult for children to understand as the ideas are often abstract and applicable to situations beyond which the child may have had direct experience with (Göncü & Gauvain, 2012; Vygotsky, 1986). Additionally, children may have to alter their current conceptions of the world in favour of scientific explanations (Göncü & Gauvain, 2012; Vygotsky, 1986). This chapter reviews and discussed the role that both gesture and adult initiated *wh*-questions may play in supporting children's learning and recall of scientific content.

Science and gesture

Gesture is ubiquitous in educational settings, and particularly science lessons. When adults were asked to explain scientific systems to children, they automatically created large virtual diagrams in their gesture (Kang et al., 2015). Similarly, children have also been found to use a variety of gestures while conversing about science, typically progressing from presenting incorrect information in both speech and gesture, to displaying correct information only in gesture, before correctly depicting the information in both gesture and speech (Roth, 2003). Despite the potential prevalence of gesture within science education, the role gesture may play in scaffolding children's learning and recall of science content is not well studied.

Function of gesture in science

One of the major issues within science lessons is that children may lack the necessary knowledge to understand the novel scientific terms used. Gesture may disambiguate the meaning of this new content by providing children with a second, visuospatial representation of the information (Sekine et al., 2015; Thompson & Massaro, 1986). Children's own gesture production could also provide them with actions to represent the new terminology, until the children learn to use the appropriate words (Roth, 2003). For example, children may use a forward motion to represent velocity until the word is assimilated into their verbal repertoire.

A second major challenge for children is the abstract nature of scientific concepts. Abstract ideas are not tangible or directly observable (Wiemer-Hastings & Xu, 2005), potentially making them difficult for children to understand. When children observe gesture, however, the content can be grounded and made more concrete (Alibali & Nathan, 2012; Alibali et al., 2014; Goldin-Meadow, 2015). For example, by observing or producing pointing gestures, speech can be linked to the observable environment (Glenberg & Robertson, 1999).

The complexity of scientific concepts is also problematic, as children may lack the working memory capacity to hold and process the variety of facts, concepts, and terms taught. Gesture, whether observed or produced, appears to reduce children's cognitive load (Goldin-Meadow et al., 2001; Ping & Goldin-Meadow, 2010). Additionally, when children gesture, they are potentially able to test out burgeoning ideas, and advance their understanding (Alibali & Nathan, 2012; Broaders et al., 2007; Roth & Lawless, 2002). As children's hands travel with them, they can scaffold their thinking, and reduce cognitive load, regardless of the context around them (Novack & Goldin-Meadow, 2015).

Finally, spatial content is particularly prevalent within science education, especially the physical sciences. Verbally describing spatial content is often cumbersome. The sequential and rule bound nature of speech means that each piece of information must be described one after the other, in the correct order (McNeill, 1992). The listener then has to hold all the features in mind in order to put them together in a visual representation (Rapp & Kurby, 2008). Conversely, the analogue nature of gesture means that spatial content, such as size, shape, and location, can be simultaneously presented to the listener (Goldin-Meadow, 2015; McNeill, 1992). As such, gesturing while describing spatial content may reduce cognitive load, and help organise spatial information in preparation for verbalisation (Kita, 2000).

Few studies have directly examined the influence of gestures upon children's understanding and recall of scientific concepts, but observing and performing gesture potentially plays a significant role in scaffolding learning.

Science and adult talk

Just as the pictorial information conveyed through gesture likely scaffolds children's learning and recall of scientific content, what adults say to children may also be important for the development of children's scientific understanding. This thesis focused

on adults' use of elaborative talk, and in particular their use of *wh*-questions, in developing children's ability to remember and understand scientific information.

The use of questioning has, and continues to be, a pervasive feature of science education. In science classrooms, primary school teachers have been found to ask between 15-20 content-related questions per lesson, many of which were likely *wh*-questions (Eshach, Dor-Ziderman, & Yefroimsky, 2014). Equally, informal parent–child interactions in science exhibits are also dominated by explanations and questions (Tare, French, Frazier, Diamond, & Evans, 2011). Adult-talk, and particularly the use of *wh*-questions, may play an important role in how children understand and apply their knowledge. When children were engaged in elaborative talk during a building exhibit (Benjamin et al., 2010), or when they were asked *wh*-questions during a lesson on tide pools (King & Rosenshine, 1993), children were better able to recall relevant information, and exhibited greater comprehension and transfer of knowledge than children who were not questioned. The benefits from asking questions and providing explanations have even been maintained years later (Tenenbaum, Snow, Roach, & Kurland, 2005).

Function of adult talk in science

Wh-questions are a common and influential feature of science education, likely because questioning allows adults access to what the children know and can scaffold the advancement of accurate thinking and understanding (Yoon & Onchwari, 2006). In order for children's conceptual knowledge to change, their current understanding has to be identified and challenged (Dreyfus, Jungwirth, & Eliovitch, 1990). Questioning with feedback offers an ideal opportunity for this cognitive change to occur.

The exact role *wh*-questions may play in advancing children's scientific thinking and understanding likely depends on the type of *wh*-question asked. One type of *wh*-question commonly posed is the orienting question. These questions aim to access children's prior knowledge in advance of teaching the related scientific content, for example, "What do you think would happen if…" (Osman & Hannafin, 1994). Such questions have been found to improve 15- to 17-year-old participant's recall of a scientific text and transfer of learning, compared to participants who simply read the text (Osman & Hannafin, 1994). The authors suggested that orienting questions might focus individuals on the relevant upcoming content and highlight differences between misconceptions and the appropriate scientific knowledge. This could then ensure that prior knowledge is either rejected, updated, or integrated with the new information. In science education, updating children's ideas is particularly important, as children often hold common sense ideas that

do not necessarily match the scientific principles (Duschl, Schweingruber, & Shouse, 2007).

A second question type is the retrieval question. These questions test children's understanding of the information they have been taught (e.g. "What is it called when…?"). Psychological literature commonly finds that testing improves recall in adults (e.g. Tulving, 1967; Wheeler, Ewers, & Buonanno, 2003), and children (Goossens, Camp, Verkoeijen, Tabbers, & Zwaan, 2014; Lipko-Speed, Dunlosky, & Rawson, 2014). These types of retrieval *wh*-questions are likely effective during science instruction as children can practice retrieving important content, and adults can check that the information children have encoded is accurate, and correct any misconceptions.

When children are asked questions while learning about scientific content, adults may be able to direct children's attention to the appropriate content and ensure that children are storing accurate information. Adults' questions may also provide children with practice at retrieving the information they have just been taught.

The current thesis

Based on research from a number of scholars, Ramsey and Fowler (2004) proposed that science learning is enhanced when lessons relate to children's current knowledge, when concrete materials are used, when children are active in the lesson, and when adults are present to scaffold learning. In this thesis, children were taught about the solar system in a way that included all these factors. The content was designed so that conceptual information was linked to real world examples, and concrete props were included for the children to manipulate.

Specific aims of the thesis

As outlined throughout the introductory chapters, this thesis examined the influence of gesture, both observed and performed, in supporting children's verbal recall of scientific content. The influence of gesture was examined both separately, and in combination with *wh*-questions. This thesis aimed to address a number of specific questions. Specifically, these included:

1. Does observing gesture during encoding improve children's verbal recall of a complex learning experience? (Studies 1, 2, and 3)
    a. Does seeing gesture at encoding improve the amount (total recall), depth (percentage of facts, and concepts recalled), and accuracy (percentage accurate) of children's verbal recall of the solar system lesson the next day?

b. Are any beneficial effects of observing gesture maintained across delays of 2 weeks and 7 months?

c. Does observing gesture influence the amount of spatial information children report?

2. Do children's own gestures scaffold and enhance their verbal report? (Studies 2 and 3)

a. Do children who gesture at greater rates report more information about the solar system, with greater depth and accuracy?

b. Do children who gesture more frequently verbally recall more information both within the same interview and across a delay?

c. Do children who gesture more frequently also report more spatial information?

3. Do children, who have observed a variety of representational gestures accompanying interrelated facts and concepts, mimic the gestures they have seen? (Studies 2 and 3).

Together these main questions examined both the direct effect of observing gesture upon children's verbal recall, but also the indirect effect observing gesture may have on verbal recall via influencing children's gesture production. In answering these main questions, a number of additional queries were also examined, including:

4. Whether additive improvements in children's verbal recall were found when observed gesture was used in conjunction with asking *wh*-questions (Study 1).

5. Whether the presence of visual aids within the learning experience decreased the effectiveness of gesture (Study 1 versus Study 2 and 3).

6. Whether instructing gesture or restricting gesture further influenced children's verbal recall (Study 3).

Chapter 6

Study 1: Scaffolding children's verbal recall of scientific content: The role of observing gesture and answering *wh*-questions

When a child experiences an event, there are any number of elements that they may recall. Take, for example, a child watching a baking soda volcano erupt; the child could recall the room, who was there, what the volcano looked like, what ingredients were added, what happened when the ingredients were combined etc. Knowing what to prioritise during encoding and what to retrieve during recall is something that children must learn. Vygotsky's sociocultural theory suggests that through experiences with more knowledgeable adults, children are supported to think and act in more sophisticated ways (Vygotsky, 1978, 1986). During the volcano experience, children's encoding and recall of the most salient aspects of the lesson (i.e. what happened and why) may have been enhanced if adult scaffolding was provided. For example, adults may point to the important information during encoding, or encourage children to elaborate on certain pieces of information during recall.

This study examined how adults may scaffold children's learning and recall of educational content, particularly relating to science. Children are often curious about science information, as it helps them to understand the world around them (Duschl et al., 2007). At the same time, however, children often find science lessons challenging and perceive themselves as less competent compared to other subject areas, such as mathematics and reading (Andre, Whigham, Hendrickson, & Chambers, 1999). This study examined whether manipulating adult's verbal and non-verbal scaffolding may aid children's learning and recall of the complex scientific information taught. Two forms of scaffolding were examined during a science lesson: 1) the influence of observing gesture and 2) the effect of posing *wh*-questions to children.

Gesture and learning

Gesture and speech are typically bound together during communication. Even without explicit awareness, individuals will typically produce gestures which are temporally and semantically linked to their speech (McNeill, 1992, 2005). These accompanying hand movements are more than just a side-effect of talking, instead they often provide important information to the listener (Beattie & Shovelton, 1999).

During adult–child interactions, adults seem to be implicitly aware of the scaffolding their gestures may provide. When task demands are difficult or children have disabilities, adults have been found to gesture more in order to support children's learning and task completion (Grimminger et al., 2010; Iverson et al., 2006). Similarly, when adults were instructing typically developing children about scientific ideas, they were also found to use more gesture than when teaching the same content to expert adults (Kang et al., 2015). Adults appear to be sensitive to the abilities of the listener, and implicitly adjust their use of gesture accordingly.

Observing gesture has also been found to improve children's learning and recall (see Chapter 4 for more information). Children who observe gesture during learning show enhanced verbal recall of isolated content compared to children who did not see gesture (So et al., 2012; Thompson et al., 1998). Equally, problem solving on a variety of different tasks has also been found to improve when children have observed gesture during learning (e.g. Cook et al., 2013; Ping & Goldin-Meadow, 2008; Valenzeno et al., 2003). In many of these tasks, however, children were simply required to identify a correct answer and possibly provide a short explanation. To date, few studies have examined whether observing gesture during encoding improves children's ability to verbally recall integrated, discursive information. Two studies have required 3- to 5-year-old children to recall narrative information about the activities carried out by a character (Austin & Sweller, 2014; Macoun & Sweller, 2016). In both cases, the presence of representational gestures aided children's recall compared to when gesture was not present. Neither study, however, required children to recall novel or difficult information. The current study extended upon extant research, by examining whether observing gesture could aid children's ability to recall narrative scientific information that was both unfamiliar and complex.

Observing gesture may be especially helpful for scaffolding children's learning and recall of difficult, discursive scientific information. First, the presence of gesture may direct and maintain children's focus on the relevant information taught (Alibali & Kita, 2010; Valenzeno et al., 2003). Second, observing gesture may disambiguate any information the children do not understand. Studies have found that when speech is ambiguous, the presence of gesture may aid children's comprehension (Thompson & Massaro, 1986). It is also theorised that gesture can ground the abstract scientific ideas taught within a concrete visual representation, potentially making the concepts easier to understand (Alibali & Nathan, 2012). Third, when gesture imparts information in a visuospatial format, children have an additional representation of the content, which they

can process, store, and retrieve without overtaxing working memory (Goldin-Meadow et al., 2001; Singer & Goldin-Meadow, 2005; Sweller et al., 2011). The visuospatial nature of gesture also makes it particularly efficient for conveying spatial information, which is often pervasive throughout science lessons (Alibali, 2005; Austin & Sweller, 2014; Sauter et al., 2012). Finally, observing gesture may aid children in creating a comprehensive and integrated understanding of the information taught. Cutica and Bucciarelli (2008) found that adults who observed gesture while encoding a story, showed greater recall and a richer understanding of a story content compared to those who did not observe gesture. When gesture is observed during learning, children may encode more information, more accurately, and with a greater depth of understanding.

Wh-questions and learning

The way adults talk to children also plays an important role in scaffolding children's learning and remembering (Göncü & Gauvain, 2012; Nelson & Fivush, 2004). While there is large variability in the style of talk adults use, an elaborative style has been found to be particularly helpful for supporting children's encoding and recall. This involves adults extending the conversation, making associations, and posing *wh*-questions (i.e., who, what, where, when, why, how). When adults engage children with an elaborative style of talk during an experience, children (across a variety of ages) are better able to report personal experiences (Boland et al., 2003; Hedrick, Haden, et al., 2009; Hedrick, San Souci, et al., 2009; McGuigan & Salmon, 2006), and information they have learnt (Benjamin et al., 2010; Jant et al., 2014). Research suggests that *wh*-questions are potentially the key ingredient of elaborative talk (Boland et al., 2003; Cleveland & Reese, 2005; Farrant & Reese, 2000). The use of *wh*-questions during encoding, as a verbal form of scaffolding, was the focus of the current study.

In science education, questioning is also a very common technique. Teachers typically ask around 15-20 content-related questions per lesson, including many *wh*-questions (Eshach et al., 2014) and even peer-led *wh*-questions can improve children's learning in science education settings (King & Rosenshine, 1993). The types of *wh*-questions asked may include orienting questions or retrieval practice questions. Orienting questions come before the content information, and aim to activate the children's current knowledge before providing the new content (Osman & Hannafin, 1994). Such questions have resulted in 15- to 17-year-old adolescents showing improved understanding of genetics (Osman & Hannafin, 1994). Retrieval questions, on the other hand, test children's knowledge of the information they have been taught. These questions have also been

found to improve the recall of science information in 10- to 11-year-old children (Lipko-Speed et al., 2014). In the current study, both types of questions were used throughout the event.

During instruction, questioning may improve children's recall through a number of mechanisms. First, *wh*-questions might ensure that *relevant* information is encoded. Teachers often ask questions in order to direct and maintain children's engagement and attention (Eshach et al., 2014), and such questioning may also support children in identifying the most salient information to encode and store (Boland et al., 2003). Second, questions may guide children to encode *accurate* information. By asking children questions, their current understanding can either be corrected through feedback or built upon to advance their existing knowledge (Benjamin et al., 2010; Chin, 2006). As children often hold common-sense misconceptions about scientific content (Duschl et al., 2007), questioning may be an effective means of highlighting discrepancies between prior knowledge and accurate information (Duschl et al., 2007; Osman & Hannafin, 1994). Finally, by answering questions children may be supported to *retrieve* important information. Testing allows information to be rehearsed, and may set up cues and retrieval pathways to make the memories more accessible in the future (Karpicke & Blunt, 2011).

Adult gesture and *wh*-questions as scaffolds during science instruction

Wh-questions and gestures are typically present when adults instruct children about scientific content. When interacting in a science museum, parents were found to question their 6- to 12-year-old children throughout the exhibits (Tare et al., 2011). Similarly, when adults were instructed to explain biological systems to children, all participants created large virtual models with their gestures (Kang et al., 2015). *Wh*-questions and gesture may also co-occur. Jant et al. (2014) found that parents who employed a greater number of *wh*-questions also produced more non-verbal behaviours, such as pointing. Allen and Shatz (1983) also found that 1-year-olds infants' responses to *wh*-questions were affected by the presence of gesture. Research to date, however, has not examined the effects of gesture and *wh*-questions in combination. Together, these forms of scaffolding may provide additive benefits. Answering *wh*-questions could ensure that children encode and retrieve appropriate and accurate information, while the complementary and supplementary information provided through gesture may support children's comprehension, and enhance their depth of encoding.

The current study

This study examined the unique and combined benefits of *wh*-questions and hand gestures in scaffolding children's learning and recall of scientific content. Children, 7 – 9 years old, individually participated in an interactive, staged scientific learning experience about the solar system. The age range was selected to be similar to that employed by previous research examining observed gesture (e.g. Cook et al., 2013; Ping & Goldin-Meadow, 2008) and elaborative talk (Benjamin et al., 2010).

The content was delivered in one of four ways: verbal information only, verbal information and observed gesture, verbal information and *wh*-questions, or verbal information and *wh*-questions and observed gesture. Children were interviewed to assess their recall the following day. This delay was chosen for two reasons: 1) beneficial effects of *wh*-questions have previously been found when a single day has passed between encoding and recall (e.g. Benjamin et al., 2010), and 2) gesture research typically has not examined delays longer than one day.

It was hypothesised that children who answered questions or saw gesture during learning would recall more information, with greater accuracy, than those who were not provided with such scaffolding. It was also expected that when gesture and elaborative talk were used in tandem, even greater verbal recall benefits would be observed than when either communication device was used in isolation.

Children's separate recall of both the facts and concepts taught were also examined. This was done to assess whether observing gesture and answering *wh*-questions differentially supported children's recall of isolated facts, versus the integration and understanding of why these facts were the case (i.e. concepts). It was hypothesised that observed gesture and *wh*-questions would separately improve both children's recall of facts and concepts. Additional recall benefits were again hypothesised when both the scaffolds were used in combination.

Method

Participants

Eighty-one children, 7 – 9 years of age, were recruited from two local Wellington primary schools. Two children were excluded from the original sample; one child was absent on the days the interviews were conducted, and the other withdrew before the conclusion of the interview. The final sample included 79 children (M_{age} = 8.5 years, SD = 0.50). There were approximately 20 children per condition (see Table 6.1 for sample

characteristics). Children had not learnt about the solar system at school during the school year.

All children had parental consent to participate in the study, and assented to their own involvement. The study was approved by the School of Psychology Human Ethics Committee under delegated authority of Victoria University of Wellington's Human Ethics Committee.

Materials

Parental rating scales. In order to control for children's prior knowledge and interest in the solar system, parents completed two Likert scales. The child's interest in the solar system was rated on a 9-point scale, from 1 ("not at all interested") to 9 ("extremely interested"), and children's knowledge about the solar system was also rated on a 9-point scale from 1 ("limited understanding") to 9 ("advanced understanding") [1]. Ratings could not be obtained from one parent.

Procedure

Parents of children participating in the study were asked to complete the parental rating scales attached to the consent form. These ratings, as well as children's age, were taken into account when assigning children to the four conditions. Parental ratings of children's interest in and knowledge about the solar system did not significantly differ across conditions and neither did child age ($p > .05$ for all main effects and interactions), see Table 6.1.

Solar system lesson. Once children were assigned to a condition, they individually learnt about the solar system with the author. The room was set up with the instructor and child standing on the edge of a large 3×3m mat depicting a model of the solar system. A camera recorded their interactions. Throughout the lesson, the mat was referenced in accordance with the lesson script.

---

[1] See Appendix A for information on the validity of the parental rating scales (p.173).

Table 6.1

Means and Standard Deviations of Children's Age and Parental Ratings of Children's Interest and Knowledge Across the Conditions

	Number of participants			Child's age (in months)		Parental ratings of child interest		Parental ratings of child knowledge	
	Males	Females	Total	*M*	*SD*	*M*	*SD*	*M*	*SD*
Verbal-only	9	10	19	101.60	6.03	5.32	2.36	4.26	1.63
Verbal+*Wh*-questions	10	10	20	102.54	6.93	5.55	1.47	4.30	1.87
Verbal+Gesture	10	10	20	102.49	5.66	5.55	1.23	4.30	1.30
Verbal+Gesture+*Wh*-questions	9	11	20	101.81	7.21	5.50	1.85	4.20	1.58
Total	38	41	79	102.12	6.38	5.48	1.74	4.27	1.58

The lesson script was informed by children's non-fiction books and, through collaboration with the Carter Observatory and the implementation of pilot testing[2], was designed to be both novel and developmentally appropriate. The lesson script was divided into four sections. An introductory section provided children with basic knowledge about the solar system to ensure all children started with the necessary knowledge to understand the main lesson. This section named and located the eight planets and highlighted the differences between the inner and outer planets. The remaining three topics were the primary focus of the lesson, and included information on planetary movements, the role of gravity in space, and the consistency of the surfaces of the planets. Each section included an interactive demonstration and a fact uncommonly known (see Table 6.2 for more information). The lesson took approximately 15 minutes and children were asked not to talk unless they were asked a question. If the child did ask a question, the instructor declined to comment.

The only difference across the conditions was the type of support provided (i.e. gesture and/or *wh*-questions). In the *Verbal-only* condition, the instructor taught the solar system verbal content without any additional support. In the *Verbal+Wh-questions* condition, children heard the same verbal information interspersed with scripted *wh*-questions, to which children were encouraged to respond. In both the *Verbal-Only* and *Verbal+Wh-questions* conditions, the instructor kept her hands in her pockets while speaking. In the *Verbal+Observed-gesture* condition, the verbal information was accompanied by informative hand movements and in the *Verbal+Observed-gesture+Wh-questions* condition children both answered *wh*-questions and observed hand gestures (see Table 6.3 for an example of how the conditions differed).

Children who observed gesture during the lesson saw the instructor produce a variety of representational gestures that complemented the verbal information. These gestures included: 1) iconic gestures, conveying information about concrete objects or actions (e.g. a ball shape was created with both hands to represent a planet), and 2) metaphoric gestures, conveying information about abstract ideas (e.g. fists were pulled inwards, towards the body, to represent gravity's pull). Deictic gestures were also used to point out the locations or objects on the mat. See Figure 6.1 for more information.

---

[2] See Appendix A for information about pilot testing the solar system lesson (p.173).

Table 6.2

Description of Lesson Topics

	Topic 1	Topic 2	Topic 3
Primary focus	Planetary movements	Role of gravity in space	Consistencies of the rocky and gas planets
Information taught by the instructor	The instructor defined an orbit, outlined the three rules an object has to meet to be counted as a planet (i.e. spherical shape, orbit the sun, and a clear orbit path), and described why planets vary in the time taken to orbit the sun.	The instructor taught children how bigger things have more gravity, as well as the sun's role in keeping the planets in place while they orbit. They were also taught how the black hole keeps the orbiting suns in place.	The instructor explained the differences between solids and gases and linked these explanations to the consistencies of the rock and gas planets.
Demonstration and purpose	Children were timed walking around the mat on Mercury and Neptune's orbit lines to highlight the different orbit lengths	Children swung around a ball tied to a piece of string to demonstrate how the string holds the ball in place like the sun's gravity holds the moving planets in place.	Children stood in a box filled with dirt and rock, and a box which the experimenter blew into. This demonstrated how you could stand on the rock planets but not the gas planets.
Uncommon fact	Why Pluto isn't counted as a planet	The black hole holds the sun in place and the length of time taken for the sun to orbit the black hole	There might be diamonds in the middle of gas planets.
Prop			

For each of the three topics, children in the *wh*-questions conditions were asked five *wh*-questions, resulting in 15 questions being asked throughout the lesson. The questions all began with who, what, where, when, how or why, but varied in their purpose. Some questions were *orienting* questions, and were asked before delivering the scripted educational content. For example, children were asked, "What do you think makes something a planet?" before they were taught about the rules an object has to meet in order to be considered a planet. Children were also asked retrieval based *wh*-questions, which highlighted particular information they had encountered during the demonstrations. For example, children were asked, "What did you notice happening when you were spinning the ball around?" If children responded incorrectly to a question, they were told, "That's not quite right" and were then told the relevant scripted information which explained the correct answer. If they responded correctly, the child was told, "Yes, that's right" and they heard the same scripted content.

Interviews. The next day, children were interviewed, by a different research assistant, to assess their recall of the lesson content. Four of the 79 children were interviewed two days later due to illness or participation in other school activities during the time of interviewing. The interview followed a standardized protocol, which involved three phases of questioning. First, children were asked to freely recall everything they remembered, followed by two additional prompts (e.g. "tell me some more things you remember"). Phase two encouraged children to elaborate on previous responses (i.e. cued recall). The interviewer summarised each piece of information provided by the child and asked for more information (e.g. "So you said you learnt about gravity and the sun. Tell me all about that"). The third phase prompted children to recall the three solar system topics (i.e. prompted recall): 1) "I heard you learnt about the planets and how they move in orbits. Tell me all about that," 2) "I heard you learnt about why everything stays in the solar system. Tell me all about that" and, 3) "I heard you learnt about the types of the planets. Tell me all about that." After each prompted recall question, children received two further open-ended prompts (e.g. "Have another big think and tell me anything else you remember about the types of planets"). Children received a small gift after both the lesson and interviews, and each classroom received a $20 voucher from a local retail store.

Table 6.3

Outline of Solar System Script Across Conditions

		Conditions		
Introduction	Verbal-only	Verbal+ *wh*-questions	Verbal+ gesture	Verbal+ *wh*-question + gesture
Now let's talk about the planets and how they move	For something to be a planet it has to have a round shape and it has to move around the sun in a circle	**What do you think makes something a planet?** For something to be a planet it has to have a round shape and it has to move around the sun in a circle	For something to be a planet it has to have a round shape **[both hands joined in a sphere shape]** and move around the sun in a circle **[pointer finger from right hand moving around fist of left hand]**	**What do you think makes something a planet?** For something to be a planet it has to have a round shape **[both hands joined in a sphere shape]** and it has to move around the sun in a circle **[pointer finger from right hand moving around fist of left hand]**

Point to sun on mat					
The sun	is the biggest thing	in all of the solar system	and because the sun is so big	It has lots of gravity to pull the planets in	And stop them from floating away into space

Figure 6.1. Example of gestures used when teaching children about the role of gravity in the solar system

Coding

Based on the event script and information about the event set-up, it was calculated that children could report 224 pieces of information about the event. Within each interview transcript, it was identified whether each unit of information was *present, absent*, or *developing*. *Developing* codes were assigned when children: a) used terms incorrectly but showed an understanding of the concept (e.g., saying "air planets" would receive a developing code for naming the gas planets), b) reported partial information which showed a correct understanding (e.g., saying "the sun has pull" would receive a developing code for the fact that the sun has gravity), or c) when children implied information without specifically stating it (e.g., saying "the first four planets" would receive a developing code for the idea that there are a group of four planets closer to the sun). Thus, for each of the 224 pieces of information children were given a recall score of 0 (absent), 1 (developing), or 2 (complete), resulting in a total possible score of 448. Children's total recall was calculated as the sum of item scores (i.e. total possible score of 448), collapsed across the interview phases.

Information was coded as an error if it was related to the solar system or to space but included either *intrusions* of information that the children knew from other sources (e.g. mentioning moons) or *distortions* of information they were taught (e.g. talking about gravity as a push rather than a pull). Repetitions, the child's opinion, and unrelated pieces of information were not coded. Accuracy (i.e. proportion correct; total correct recall/total correct recall + errors) was calculated to examine errors relative to the correct information provided by the child.

Lesson information was also identified as either a fact or a concept. *Facts* included describing the properties of the solar system and the objects in and around it (e.g. "planets move around the sun" or "the black hole has gravity"). *Concepts*, on the other hand, included information that linked ideas together and explained information (e.g. "planets move around the sun because of the sun's gravity" or "the middle of the gas planets is liquid because of the pressure inside the planet). Children's recall of facts and concepts were tallied across the three main solar system topics (i.e. planetary movements, role of gravity, and planet consistencies). Of the 159 codes assessing children's recall of the three main topics, 114 were identified as either facts or concepts[3]. Totals for facts and concepts

---

[3] Coding and analyses were also carried out for children's percentage recall of names (i.e. scientific terms) and contextual content. See Appendix B, Supplementary analysis S1.1 (pp.175-176).

were converted to percentages by adding scores across the relevant event components and dividing by the maximum score children could obtain (2 points for every relevant event component). There were 75 codes for facts (total = 150) and 39 codes for concepts (total = 78).

A research assistant blind to the hypotheses of the study coded 20% of the transcripts to establish reliability. Interrater reliability was calculated using Cohen's Kappa and was examined two ways. The agreement between coders when information was coded as either present or absent across the three phases of the interview was examined ($k = .80$) as well as agreement when information was coded as present, absent or developing ($k = .76$). Finally, coders' interrater reliability was also assessed when identifying errors versus correct information ($k = .62$).

Results

The individual and interactive effects of observing gesture and answering *wh-*questions during learning were examined in terms of the amount (i.e. total recall) and accuracy (i.e. total/(total recall +errors)) of children's verbal recall. It was also of interest to examine children's percentage recall of facts and concepts across the conditions. Prior to analyses, the distributions of the verbal recall measures were examined. Children's total recall and recall accuracy were not normally distributed[4]. Total recall was corrected with a square root transformation, and children's accuracy scores were corrected by reverse scoring the items, applying a square root correction, and then back transforming the data (see Field, 2013).

To examine whether possible ceiling or floor effects may have been occurring, the amount of information the children recalled was examined when collapsed across the conditions. The event was divided into 224 items, and each item could receive a maximum score of two. Of the 448 items, children were reporting between one-quarter and one-third of the content ($M = 120.19$, $SD = 39.99$). Children's recall was also examined in terms of the number of lesson items recalled (regardless of whether they were recalled in a complete or developing manner). Of the 224 items, children were similarly reporting between a quarter and a third of the content ($M = 64.86$, $SD = 20.99$). Children's recall was highly accurate, however (90% accurate, see Table 6.5). While children's reporting was at

---

[4] See Appendix B, S1.2 for Kolmogorov-Smirnov normality information for the verbal recall measures (p.176).

the lower end, there was still large variation within the sample, indicating that floor effects were unlikely to have been an issue.

Total verbal recall.

To test the hypothesis that children who observed gesture or answered *wh*-question would recall more information, a 2 (Gesture: Observed vs Not observed) × 2 (*Wh*-questions: present vs absent) factorial ANOVA was conducted. Children's knowledge rating was included as a covariate, see Table 6.4. Contrary to hypotheses there was no significant main effect of observing gesture ($F(1, 74) = 0.09$, $p = .769$, $\eta_p^2 = .001$), or answering *wh*-questions ($F(1, 74) = 3.57$, $p = .063$, $\eta_p^2 = .05$), and no significant interaction ($F(1, 74) = 0.41$, $p = .526$, $\eta_p^2 = .01$). See Table 6.5 for raw means and SDs.

Table 6.4

Correlations Between Age, Parental Ratings of Children's Interest, Parental Ratings of Children's Knowledge and Verbal Recall

	Interest	Knowledge	Total Recall	Accuracy	Facts	Concepts
Age	-.01	.23*	.14	.01	.19	.14
Interest		.48*	.20	.18	.17	.26*
Knowledge			.33**	.16	.34**	.32**

* $p < .05$ ** $p < .01$

Accuracy.

It was also hypothesised that children who observed gesture or answered *wh*-questions would recall the solar system information with greater accuracy. A 2 (Gesture: observed vs not observed) × 2 (*Wh*-questions: present vs absent) ANOVA was conducted to test this assertion. There was a significant main effect of gesture condition ($F(1, 75) = 4.67$, $p = .034$, $\eta_p^2 = .06$), with those who observed gesture recalling information with greater accuracy than those who did not. There was no main effect of *wh*-questions, however ($F(1, 75) = 0.004$, $p = .950$, $\eta_p^2 < .001$) and no significant interaction ($F(1, 75) = 0.001$, $p = .981$, $\eta_p^2 < .001$), see Table 6.5.

Table 6.5

Means (Standard Deviations) for Children's Verbal Recall Across The Conditions

		Total recall	Accuracy	Facts	Concepts
Gesture-not observed	No *wh*-questions	130.47 (45.27)	88.25 (8.86)	30.72 (12.67)	22.32 (10.62)
	Wh-questions	109.10 (39.27)	88.96 (7.45)	26.19 (11.83)	16.79 (9.81)
Gesture-observed	No *wh*-questions	127.25 (40.37)	92.29 (5.51)	32.61 (10.59)	20.67 (10.49)
	Wh-questions	114.45 (33.59)	91.65 (8.30)	27.47 (11.13)	21.32 (12.74)
Total	No *wh*-questions	128.82 (42.29)	90.32 (7.52)	31.69 (11.53)	21.47 (10.45)
	Wh-questions	111.78 (36.17)	90.31 (7.90)	26.83 (11.36)	19.06 (11.46)
	Gesture-not observed	119.51 (43.11)	88.61 (8.07)$_a$	28.40 (12.30)	19.49 (10.46)
	Gesture-observed	120.85 (37.23)	91.97 (6.96)$_b$	30.04 (11.04)	20.99 (11.52)
	Total	120.19 (39.99)	90.31 (7.67)	29.23 (11.63)	20.25 (10.97)

Note. Conditions with different subscripts were found to significantly differ.

Recall of facts and concepts

To examine the hypothesis that observing gesture and answering *wh*-questions would enhance both children's recall of facts and their recall of concepts, two 2 (Gesture: Observed vs not observed) × 2 (*Wh*-questions: present vs absent) factorial ANOVAs were run. Percentage recall of facts and concepts were separately examined as the dependent variables and knowledge ratings were included as covariates, see Table 6.4. While interest scores also correlated with the recall of concepts, given the significant correlation between knowledge and interest, only knowledge was included as a covariate.

Contrary to hypotheses, there were no significant main effects of observing gesture for the recall of facts ($F(1, 74) = 0.46$, $p = .500$, $\eta_p^2 = .01$), or the recall of concepts ($F(1, 74) = 0.41$, $p = .522$, $\eta_p^2 = .01$). Similarly, there was no significant main effect of *wh*-questions on the recall of concepts ($F(1, 74) = 1.02$, $p = .316$, $\eta_p^2 = .01$), but there was a trend that those who answered *wh*-questions actually recalled fewer facts than those who did not answer *wh*-questions ($F(1, 74) = 3.75$, $p = .057$, $\eta_p^2 = .05$). This was the opposite effect to what was hypothesised. There were no significant interactions between the observed gesture and *wh*-questions conditions for either the recall of facts, or concepts (all *p*'s > .05), see Table 6.5.

Post hoc analyses examining children's responses to the *wh*-questions

Responding to *wh*-questions during the lesson. Previous studies examining elaborative talk and *wh*-questions have found that children mostly respond correctly to the questions asked (Hedrick, San Souci, et al., 2009). This study also examined children's responding to the *wh*-questions during the solar system lesson. Descriptive statistics revealed that children's answers to the *wh*-questions posed during the lesson were only accurate 45% of the time ($M = 44.24$, $SD = 13.99$).

Recalling *wh*-question answers during the interview. During the interview, it was also examined whether children in the *wh-questions* learning conditions better recalled the answers to the lesson questions compared to children in the *no-wh-questions* conditions. Of the 159 codes that related to the three main event components, 37 assessed information that would directly answer the *wh*-questions posed at learning. An independent samples *t*-tests found that children in the *wh-questions* conditions ($M = 25.98$, $SD = 13.10$) and *no-wh-questions* conditions ($M = 29.78$, $SD = 11.84$) did not significantly differ in their interview recall of the answers to the questions posed at learning ($t(77) = 1.35$, $p = .180$, $d = .30$).

Across the two conditions, children's recall of lesson content during the interview that did not answer the *wh*-questions was also assessed (i.e. the other 122 codes of 159 total). Interestingly, children in the *wh-questions* conditions ($M = 24.63$, $SD = 9.90$) showed significantly reduced recall of content that was unrelated to the *wh*-questions, compared to children in the *no-wh-questions* conditions ($M = 30.06$, $SD = 10.50$; $t(77) = 2.37$, $p = .020$, $d = 0.53$). In summary, compared to the *no-wh-questions* conditions, children who were asked *wh*-questions at learning were no better at recalling the question answers during the interview but were worse at recalling content unrelated to the questions.

Accuracy of *wh*-question responses in the event and recall during the interview. Finally, it was examined whether children's accurate responding during the *lesson* was associated with children's correct recall during the *interview*. Pearson's correlations revealed that when children correctly responded to the *wh*-questions during the lesson they also recalled more of the *wh*-question answers during the interview ($r(38) = .43$, $p = .006$). Similarly, children who correctly responded to more of the *wh*-questions during the lesson also recalled more information during the interview (i.e. total recall; $r(38) = .51$, $p = .001$), as well as more facts and concepts ($p < .05$). Correct responding during the lesson was not related to children's overall accuracy during the interview, however ($r(38) = .25$, $p = .122$).

Discussion

Adults commonly use hand gestures and *wh*-questions when they interact with children in educational settings. This study examined the separate and combined benefits of these two forms of scaffolding for improving children's recall of a complex, discursive scientific lesson.

It was hypothesised that when children observed gesture or answered *wh*-questions during learning they would recall more information, with greater accuracy. Further recall benefits were also expected when the two scaffolds were used in tandem. While children who observed gesture during learning were found to be more accurate in their verbal recall, no improvements were found in the amount children reported. Comparatively, no beneficial effect on either the amount or accuracy of children's recall were found when answering *wh*-questions, and no additional improvements occurred when the two scaffolds were engaged in combination.

Across the conditions, children's differential recall of facts and concepts were also examined. It was expected that both children's recall of isolated facts and their recall of connections and explanations (i.e. concepts), would improve with adult scaffolding. Again, neither observing gesture nor answering *wh*-questions improved children's recall of facts or concepts. If anything, answering *wh*-questions appeared to reduce children's recall of facts (though this was only a trend). Altogether, these findings beg the question as to why these two forms of scaffolding did not improve children's learning and recall as they have in previous studies.

Gesture as a form of scaffolding

Given that only children's accuracy improved when gesture was observed at learning, the effectiveness of the gestural scaffolding appeared to be diluted compared to what was hypothesised. This may have occurred due to the nature of the learning experience. The science lesson was designed to mimic the type of experiential learning that may occur in a museum or classroom. As such, it included visual aids (e.g. mat representing the solar system) and props for the children to manipulate (e.g. swinging around a ball tied to a string). The influence of the visual aids (i.e. props, mat, and posters) and the interactive nature of the event may have limited the influence of observing gesture, however.

By presenting visuospatial information in gesture, children's learning and memory was expected to improve. Gesture was proposed to disambiguate any verbal information the children did not understand (Thompson & Massaro, 1986) and provide a means of grounding the abstract content in the concrete (Alibali & Nathan, 2012; Goldin-Meadow, 2015). It was also theorised that gesture would reduce children's working memory load by spreading the cognitive burden across the working memory subsystems, i.e. the phonological loop and the visuospatial sketchpad (Goldin-Meadow, 2015; Goldin-Meadow et al., 2001; Ping & Goldin-Meadow, 2010; Sweller et al., 2011). Just like gesture, visual aids are also visuospatial in nature and likely act upon similar mechanisms.

While both gesture and visual aids may be separately beneficial for improving children's comprehension and recall, in combination they may overload working memory. In the current study, children likely had to hold, process, and integrate various pieces of visual information together, potentially creating substantial cognitive load (Sweller et al., 2011). Given the similarity of the information presented across the visual modalities, cognitive resources may also have been unnecessarily taxed by redundant content (Sweller et al., 2011). Because of the increased cognitive load created, children may have been left

with limited capacity to encode and store additional information. Furthermore, presenting information in a number of different visual formats may have drawn children's focus away from the fleeting gestures, reducing the likelihood that the gestural information was either attended to or encoded. These factors may explain why observing gesture was not as effective at enhancing recall as it was expected to be.

Comparing the gestural and visual content, only movement information was likely to be conveyed uniquely in gesture. Yet, any potential benefit from observing gestural movements may have been washed out by the children actually *performing* action. In the current study, children walked around the orbits, swung a ball around, and stood in the boxes with air and dirt. This provided the children with an opportunity to embody their understanding, test out their current knowledge, and alter their thinking based upon the observable effects they were exerting upon their environment. Unlike observing gesture, during which children may only be able to quasi-embody their understanding through the activation of mirror neurons (Rizzolatti et al., 1996), the direct action the children exerted was probably more effective for memory (see Baker-Ward, Hess, & Flannagan, 1990). The limited number of distinct mechanisms by which gesture was exerting its effects on recall may explain why the amount children recalled did not differ when children either did or did not observe gesture at learning.

Interestingly, however, observing gesture did improve the accuracy of children's verbal report. This may have been due to children creating a more event-specific memory. A study by Alibali and Kita (2010) found that when children (5 – 7 years old) performed gesture while solving a conservation task they talked more about the perceptually present content compared to children who were restricted from gesturing. In the current study, observing gesture may have highlighted props and aids at the appropriate time, enabling the child to integrate the visual aids with the verbal information they were hearing. Tying the verbal information to the specific visual aids and props may have created a tighter event-memory that was less prone to intrusions or distortions. While the interactive nature of the lesson may have washed out the effect observing gesture had on the amount children recalled, gesture still appeared to improve the accuracy with which children recalled their learning.

Wh-questions as a form of scaffolding

Whether or not children answered *wh*-questions during learning had no significant effect on the amount or accuracy of children's verbal recall. When children's recall of facts and concepts were separately examined, however, there was a trend that answering *wh*-

questions resulted in the recall of fewer facts. This was unexpected given that many studies find elaborative talk can be beneficial for scaffolding children's learning and recall during museum exhibits (Benjamin et al., 2010; Tare et al., 2011), and that *wh*-questions have improved children's recall of science lessons (King & Rosenshine, 1993; Osman & Hannafin, 1994).

The use of *wh*-questions was expected to allow the experimenter access to the children's current knowledge and enable the correction of misconceptions. Despite this hypothesis, literature on the benefits of feedback has been mixed. A meta-analysis found that while feedback was generally beneficial, in a third of the cases the effects were actually harmful for learning (Kluger & DeNisi, 1996). Differences in working memory capacity (Fyfe, DeCaro, & Rittle-Johnson, 2015) and prior knowledge (Fyfe & Rittle-Johnson, 2015; Fyfe, Rittle-Johnson, & DeCaro, 2012) have been found to influence the effectiveness of feedback and may similarly influence the effectiveness of *wh*-questions. The effect of each factor is separately discussed below.

Working memory. Working memory is a limited, short term store which is prone to overload (Baddeley, 2003b). When overload does occur, learning suffers (Sweller et al., 2011). In the current study, it was found that children who did and did not answer *wh*-questions did not significantly differ in their recall of content directly relevant to the questions asked. When their recall of the content unrelated to the questions was examined, however, children who answered *wh*-questions recalled less information than those who did not. Answering *wh*-questions and updating one's conceptual understanding requires cognitive resources (Fyfe et al., 2015; Fyfe & Rittle-Johnson, 2015). It could therefore be speculated that when children were required to answer *wh*-questions, they lacked the working memory capacity to process and store both the question information *and* the additional content being taught (Sweller et al., 2011). This may explain why *wh*-question did not improve children's verbal recall overall.

Children in the *wh*-questions condition were also typically responding incorrectly to the questions asked during the lesson. In response to incorrect answers, children were told, "That's not quite right" and then heard the standardised, scripted lesson content. In updating their understanding, children likely had to hold their own answer in mind while also searching through the dense incoming narrative for the correct information. This likely created substantial cognitive load (Sweller et al., 2011). In naturalistic conversations, where beneficial effects of *wh*-questions have been found, cognitive demands may have been reduced for the children. Adults could provide specific examples

to counter children's misconceptions and follow-up questions could be asked to guide the children towards an appropriate understanding (Yoon & Onchwari, 2006). The cognitive load created when answering questions and the standardised nature of the lesson may have reduced the effectiveness of the *wh*-questions asked.

Prior knowledge. During the lesson, the accuracy with which children responded to the questions was significantly and positively associated with the amount of information recalled during the interview. Correctly responding to the *wh*-questions may indicate that the retrieval practice was effective for enhancing recall. Alternatively, correct responding could reflect the child's prior knowledge, and their increased likelihood of correctly recalling more information regardless of any intervention (see Alexander, Kulikowich, & Schulze, 1994).

Similarly, when children responded incorrectly to the lesson questions, their poorer recall may have reflected the children's limited prior knowledge and reduced likelihood of correctly recalling information. In contrast, children who answered lesson questions incorrectly may have created two representations of event-information; the incorrect content they verbalised, and the correct information they heard. The *production effect* suggests that when an individual produces information in speech they will better recall that information compared to what they have heard (MacLeod, 2011). When children were asked to retrieve the content they heard during learning, there may also have been competition from what they said, and recall of the spoken information likely won out.

Additional effects. Additional influences may also explain the limited effectiveness of the *wh*-questions on children's verbal recall. First, the open-endedness of the *wh*-questions may have had an effect. When children respond to open-ended questions, they can provide any number of responses, many of which may be incorrect. As a result, children may strengthen incorrect retrieval paths (Karpicke, Blunt, Smith, & Karpicke, 2014). Given the complexity of the information taught in the current study, the questions probably needed to scaffold the children towards providing a correct response more than they did.

The ineffectiveness of the *wh*-questions in improving children's recall could also be explained by retrieval-induced-forgetting (Anderson, Bjork, & Bjork, 1994). This account suggests, that when individuals encode interrelated content (e.g. names of fruit), the information that is rehearsed (e.g. apple) is better recalled than information that was never studied (e.g. names of insects). Comparatively, non-rehearsed information from the same category (e.g. banana) is worse recalled than content never studied. The current study

also found that when children practiced retrieving specific solar system content, by answering *wh*-questions, their recall of the content unrelated to the questions decreased compared to children who never answered questions. Findings suggest that *wh*-questions may not only be ineffective in enhancing recall, but may actually impair the recall of content not directly related to the questions asked.

Children's affect may also have influenced the effectiveness of the *wh*-questions. As children were primarily answering the questions incorrectly, they were receiving a lot of corrective feedback. Telling a child that their answer is incorrect can hurt children's confidence and self-image and create anxiety (Chin, 2006; Dreyfus et al., 1990). The effect of cognitive conflict, however, tends to differ based on the academic achievement level of the student. While high achievers typically benefit from adults highlighting disparities between their understanding and the correct information, low achievers show hindered progress following conflict (Dreyfus et al., 1990; Zohar & Aharon-Kravetsky, 2005). As the solar system lesson was designed to be difficult for the majority of children, the amount of corrective feedback given may have hampered the children's learning and recall.

Finally, the length of the delay in the current study may also have contributed to the null effects. While beneficial effects of *wh*-questions on children's recall have been found at the 1-day delay (Benjamin et al., 2010), greater benefits have typically been found after two weeks (Benjamin et al., 2010; Jant et al., 2014). It may be the case that, over this period, children were able to consolidate the new knowledge in order to create a more advanced understanding. The early interview, in the current study, may have caught children in the middle of assimilating and accommodating the information taught. Improvements in children's verbal recall may still have been under-construction.

Summary. The effectiveness of the *wh*-questions asked seemed to depend on the children's prior knowledge, the cognitive demands placed upon the children, and the scaffolding the questions provided. In naturalistic conversations, *wh*-questions may have been effective for supporting children's learning and recall because adults could tailor feedback to each child, reducing the cognitive demands children were placed under. Such settings also allow information to be targeted within each child's zone of proximal development, ensuring that new information is neither too easy nor too difficult. The nature of the current study, and the need for experimental control, removed the possibility for individualised scaffolding and this may have diluted the effectiveness of the *wh*-questions asked.

***Wh*-questions and gesture together**

When engaged in combination, the two types of scaffolding did not provide further benefits to what was seen individually. Due to the outlined reasons for why both gesture and *wh*-questions were not individually effective, there were unlikely to be additional benefits when the two forms of scaffolding were used in tandem.

Conclusions

Across the interventions initiated in this study, limited beneficial effects were found. When children observed gesture during learning, no improvements were found in the amount of information children recalled, and there were no improvements in children's recall of either facts or concepts. This likely occurred due to the overlap in mechanisms through which gesture, visual aids, and the interactive demonstrations exerted their effects. Observing gesture, however, did improve the accuracy of children's verbal report, possibly due to the gesture tying the information to the specific learning experience.

The subsequent studies in this thesis focused exclusively on influence of observing gesture on children's learning and recall. While previous research has fairly consistently found that observing gesture can improve children's problem solving, understanding, and recall, the current findings indicated that observing gesture may not always be an effective form of scaffolding. These disparate findings highlight the need for further research examining the boundaries to when observed gesture is and is not effective. In particular, it was hypothesised that, in the current study, the effectiveness of gesture was reduced due to the presence of other visual supports. Study 2 therefore examined the effect of observing gesture when visual aids were absent. In many educational settings, props may not be present and it is possible that, under such circumstances, observing gesture would play a more important role in aiding learning and recall. Study 2 also included both a short and long delay following learning in order to assess whether observing gesture could improve the maintenance of children's learning.

While this thesis moved away from the effects *wh*-questions may exert in scaffolding children's learning and recall, further research is also needed in this area. Future studies should manipulate the openness of the questions asked and the targeted nature of the feedback provided. When questions are more directive, children may be supported in providing correct answers, enhancing the effectiveness of the retrieval practice. Feedback could also be better tailored to the child by including sub-questions, which guide the children away from incorrect answers and towards appropriate understandings. It would then be interesting to compare the effectiveness of adults

addressing the child's specific misconception compared to providing the child with a generic standardised response. Manipulating both the questions and feedback may further illuminate when *wh*-questions are and are not effective.

While studies examining both gesture and elaborative talk highlight the robust nature of the memory benefits observed, the extent to which the findings apply to diverse learning settings are less clear-cut than originally presumed. When children were engaged in experiential learning during a science lesson, the benefits of both observed gesture and *wh*-questions were diluted or absent. Both educators and researchers need to be aware of how gestures and *wh*-questions are used, and what other forms of scaffolding are present. It may be the case that *wh*-questions are only effective when tailored to the individual child and that children can have too much of a good thing, when scaffolding techniques like visual aids and gesture are used in combination.

Chapter 7

Study 2: The role of hand gestures in supporting children's learning and recall of scientific concepts

When individuals communicate, they use more than just their words; they also use their hands. People often take the gestures accompanying speech for granted, implicitly producing and processing the information they contain. The presence of gesture, however, can enhance immediate comprehension and understanding, as well as memory for information across a delay (Hostetter, 2011). As in everyday conversations, gestures are also prevalent and beneficial in educational settings (e.g. Goldin-Meadow et al., 1999). The current study examined the influence of observing gesture on 7- to 9-year-old children's recall of complex scientific information across both short (1-day) and long (7-month) delays.

As previously discussed, the benefits of observing gesture are evident across a variety of learning tasks (see Chapter 4 for more information). In typical classroom settings, however, the information taught to children is unlikely to be as tightly packaged as it has been in laboratory-based studies. Extant research has demonstrated the influence of observing gesture on children's recall and problem solving by teaching a single concept linked to limited verbal content (e.g. Cook et al., 2013; Ping & Goldin-Meadow, 2008), or by requiring children to remember isolated pieces of verbal information (e.g. So et al., 2012; Thompson et al., 1998). Even when children are required to recall narrative information, the content is typically simple and familiar, such as a story about a character carrying out common tasks (e.g. Austin & Sweller, 2014; Macoun & Sweller, 2016). This study extended upon extant literature by examining whether observing a multitude of representational gesture would improve children's verbal recall of a rich and complex science lesson, made up of a variety of facts and concepts.

While studies have not yet examined the potential benefits that observing gesture may exert on *children's* recall of novel and difficult science content, some research has been conducted with adults. Kang et al. (2013) found that while adults who observed gesture accompanying a science lesson on mitosis did not show superior recall, they did exhibit enhanced depth and flexibility of their understanding compared to those who did not see gesture. Comparatively, a study by Stieff et al. (2016) found that those who observed gesture during a lesson on molecule structure were no better at drawing correct

molecules compared to those who simply read the instructional text. In both these studies, however, the diluted effectiveness of observing gesture may be explained by the presence of diagrams within the lessons. Visual aids were also present in Study 1 (Chapter 6), and it was hypothesized that this may have reduced the potentially favorable effects of observing gesture. In addressing this issue, the current study examined the benefits that observing gesture may exert on children's verbal recall of scientific content, when visual aids were not present.

During science lessons, adults often automatically gesture when explaining scientific ideas to children (Kang et al., 2015). Gesturing seems a natural component of science instruction, and is likely a helpful one. The imagistic gestural content likely disambiguates novel scientific terms (Thompson & Massaro, 1986), makes abstract scientific ideas more concrete (Alibali & Nathan, 2012; Goldin-Meadow, 2015), and provides an additional visual representation of the information to process and store without creating cognitive overload (Cook et al., 2013; Singer & Goldin-Meadow, 2005). Science lessons, especially the physical sciences, also tend to include a lot of spatial content, which gesture can be particularly effective in conveying (Alibali, 2005; Sauter et al., 2012). As well as aiding the recall of isolated information, gesture has also been found to improve adults' understanding of a story in its entirety (Cutica & Bucciarelli, 2008, 2011). Observing gesture may also aid children in creating a comprehensive and integrated understanding of the information taught.

An additional aim of this study was to examine whether observing gesture could enhance children's recall and reduce forgetting after a long delay. Existing findings typically assess children's recall and learning immediately after encoding, with a long delay considered a number of hours. Cook et al (2013) examined children's problem solving across a 24-hour period and found improvements in post-test performance across the delay when children had observed gesture during learning, compared to when gesture was absent. This finding suggests that gesture may play an important role in the consolidation of learning, but the effects following longer delays are yet to be examined.

The current study

This study examined whether observing gesture during a comprehensive, verbally rich lesson about the solar system would influence children's learning and recall of the information taught. Children, 7 – 9 years old, individually participated in a lesson including a variety of abstract concepts, terminology, and spatial content, either with or

without accompanying hand gestures. The next day, and again 7 months later, children were interviewed to assess their learning and memory of the information taught.

It was hypothesised that observing gesture would support children in recalling more information, more accurately. Children's separate recall of the facts and concepts taught was also examined. It was hypothesised that observing gesture at learning would improve both children's recall of the isolated facts as well as comprehensive explanations of the solar system information (i.e. concepts). The amount and accuracy of children's recall, as well as their separate recall of facts and concepts, was also expected to be superior at the 7-month delay if children observed gesture during learning. Finally, reductions in forgetting were also expected if children observed gesture during the lesson, compared to when they did not.

Method

Participants

Participating children were recruited from a local school in Wellington, New Zealand (children had not participated in Study 1). Discussion with teachers established that none of the children recruited had learnt about the solar system during the year. The final sample comprised forty-two children (five children were excluded due to an error in the interview questioning) 7 – 9 years of age, with an average age of 8.1 years ($SD = 0.5$). There were 21 children in the *gesture-observed* condition (10 males, 11 females) and 21 in the *gesture-not observed* condition (9 males, 12 females). Thirty children from the sample participated in the Time 2 interviews, approximately 7 and a half months later ($M = 234.70$, $SD = 5.94$ days). Attrition occurred due to children moving to a new school, illness, or not returning the second consent form. There were 17 children in the *gesture-observed* condition (10 males, 7 females) and 13 in the *gesture-not observed* condition (6 males, 7 females) at the 7-month delay.

All children had parental consent to participate in the study and assented to their own involvement. The study was approved by the School of Psychology Human Ethics Committee under delegated authority of Victoria University of Wellington's Human Ethics Committee.

Materials

Parental rating scales. In order to control for children's prior knowledge and interest in the solar system, parents completed two Likert scales (as in Study 1, Chapter 6). Parents rated their child's interest and knowledge in the solar system on 9-point scales,

from 1 ("not at all interested") to 9 ("extremely interested") and children's knowledge about the solar system was also rated on a 9-point scale from 1 ("limited understanding") to 9 ("advanced understanding"). One parent did not return their ratings.

Procedure

Parents of participating children completed the parental rating scales attached to the consent form. These ratings, as well as children's age, were taken into account when assigning children to the learning conditions. Parental ratings of children's interest and knowledge did not differ between the *gesture-observed* and *gesture-not observed* groups at either Time 1 or Time 2 (p's > .05). An independent samples t-test also revealed that there were no significant age differences across the two conditions at either Time 1 ($t(40) = 0.53$, $p = .601$, $d = .16$) or Time 2 ($t(28) = 1.88$, $p = .070$, $d = .71$).

Solar system lesson. As in Study 1, children individually learnt about the solar system with the first author. In this study, however, the mat and props were removed. The lesson contained the same content, though some of the verbal information and associated gestural content was altered to make the lesson clearer. The child and research assistant sat at a 90-degree angle and half the children heard the verbal information accompanied by hand gestures, while the other children heard the same information with the instructor sitting on her hands.

The lesson was divided into four sections. The first section provided children with background information about the solar system to ensure all children could understand the remainder of the lesson. The second section taught children the three requirements for an object to be considered a planet, and why dwarf planets are not considered planets. The third section taught children about the role of gravity in the solar system. The final section taught children about the differences between rocky and gas planets (see Table 7.1)[5]. The lesson took approximately 15 minutes and children were asked not to talk.

The gestures accompanying the verbal content included: 1) iconic gestures, conveying concrete objects or actions, e.g. creating a ball shape representing a planet, 2) metaphoric gestures, conveying abstract ideas, e.g. fists pulled inwards to represent gravity's pull, and 3) deictic gestures, e.g. pointing out planet locations relative to the sun (see Figure 7.1 for examples of the gestures used).

---

[5] The full script used in Study 2 and Study 3 is also included in Appendix C (pp. 201-206).

Table 7.1

Topics Taught During the Solar System Lesson

Topic	Content

What makes something a planet?	• Three rules for something to be a planet
	   1. Round shape
	   2. Orbit the sun
	   3. Clear orbit
	• Pluto and other dwarf planets do not have a clear orbit
	• Orbit sizes dictate time taken to orbit

Gravity	• Gravity is a pull • Bigger things have more gravity. • The sun's gravity holds the moving planet in place o E.g. swinging around a ball tied to string. • The black hole keeps the moving suns in place

Surfaces of the planets	• Four closest planets to the sun are rocky
	   o Rocky planets are solid so you can stand on them
	• Four furthest planets are gas
	   o Gas planets are gassy so you can't stand on them
	   o Pressure inside planet creates liquid centre and diamonds
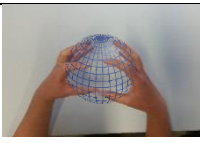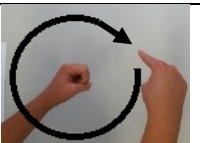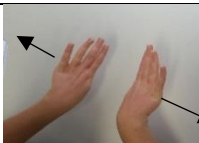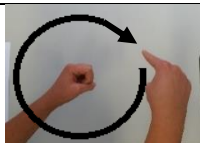
[One finger]		[Two fingers]		[Three fingers]		
The first rule	is that a planet has a round shape.	The second rule	is that it moves around the sun in a circle.	The third rule	is it has a **clear** pathway	as it orbits the sun.

Figure 7.1. Example of lesson gesture (scripted information included above has been condensed)

Interviews. Children were interviewed by different interviewers at two time points; the next day (two children were interviewed 2 days later due to illness) and approximately 7 and a half months later. The longer delay was, in part, determined by the availability of the children and the school once the need for a follow-up data was identified.

The interviews followed a standardized protocol. Children were first asked to recall everything they remembered about the time when "the lady came and taught you about the solar system," and this was followed by two open-ended prompts (e.g. "tell me some more things you remember about that time"). Children were then cued to recall information about each of the three main event topics. These questions were "I heard you learnt about what makes something a planet, tell me all about that,"; "I heard you learnt about why everything stays in the solar system, tell me all about that"; and, "I heard you learnt about the types of planets, tell me all about that." Each question was followed by two additional open-ended prompts (e.g. "Tell me more…"). If children reported information they had clearly obtained from another source (e.g. "my dad told me…") they were reminded that the interviewer only wanted to hear what they learnt during the event with the woman. Children received a small gift for participation, and each classroom received a $20 voucher.

Coding

Verbal coding. Coding was carried out using the same coding scheme as in Study 1 (see Chapter 6), with slight changes made related to the changes in the script. The lesson information was divided into 159 items. Coders identifying whether each unit of information was *present, absent* or *developing*. *Developing* codes were assigned when children showed understanding of the concept but: a) used terms incorrectly (e.g., "air planets"), b) reported partial information (e.g., "the sun has pull"), or c) reported information in an ambiguous way (e.g., "the first four planets"). See Study 1 (Chapter 6) for more information. For each of the 159 pieces of information children were given a recall score of 0 (absent), 1 (developing), or 2 (present).

Children's total recall was calculated as the sum of all scores across the four event components (i.e. total possible score of 318). For the three main event topics (i.e. requirements for a planet, gravity, and the surfaces of the planets), each of the 117 pieces of information was assigned to one of three categories: names, facts or concepts. The analysis of name information was not of interest within this study[6]. *Facts* involved describing the properties of the solar system and the objects in and around it (e.g. "planets move around the sun" or "the black hole has gravity"). *Concepts* included information that connected facts together and explained information (e.g. "planets move around the sun

_____

[6] See Appendix B, Supplementary Analysis S2.1 for information about the coding and analysis of children's recall of names (pp. 176-177).

because of the sun's gravity" or "the middle of the gas planets is liquid because of the pressure inside the planet"). Only the three main topics were divided into the categories as the introductory topic (i.e. background information) included content that was typically understood by the children prior to the event.

Category scores for facts and concepts were converted to percentages by adding scores across the relevant event components and dividing by the maximum score children could obtain (2 points for every relevant event component). There were 80 codes for facts (total score = 160), and 28 codes for concepts (total score = 56). Research assistants, blind to the hypotheses, conducted reliability coding. One research assistant coded 20% of the transcripts at Time 1, and the other coded 30% of the transcripts at Time 2. Inter-rater reliability (Cohen's Kappa) was calculated both when information was rated as present or absent at Time 1 ($k = .84$) and at Time 2 ($k = .78$), and when rated as present, absent or developing, both at Time 1 ($k = .82$) and Time 2 ($k = .75$).

Errors included intrusions and distortions (see Study 1 for more information). Repetitions and content unrelated to the solar system were not coded. The same research assistants conducted error coding on the same transcripts to establish reliability. Cohen's Kappa was calculated both at Time 1 ($k = .73$) and Time 2 ($k = .61$). Children's accuracy score was calculated from children's error scores (i.e. as proportion correct (total recall/total recall+errors *100).

Results

Preliminary analyses

Before examining the influence of the *gesture-observed* and *gesture-not observed* conditions on the amount (total recall), accuracy, and content (facts and concepts) of children's verbal recall, Kolmogorov-Smirnov tests of normality were conducted. At Time 1, children's accuracy and recall of concepts were not normally distributed. Accuracy scores were corrected by recoding a significant outlier (> three SD from the mean) while retaining rank order (see Field, 2013 for information about winsorizing scores). The recall of concepts was corrected with a square root transformation. At Time 2, children's recall of facts and concepts were not normally distributed. A square root transformation corrected the distribution of facts, but not concepts. Children's percentage recall of concepts was therefore analysed using non-parametric tests where possible, otherwise raw scores were assessed under the assumption that ANOVA is robust to violations of normality (Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). When analysing variables

across the two time points, if a square root correction was applied to a variable at Time 1 (1-day delay) it was also applied at Time 2 (7-month delay), and vice versa. This resulted in all distributions being normal except the recall of concepts at Time 2[7]. Raw means and standard deviations are reported.

Descriptive statistics

The lesson was divided into 159 items. Each item could receive a maximum score of two, resulting in a total possible score of 318. At Time 1, children were reporting close to one-third of the content ($M = 93.12$, $SD = 39.97$). When the number of items children reported was examined out of 159, children were also reporting almost a third of the information ($M = 50.19$, $SD = 21.38$). Children were highly accurate in what they recalled, with just over 86% of what they recalled being correct ($SD = 10.87\%$). While there was substantially more information that the children could report, the variability in responding suggests floor effects were not an issue.

At Time 2, however, 7 months had passed since the learning experience. At this point, out of a total possible score of 318, children were only reporting about one-seventh of the information ($M = 46.17$, $SD = 24.97$). Similarly, out of the 159 items, children were reported about one-eighth of the items ($M = 18.55$, $SD = 16.25$). This was a dramatic decrease in the amount of information the children recalled. Accuracy also decreased, as only 68% of the information children reported was correct ($SD = 17.34\%$). Again, there was large variability in the information recalled across the sample, suggesting that the effects of gesture were unlikely to have been washed out by the poor rate of responding. Children's minimal recall, however, may have hampered the ability to find subtle differences between the conditions.

Short delay

The first hypothesis, that children in the *gesture-observed* condition would recall more information with fewer errors compared to the *gesture-not observed* condition, was not supported. Independent samples *t*-tests revealed no significant differences between the conditions for total recall ($t(40) = -0.16$, $p = .871$, $d = .05$) or accuracy ($t(40) = 0.67$, $p = .507$, $d = .21$). Similarly, the children's recall of facts ($t(40) = -0.17$, $p = .864$, $d = .05$) and concepts ($t(40) = -1.22$, $p = .229$, $d = .38$), also did not differ across the conditions (see Table 7.2).

---

[7] See Appendix B, Supplementary Analysis S2.2 for information about Kolmogorov-Smirnov tests of normality for children's Time 1 and Time 2 verbal recall (p. 178).

Table 7.2

Means (Standard Deviations) for Verbal Recall at Time 1 and Time 2

	Total recall		Accuracy		Facts		Concepts	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Gesture not observed	92.10	42.62	87.23	70.11	20.93	8.37	10.86	2.48
	(48.81)	(23.75)	(9.26)	(15.13)	(12.73)	(6.40)	(10.38)	(3.72)
Gesture observed	94.14	48.88	84.97	67.09	21.52	9.17	13.01	2.83
	(29.84)	(26.24)	(12.39)	(19.21)	(9.04)	(6.87)	(8.85)	(3.67)

Long delay

The second hypothesis, that children would show superior recall after a long delay if they had observed gesture at learning, was also not supported. Independent samples t-tests revealed the two conditions did not differ 7 months later in their total recall ($t(28) = -0.68$, $p = .505$, $d = .25$) or accuracy ($t(28) = 0.47$, $p = .644$, $d = .17$). Again, children's separate recall of the facts and concepts taught were also analysed. Children's recall of concepts across the conditions was assessed using Mann-Whitney test, due to problems with the normality of the distribution. There was no difference across the *gesture-observed* and *gesture-not observed* conditions for either the recall of facts ($t(28) = -0.14$, $p = .892$, $d = .05$) or concepts ($U = 97.00$, $z = -0.59$, $p = .570$, $r = -.11$), see Table 7.2.

The hypothesis that children who observed gesture during learning would show decreased forgetting across the delay relative to children who did not observe gesture, was also not supported. Repeated measures ANOVAs were run with delay (i.e. verbal recall at Time 1 vs Time 2) as the within-subjects variable, and learning condition (i.e. *gesture-observed* versus *gesture-not observed*) as the between-subjects variable. Results indicated that there was a main effect of delay, with children recalling less correct information (total recall, and facts, and concepts) and less accurate information over time (all p's < .001). There were no significant interactions between condition and delay for total recall ($F(1, 28) = 0.89$, $p = .353$, $\eta_p^2 = .03$) or accuracy ($F(1, 28) = 0.09$, $p = .768$, $\eta_p^2 = .003$). Similar null effects were also found when the interactions between delay and children's recall of facts ($F(1, 28) = 0.27$, $p = .611$, $\eta_p^2 = .01$), and concepts ($F(1, 28) = 0.10$, $p = .761$, $\eta_p^2 = .003$) were separately examined, see Table 7.2.

Conclusions

Observing gesture did not influence the amount, accuracy, or type of content children verbally recalled at either time point. This finding was unexpected, as previous studies have found that children who observe gesture during learning typically exhibit superior comprehension, recall, and problem solving compared to children who have not observed gesture. The developmental stage of the children in the current study, combined with the nature of the learning experience, may offer a potential reason why observing gesture did not improve children's verbal recall.

Across development, children become more proficient at amalgamating gesture and speech to form a cohesive understanding. By age five, children can integrate both matching and mismatching information presented in gesture and speech (Kelly, 2001; Sekine et al., 2015), and, in adulthood, individuals have been found to do this with such proficiency that they recall gestural information as if it was presented in speech (Kelly et al., 1999). In middle childhood, however, children seem to go through a transitional developmental stage. At approximately 9 to 10 years old, children appear to encode both the verbal and gestural content imparted, but find it difficult to convert information across different representational formats. Church et al. (2000) found that 9- to 10-year-old children were able to retrieve the gestural content encoded if they were tested non-verbally, but not when tested verbally. The younger children, however, could combine the information across the modalities, converting gestural information into speech and vice versa.

The participating children in the current study were typically younger than 9 years of age, and were therefore not expected to be effected by this developmental transition. Previous studies with 7- to 10-year-old children have found that observing gesture during learning has improved both children's problem solving (Cook et al., 2013; Singer & Goldin-Meadow, 2005) and verbal recall (Thompson et al., 1998). If gesture was helpful for children of a similar age in previous studies, why was it not helpful in the current study?

It is commonly understood that development is typically not a series of steps, but rather a gradual progression over time (Berk, 2009). Competency may be apparent in simple tasks but take longer to emerge when nested within more complex cognitive operations (e.g. Deloache & Marzolf, 1995). Difficult tasks can therefore result in individuals revealing subtle deficits that may not have been otherwise observed. In previous studies, where beneficial effects of gesture on recall and problem solving have

78

been found, the gestures observed were typically repeated a number of times with the same verbal content (Cook et al., 2013; Ping & Goldin-Meadow, 2008; Singer & Goldin-Meadow, 2005). This may have created strong links across the modalities, possibly making it easier for the children to convert the gestural content into speech and vice-versa. In this study, however, children were taught a variety of facts and concepts, associated with a number of different and fleeting gestures. It may be the case that when children were required to recall the complex narrative information that they were taught in the current study, subtle deficits in combining gestural and verbal information became more apparent.

It could be hypothesised that children's own gesture production during the interviews in the current study might have acted as a form of non-verbal recall. When children recall visuospatial content in their gesture, they might receive visual and proprioceptive feedback (Cooperrider et al., 2015). This may then facilitate children in organising and chunking the visuospatial information in preparation for verbalising the content (Kita, 2000; Kita et al., 2007; Wagner et al., 2004). Children's gesture production may bridge the gap between observing gesture at learning and improved verbal recall during the interview.

An alternative hypothesis, however, is that the observed gesture in the current study was redundant. When gestural content is complementary (as it was in the current study), rather than novel, the beneficial effects gesture exerts on learning and memory are typically found to be smaller (Hostetter, 2011). The current study also taught children interconnected pieces of information, rather than isolated sentences or words. This may have meant children had sufficient verbal cues to access their learning without the need for gestural content to spur further recall.

Supplementary analyses were required to disentangle whether the observed gestures were ineffective in the current study, or whether the effectiveness of gesture was diluted due to the developmental stage of the participating children and the task difficult. Analyses were carried out to examine whether children's own gesture production may form an intermediary step between encoding gesture at learning and being able to utilize the gestural information at recall. The influence producing gesture may exert on children's recall and comprehension was examined at both the short and long delays.

Study 2B: Children's gesture production and verbal recall

Children's own gesture production may also act as a form of scaffolding. When tasks fall within children's zone of proximal development, children appear to gesture more

in order to propel themselves to higher levels of understanding (Bates & Dick, 2002). For example, when infants showed moderate task comprehension they were more likely to imitate familiar gestures as a form of scaffolding than when their task comprehension was either high or low (Shore, Bates, Bretherton, Beeghly, & O'Connell, 1990). Children appear to be implicitly aware of the benefits gesturing can offer as they grapple with understanding and describing information.

When children do produce gestures during retrieval, their verbal recall typically improves. Findings indicate that when children were instructed to gesture, they recalled more information about a story they had heard (Cameron & Xu, 2011), or an event they participated in (Stevanoni & Salmon, 2005) compared to when children did not gesture. Children also showed superior recall of novel verbs when they were instructed to gesture versus simply permitted to gesture (de Nooijer et al., 2013). By gesturing children are able to offload information into their hands (Cook, 2011; Pouw et al., 2014). They can then use the visual and proprioceptive feedback to assess new relationships between the content and organise their verbal report (Alibali et al., 2000; Kita, 2000; Kita et al., 2007). Children's own gesture production may even act as a retrieval cue (Frick-Horbury, 2002).

Despite the benefits of performing gesture, it is rare for studies to examine the interaction between observing gesture during learning and producing gesture at retrieval. Studies typically require *all* participants to observe gesture during encoding and then manipulate children's gesture production (Cameron & Xu, 2011; Cook et al., 2008; de Nooijer et al., 2013). Alternatively, some studies have examined the influence of gesture production without first modelling gesture (Sauter et al., 2012; Stevanoni & Salmon, 2005). Two studies, however, have examined the interaction between observing and performing gesture. First, Cook and Goldin-Meadow (2006) manipulated both the gesture children observed and the gesture children produced, but all during encoding. They found that the 9- to 10-year-old children, who saw gesture while learning how to solve mathematics problems, mimicked the instructor's gesture during the lesson. The gestural mimicry was then associated with improved mathematics problem solving. The second study, conducted by Austin and Sweller (2014), manipulated the gesture children observed while encoding a story, and examined children's gesture production while verbally recalling the narrative. They found that 3- to 4-year-old children did *not* mimic the gesture they observed. The differences found across the studies may reflect the different ages of the participating children. Middle childhood may be a special developmental period during

which children's gesture production potentially forms an intermediary step between observing gesture at encoding and improved recall.

Finally, it was also of interest to examine the influence of performing gesture across a longer delay. Studies have found that when children perform gesture the benefits tend to be long lasting. Cook et al. (2008) found that children who gestured during learning were more able to retain their learning across a four-week delay and Stevanoni and Salmon (2005) found that when children gestured while recalling an event that occurred two weeks earlier they also recalled more. Longer delays have not been examined to date.

The current study

Additional analyses were carried out to examine whether children's gesture production mediated the relationship between observing gesture at learning and improved verbal recall at both the short (1-day) and long (7-month) delays. This involved coding and analysing children's gesture production during the Time 1 and Time 2 interviews.

It was hypothesised that the reason children were not benefiting from observing gesture was not a result of the gesture being redundant. Instead, it was speculated that the complexity of the narrative content the children were taught revealed subtle deficits in the 7- to 9-year-old children's ability to combine information across the modalities. For the children to access the gestural content observed it was hypothesised that they would need to gesture themselves in order non-verbally retrieve the content before it could be converted into speech.

In examining children's gesture production as a potential mediator, it was of interest to assess: 1) whether children mimicked the gestures they observed during encoding, at both the short delay and the long delay, 2) whether children's gesture production during the short delay interview predicted children's verbal recall during the same interview, and the interview 7 months later, and 3) whether children's gesture production at the 7-month interview predicted how much the children verbally reported during the same interview.

At the 1-day delay, it was hypothesised that children in the *gesture-observed* condition would produce more gestures similar to those they had observed than children in the *gesture-not observed* condition. Increases in such gesture production were expected to be associated with increases in the amount and accuracy of children's verbal recall within the same interview. Similarly, a positive association was also expected between children's gesture production, and their separate recall of facts and concepts within the interview. Children's gesture production during the short delay interview was therefore expected to

81

mediate the relationship between observing gesture at learning and children's verbal recall during the same interview the next day.

As the 7-month delay was substantially longer than in previous literature, no specific hypothesis was made about whether children in the *gesture-observed* and *gesture-not observed* conditions would differ in their gesture production. Similarly, no specific hypothesis was made about whether children's Time 1 gesture production would predict verbal recall 7 months later. It was hypothesised, however, that children who produced relevant gestures during the delayed interview would recall more information, more accurately, and separately recall more facts and more concepts.

Method

Procedure

Gestural coding. Following Chui (2014), children's gesture during the Time 1 and Time 2 interviews were coded with children's verbal interview responses present. A gesture was defined as beginning when the hands were moved away from a position of rest to perform a stroke or action and then returned to rest. Children may have performed more than one gesture without a rest in between, however. In these cases, a gesture was coded as new when it had changed form, location, or function. Two different research assistants, blind to the children's condition, each coded approximately 25% of the videoed interviews to establish reliability. For the Time 1 interview data, 24% of the interviews were reliability coded, and 30% were reliability coded at Time 2.

Coders first established the presence or absence of representational gestures. *Representational* gestures were only coded if they accompanied speech and were iconic, metaphoric, or deictic. Gestures were classified as absent when: 1) both coders identified a beat or counting gesture (where the child successively held up fingers) or, 2) a phrase of three or more words, which included verbally codable content, was identified by both coders as being void of representational gesture. Coders percentage agreement (and Cohen's Kappa) at Time 1 was 93% ($k = .80$) and 89% at Time 2 ($k = .70$).

Coders also identified whether the representational gestures were event-relevant or novel. A representational gesture was coded as event-relevant if it was the same or similar in form and content to the scripted gestures observed during learning, and produced with similar verbal content[8]. Novel gestures were coded when the child produced: 1) relevant

---

[8] See Appendix D for more information about the gesture coding manual used in Study 2 and Study 3 (pp. 207 – 211).

verbal content with gesture that differed to those scripted, 2) similar gestures to those scripted but with different verbal content or, 3) novel verbal and gestural information. For the gestures both coders identified as representational, coder agreement as to whether the gesture was event-relevant or novel was established at both Time 1 with 86% agreement (k = .62) and Time 2 with 80% agreement (k = .61).

Results

Short delay

Preliminary analyses. Both children's representational (Rep) gesture production and event-relevant (ER) representational gesture production were assessed. While the representational gesture category included all informative gesture (regardless of content), the event-relevant gesture category included gestures similar in form and function to the scripted gestures children observed throughout the lesson.

Prior to analyzing children's representational and event-relevant gesture counts, a significant outlier (> 3 SD from the mean) was identified for both measures and was recoded (see Field, 2013). Pearson's bivariate correlations were then conducted between children's raw gesture counts and the number of words spoken by the children during the interview. There was a significant correlation between children's interview word count and their representational gesture count ($r(40) = .31$, $p = .045$). In order to account for the children's verbosity and opportunities to gesture, representational gesture production was examined as a rate per 100 words. Children's event-relevant gesture count did not correlate with their interview word count, however ($r(40) = .28$, $p = .070$). Given there was no need to control for the children's verbosity, event-relevant gesture production was analyzed both as a raw count, and as a rate per 100 words. The raw count provides a measure of the number of times content was offloaded into imagistic hand movements, while the gesture rate provides a measure of the consistency of gesture throughout the interview. For example, while two children may each have produced 15 gestures, if one child spoke 100 words and another child spoke 500, their count would be the same but the density of their gestures throughout the interview would be different.

This study examined three gesture measures: 1) representational gesture rate, 2) event-relevant gesture count, and 3) the event-relevant gesture rate. The distributions of the gesture measures were assessed using the Kolmogorov-Smirnov test of normality[9]. The

---

[9] See Appendix B, Supplementary Analysis S2.3 for information about Kolmogorov-Smirnov tests of normality for children's Time 1 and Time 2 gesture production (p. 178).

distributions of the event-relevant count and event-relevant rate were significantly different from normal and were corrected with square root transformations. Raw means and standard deviations are reported in Table 7.3.

Table 7.3

Means (Standard Deviations) for Gesture Production at Time 1 and Time 2

	Representational rate		Event-relevant count		Event-relevant rate	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Gesture not observed	1.71	2.58	5.71	N/A	1.18	1.37
	(1.70)	(2.36)	$(6.90)_a$		(1.22)	(1.55)
Gesture observed	2.73	1.96	11.19	N/A	2.06	0.78
	(2.10)	(2.32)	$(10.04)_b$		(1.60)	(0.94)

Note. Conditions with different subscripts were found to significantly differ.

It was hypothesised that children who observed gesture during the lesson would produce representational gesture at a greater rate during the 1-day delay interview, as well as higher rate and count of event-relevant representational gestures. Three *t*-tests were conducted to analyse this assertion. It was found that the *gesture-observed* and *gesture-not observed* conditions did not significantly differ in their representational gesture rate $t(40) = -1.73$, $p = .091$, $d = .53$. The *gesture-observed* condition did, however, produce a significantly greater number (i.e. count) of event-relevant gestures ($t(40) = -2.07$, $p = .045$, $d = .64$) and there was also a trend for a higher event-relevant gesture rate ($t(40) = -1.93$, $p = .061$, $d = .59$) than the *gesture-not observed* condition. There were indications that observing gesture influenced children's production of gestures that were similar to those they had observed at learning.

It was then assessed whether children's event-relevant gesture count and event-relevant gesture rate mediated the relationship between learning condition and children's verbal recall during the same interview. To avoid running many regressions, two observed variable path models were created using Analysis of Moment Structures (AMOS Ver.22)[10]. The first model included learning condition (i.e. *gesture-observed* versus *gesture-not observed*) as the predictor, children's event-relevant gesture *count* as the mediator, and children's verbal recall (i.e. total recall, accuracy, facts, and concepts) as the

---

[10] See Appendix B, Supplementary Analysis S2.4 for correlation outputs between gesture production during the Time 1 interview and verbal recall during the same interview (pp. 178 – 179).

outcome measure. The model was created with an embedded bootstrap for the mediational analysis. The size of the indirect effect was estimated with a 95% bias corrected confidence interval computed on the basis of 2,000 bootstrapped iterations with an estimand (see Figure 7.2).

It was found that children's gesture count significantly mediated the relationship between learning condition and the recall of concepts $a*b = 0.29$, $se = 0.22$, 95% CI [0.002, 0.89], $p = .048$. Children's event-relevant gesture count did not significantly mediate the relationship between learning condition and either children's total recall, recall accuracy or recall of facts ($p > .05$), see Figure 7.2. When children observed gesture during learning, they produced more event-relevant gestures, which was then associated with improved recall of concepts.
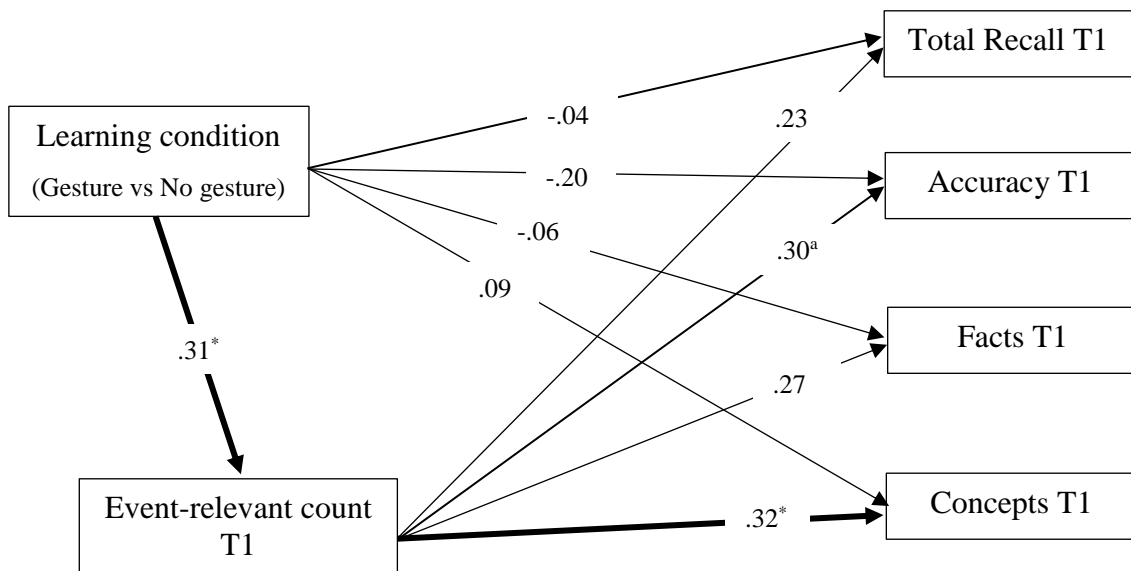
Figure 7.2. Model of event-relevant gesture count at Time 1 mediating the relationship between learning condition and verbal recall at Time 1. Standardised regression coefficients are reported.

$^{*}p < .05$ $^{a}p < .06$

The same model was also created including event-relevant gesture *rate* as the mediator. Children's event-relevant gesture rate significantly mediated the relationship between learning condition and children's accuracy during the interview $a*b = 2.20$, $se = 1.65$, 95% CI [0.05, 6.99], $p = .045$ (see Figure 7.3). No other mediational effects were significant ($p > .05$). When children produce event-relevant gestures at a higher rate throughout the interview, their verbal report was more accurate.
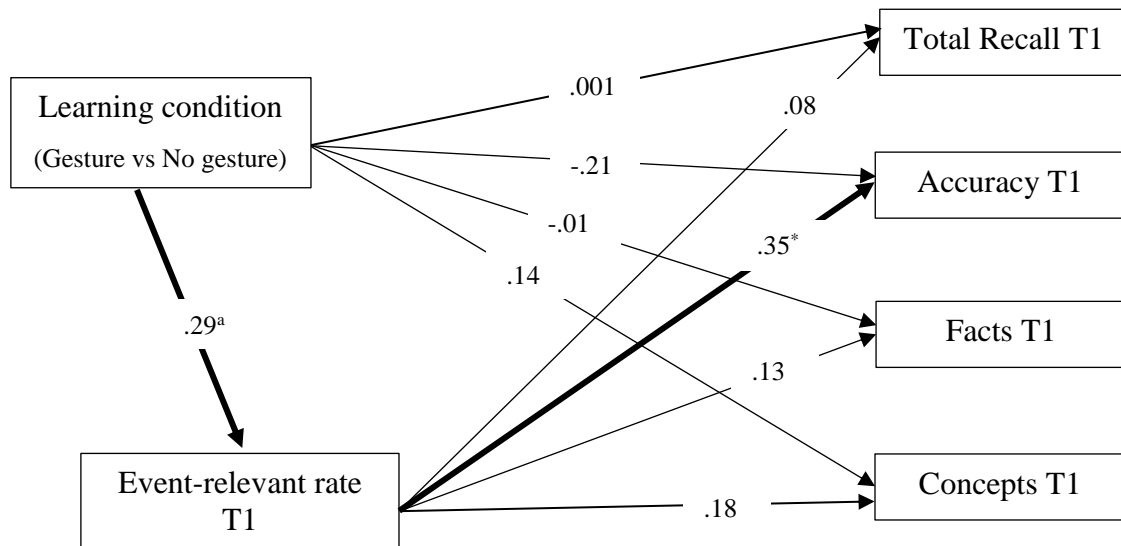
Figure 7.3. Model of event-relevant gesture rate at Time 1 mediating the relationship between learning condition and verbal recall at Time 1. Standardised regression coefficients reported.

$^{*}p < .05$ $^{a}p < .06$

No indirect effects were found when the mediator and outcome variable were swapped, for either the event-relevant count model or the event-relevant rate model (*p*'s > .05). It also makes theoretical sense that gesture aids the production of appropriate speech, rather than vice-versa, as gesture has consistently been found to slightly precede speech (Chu & Hagoort, 2014; McNeill, 2005; Morrel-Samuels & Krauss, 1992). Additionally, both the Free Imagery Hypothesis (Krauss et al., 1996; Krauss et al., 2000) and the Interface Hypothesis (Kita & Özyürek, 2003; Kita et al., 2007), suggest that gesture influences the way speech is produced.

Long delay

Children's Time 2 raw representational and event-relevant gesture counts were found to be significantly correlated with children's Time 2 interview word count (Representational count: $r(28) = .56$, $p = .001$; Event-relevant count: $r(28) = .43$, $p = .016$). To control for the child's verbosity and opportunities to gesture, only children's gesture rates per 100 words at Time 2 were examined. Kolmogorov-Smirnov tests of normality revealed that the distribution of the representational rate was not normally distributed, so

this was corrected with a square root transformation[11]. Raw means and standard deviations are reported in Table 7.3.

It was assessed whether children in the *gesture-observed* and *gesture-not observed* conditions would differ in their gesture production rates seven months after learning. Independent samples *t*-tests revealed, however, that the two learning conditions did not significantly differ in either their rate of representational gesture production ($t(28) = 0.66$, $p = .513$, $d = .24$), or rate of event-relevant gesture production ($t(28) = 1.28$, $p = .213$, $d = .47$), 7 months later, see Table 7.3.

Next, it was examined whether children's gesture production at the 1-day delay was associated with children's verbal recall seven months later. Pearson's bivariate correlations revealed no significant associations between children's Time 1 gesture production and children's verbal recall 7 months later (see Table 7.4).

Finally, it was hypothesised that children who produced representational gestures at Time 2 would verbally recall more information in the same interview. Correlations revealed that the hypothesis was partially supported. While children's representational gesture rate was not associated with improved recall during the same interview, the higher children's rate of event-relevant gestures at Time 2 the more solar system facts they recalled during the same interview (see Table 7.4). Due to the cross-sectional nature of the data, directionality of the effect could not be inferred.

Conclusions

In support of the hypotheses, it was found that when children observed gesture they were more likely to reproduce similar gestures during the short delay interview and this predicted greater recall of concepts and superior accuracy in the same interview. At the long delay however, having observed gesture did not influence children's gesture production, and their earlier gesture production did not predict later recall. When children did produce event-relevant gestures at Time 2, however, they recalled more facts during that interview.

---

[11] See Appendix B, S2.3 for information about normality (p.178).

Table 7.4

Correlations Between Children's Gesture Production at Time 1 and Verbal Recall at Time 2

	T1 ER count	T1 ER rate	T2 Total recall	T2 Accuracy	T2 Facts	T2 Concepts	T2 Rep rate	T2 ER rate
T1 Representational rate	.91**	.92**	.04	-.09	-.08	.14	.45*	.15
T1 Event-relevant count		.95**	.15	.02	.06	.24	.43*	.17
T1 Event-relevant rate			.13	.08	.01	.19	.37*	.16
T2 Total recall				.41*	.86**	.74**	.03	.16
T2 Accuracy					.55**	.19	-.22	.15
T2 Facts						.63**	.07	.36*
T2 Concepts							.20	.18
T2 Representational rate								.75**

* $p < .05$ ** $p < .01$

Interestingly, the effects of producing event-relevant gestures at the short delay appeared to differ based on how gesture production was examined. When children's event-relevant rate was included as a mediator, an indirect relationship was found between observing gesture at learning and accurate recall during the same interview. This may have occurred because children were consistently cueing themselves into correct and relevant information to talk about, and keeping themselves on topic (Frick-Horbury, 2002). The number of event-relevant gestures produced, however, mediated the relationship between observing gesture at learning and the recall of concepts. Regardless of how much the children did or did not say, producing more gestures may have increased the likelihood that the children were creating a comprehensive and accurate visuospatial model of the solar system, aiding their recall of the conceptual information taught (Alibali et al., 2000; Cutica & Bucciarelli, 2013). This supported the hypothesis that gesture (observed and produced) would aid children in creating an integrated and comprehensive understanding of what they were taught. While similar trends were generally found across the two models, the different mechanisms appear to tip the significance of the relationships.

Two potential reasons were proposed for why there was no direct effect of observing gesture on children's verbal recall: 1) children found it difficult to integrate the verbal and gestural information at recall, or 2) that the observed gestures were redundant. As hypothesised, the observed gestures did not seem to be redundant. The fact that mimicking the observed gesture enhanced children's recall of concepts and the accuracy of their verbal reports suggests that the observed gestures were beneficial, but their effectiveness was diluted. Instead, as hypothesised, it appears that the task difficulty may have revealed early developmental deficits in children's ability to verbally access gestural content in order to improve their recall. This finding mirrors what Cook and Goldin-Meadow (2006) found with slightly older children. The 9- to 10-year old children in their study showed limited direct improvement in mathematical problem solving when they observed gesture at learning, instead, what was most important for problem solving was children's production of the gestures they had observed.

Children's own gesture production appears to be the intermediate step between observing gesture and exhibiting improved verbal recall. The gestural mimicry may have allowed children to access the information they had stored visuospatially (Church et al., 2000). When children gestured, they may also have been able to use the visual and proprioceptive feedback to facilitate the conversion of the spatial information into speech,

improving the accuracy and depth of their verbal recall (Alibali et al., 2000; Cooperrider et al., 2015; Cutica & Bucciarelli, 2013; Kita, 2000; Wagner et al., 2004).

At the long delay, neither children's Time 1 nor Time 2 gesture production appeared to mediate the relationship between observing gesture at learning and improved Time 2 verbal recall. It is likely that the 7-month delay between Time 2 recall and both learning, and children's Time 1 gesture production, was too long. When children produced gesture in the moment, however, they recalled more factual information. While it is postulated that the gestures were supporting children in verbalising correct information (Kita, 2000) and providing retrieval cues (Frick-Horbury, 2002), it could also be suggested that the gestures were simply a side effect of understanding. These possibilities cannot be disentangled with the current paradigm.

While the mediational models provide an important framework for examining the data, the results must be interpreted cautiously. When the study was designed, it was of interest to examine the differences in learning and recall across the gesture-observed and gesture-not observed conditions. For this research question, a sample of 42 would be sufficient. When null effects were found, however, the hypothesis that children's gesture production may mediate the relationship between observing gesture and improvements in learning and memory was investigated. To examine the results in such a way, a larger sample would typically be used but this study was constrained by the previous research question. Further research is therefore needed to replicate these findings with a larger sample size. This is carried out in Study 3.

General discussion

These findings suggest that observed gesture may not be as influential for scaffolding learning as originally presumed. Observing gesture may only aid children's recall and comprehension in as much as it encourages children to gesture in correct ways. Before recommendations can be made, studies need to examine the potential benefits of asking children to gesture when they have or have not seen gesture during learning. Research to date suggests that when children are instructed to gesture during recall, it is beneficial (Sauter et al., 2012; Stevanoni & Salmon, 2005). Study 3 manipulated both whether or not children observed gesture during learning and their gesture production at recall. This would help tease apart the individual and interactive contributions of observed and produced gesture within a difficult, discursive lesson.

This study suggests that children's own gesture production may form a bridge between encoding the gestures observed and reaping the benefits during recall. During science instruction, whether this be informal parent child conversations or formal education settings, it is important for adults to consider how they are moving their hands, as these movements likely set the stage for children's learning and recall.

Chapter 8

Study 3: The interactive influence of observing gesture and producing gesture

Gesture is an integral part of communication, and serves many purposes. Gesture can provide information to the listener, but also plays an important role in aiding the process of speaking. As described previously, adults may scaffold children's learning and recall through their non-verbal gestural support but equally, children's own gesture production can also support their thinking and recall of information.

The findings from Study 2 suggested that whether or not 7- to 9-year-old children observed gesture at learning had no direct effect on children's verbal recall the next day. Observing gesture at learning, however, did increase children's gesture production, which was then positively associated with improved verbal recall during the same 1-day delayed interview. These results indicated that children's own gesture production played an important role in enhancing their learning and recall, but only after a short delay. To systematically examine the role children's own gesture production may play in enhancing learning and recall, this study manipulated the gesture that 7- to 9-year-old children produced when recounting what they had learnt during the solar system lesson. The influence of manipulating gesture, both observed and produced, was again examined at both short (1-day) and long (2-weeks) delays.

Manipulating children's gesture production

Research consistently demonstrates a positive effect of children's and adults' gesture on the amount of verbal information reported (even when controlling for verbosity). For example, when 3- to 4-year old children and adults heard a story about the actions conducted by a Lego man as he walked a particular path, the more they gestured while recalling the story, the more information they reported (Austin & Sweller, 2014). Similarly, when 8- to 10-year-old children and adults were asked to recall the spatial locations of toys in a room, the participants who gestured gave better quality spatial information (Sauter et al., 2012).

Given that gesture rates typically increase with age (Colletta et al., 2010), younger children may not be reaping all the benefits that spontaneous gesturing can exert on recall. Fortunately, instructing children to gesture can also increase children's gesture rate and verbal recall compared to when children are simply permitted to gesture. For example, when children were asked to verbally recall spatial locations, the 8-year-old participants

spontaneously reported significantly less information in both their speech and their gesture than the adult participants did. When a different group of 8-year-old children was *instructed* to gesture, however, children's reporting of spatial information reached near adult levels. Likewise, when recalling a personal experience, 6- to 7-year-old children who were instructed to gesture were found to gesture more, and correctly recall more information than those simply permitted to gesture (Stevanoni & Salmon, 2005). Collectively, these findings indicate that instructing children to gesture may increase both children's gesture production and verbal recall compared to when children spontaneously gesture.

As well as examining the influence of instructed versus spontaneous gesture production, the current study also examined the effect that restricting gesture may have on children's verbal recall. Including a gesture-restricted interview condition in the current study was important for two reasons. First, in educational settings there are many instances where children may be unable to gesture. For example, written tests are not conducive to gesturing as the hands are occupied with writing. As research has found that gesture production is important for producing fluent speech (Rauscher et al., 1996), and reducing cognitive load (Goldin-Meadow et al., 2001), it is important to examine the potentially detrimental effects restricting gesture may have on recall. Second, it is essential to include appropriate comparison conditions. Studies examining the effectiveness of gesture typically include only a gesture-instructed and a gesture-restricted condition. In these cases, studies have found that children who gestured show superior recall of a story (Cameron & Xu, 2011), or performance on mathematical equivalence tasks (Broaders et al., 2007). Such comparisons may, alternatively, reflect impaired performance by children when gesture was restricted, rather than improved performance by children who were encouraged to gesture.

In this study, children ($7 - 9$ years of age) were again taught complex scientific information and were asked to recall what they had learnt during an interview the next day. During the 1-day delay interview, however, children were either instructed to gesture, allowed to gesture, or were prevented from gesturing. Children in this age range are likely to spontaneously gesture at a low rate (see Sauter et al., 2012; Stevanoni & Salmon, 2005), and so a key question was whether children's recall could be improved if they were specifically instructed to gesture when recalling the lesson. Comparison conditions included both a gesture-allowed and a gesture-restricted condition. The gesture-restricted condition meant that any potentially detrimental effects on recall that occurred when hand

movements were restrained could also be examined. While the effects of manipulating children's gesture production during retrieval have previously been examined, these studies required children to recall familiar and simple information. The current study extended upon previous research, by manipulating children's gesture production as they recalled a complex and novel lesson.

Interactive effects of observing and producing gesture

Further study is also needed to examine the potentially interactive effects of observing gesture at learning and producing gesture at retrieval. When studies have examined interactions between observing and performing gesture, however, they typically do so within the learning experience. For instance, Goldin-Meadow et al. (2012) manipulated the gesture 6-year-old children observed and performed while learning how to complete mental rotation tasks. Children observed either "move" or "point" gestures, and performed either "move" or "point" gestures. Authors found that the act of performing "move" representational gestures was most important for learning, regardless of the type of gesture observed. Similarly, when adults were taught about molecule representations, those who mimicked the observed gesture during the lesson performed significantly better on the post-test than adults who only observed gesture or read the instructional text (Stieff et al., 2016). Of note, neither of these studies included a condition during which gesture was performed but not observed. It may be the case that the observed gesture indirectly affected problem solving by highlighting and enhancing the effectiveness of the produced gesture.

The gesture individuals observe can influence the gestures they perform. For example, research has found that adults will often reproduce the gestures used by their conversational partner (Chui, 2014). One study teased apart the effect observing gesture may have on children's gesture production. Cook and Goldin-Meadow (2006) manipulated both the gesture 9- to 10-year-old children observed (present versus absent) and the gesture children produced (instructed or allowed) while learning how to solve mathematics problems. Results suggested that whether the children were instructed or allowed to gesture had no effect on problem solving. Instead, when children observed gesture during learning they mimicked the instructor's gesture (also during learning), which improved their post-test performance. While the children's own gesture production was still of primary importance, the gestures that the children observed influenced the relevance and accuracy of the gestural content they produced. The more informative the children's gesture, the better their understanding and problem solving.

Results from Study 2 (Chapter 7) also indicated that children who observed gesture during learning were more likely to reproduce similar gestures at retrieval, and that this improved their verbal recall. Previous research has found that the accuracy of the information contained in children's gestures is important for learning (Brooks & Goldin-Meadow, 2016; Goldin-Meadow et al., 2009). When children in Study 2 mimicked the richly informative gesture they observed, the accuracy and relevance of the hand movements they were producing appeared to improve their recall. Extending upon previous research, the current study manipulated both the gesture children observed at learning (i.e. gesture-observed versus gesture-not observed) and the gesture children were able to produce during *retrieval* the next day. This design examined whether the effectiveness of instructing, allowing, or restricting children's gesture during recall differed depending on whether or not children had observed gesture at encoding.

Effects of gesture over time

Research examining the influence of observing or producing gesture on children's verbal recall across delays is also lacking. As previously described in Study 2, studies examining the effects of *observing* gesture on recall typically conduct the post-test immediately after learning (e.g. Austin & Sweller, 2014; Macoun & Sweller, 2016; Thompson et al., 1998; Wermeskerken et al., 2016). When the influence of observing gesture was examined following a delay of 24 hours, however, it was found that 7- to 10-year-old children who observed gesture during learning actually showed improvements in their problem solving across the delay (Cook et al., 2013).

Research studies assessing children's gesture *production* across delays are similarly sparse; with only a handful of studies including delays longer than a few hours. One such study found that when children mimicked an instructor's gesture at encoding they showed greater maintenance of their problem solving performance across a four-week delay, compared to children who mimicked an instructor's speech (Cook et al., 2008). Another study found that children who performed gesture while recalling an experience that occurred two weeks earlier verbally reported more information than those who did not (Stevanoni & Salmon, 2005). Taken together, these studies suggest that children's gesture production is important for recall across delays. To date, however, studies have not examined the influence of children's gesture production, during the recall of a complex learning experience, on children's later verbal recall of the same experience.

The current study examined children's recall both the next day and two weeks later. Compared to Study 2, when recall was assessed a second time seven months after

96

encoding, this study included a shorter delay. Research indicates that there are likely critical windows of time during which memories can be retrieved. For example, Jones and Pipe (2002) found that children were able to recall a staged personal experience with similar proficiency both days and weeks after learning. Six months after encoding, however, the memory trace had significantly deteriorated. In Study 2, the memory trace of the lesson was likely decayed to the point where, even if the original depth of encoding was enhanced by the presence of gesture, retrieving the memory was still going to be difficult. To address this issue, the current study examined whether observed and produced gesture may reduce children's forgetting of the complex science information within a shorter delay interval.

The current study

Children (7 – 9 years old) either did or did not observe gesture during a complex scientific lesson about the solar system. The next day, and two weeks later, children were interviewed to assess their recall of the information taught. Children's gesture production during the short (1-day) delay interview was manipulated (i.e. gesture-instructed, gesture-allowed, or gesture-restricted), while at the long (2-week) delay all children were allowed to move their hands as they naturally would.

Hypotheses

Effects of manipulating observed gesture. First, at the short delay, it was hypothesised that the *gesture-observed* and *gesture-not observed* learning conditions would not differ in the amount or accuracy of the information reported. Second, across the delay, children in the *gesture-observed* condition were expected to show reduced forgetting compared to the *gesture-not observed* condition.

Effects of manipulating gesture production. Three hypotheses were made about the effects of manipulating children's gesture production. First, during the short delay interview, it was hypothesised that children in the *gesture-instructed* interview condition would produce both representational and event-relevant gestures at a greater rate than the *gesture-allowed* condition. Second, children in the *gesture-instructed* condition were expected to recall more information with greater accuracy than children in the *gesture-allowed* interview condition, who in turn were expected to show superior recall compared to the *gesture-restricted* condition. Third, across the delay, it was hypothesised that children who produced gesture in an earlier interview (*instructed* or *allowed*) would show reduced forgetting and less reduction in recall accuracy than children in the *gesture-restricted* interview condition.

Interactive effects. When considering both the gesture observed at learning and the gesture manipulated at retrieval, differential effects were expected across the six conditions. At the short delay, it was hypothesised that children's gesture production would be most important for recall, with children who were able to perform gesture (*instructed* or *allowed*) recalling more information, more accurately, than children in the *gesture-restricted* condition. Across the interview conditions during which children could gesture, those who had observed gesture were expected to show superior recall compared to those who had not. This was hypothesised due to the influence observing gesture may have on the relevance and accuracy of children's gesture production.

Second, across the delay, it was expected that whether or not children had seen gesture during learning would be most important for recall, due to the picture superiority effect (see Paivio, 1971). It was expected that the *gesture-observed* condition would show reduced forgetting compared to the *gesture-not observed* condition, but that in each learning condition those who were previously able to gesture (*instructed* or *allowed*) would show less forgetting than the children previously restricted from gesturing.

Exploratory analyses of children's recall of facts and concepts. Secondary analyses examined whether children's recall of the facts taught differed from their recall of concepts. When recalling facts, children simply needed to repeat information, whereas when children reported the concepts they had learnt, they were required to recall connections and explanations for the facts. Specific hypotheses were not made across the different learning and interview conditions, but it was expected that the presence of gesture would enhance children's ability to recall the isolated facts. It was also hypothesised that gesture, both observed and performed, would be especially helpful for scaffolding children's understanding of the more complex, conceptual information.

Method

Participants

Participating children were recruited from a local school in Wellington, New Zealand. As in Study 1 and 2, none of the children recruited had learnt about the solar system during the school year, or participated in an earlier study. All children completed the solar system lesson and both interviews (1 day and two weeks later). Two children were excluded from the final sample because they asked to end the first interview before its completion. The final sample comprised 84 children with an average age of 8.45 years ($SD = 7.11$ mths). See Table 8.1 for the sample distributions across the six conditions.

All children had parental consent to participate in the study and assented to their own involvement. The study was approved by the School of Psychology Human Ethics Committee under delegated authority of Victoria University of Wellington's Human Ethics Committee.

Table 8.1

Distribution of Males and Females Across the Experimental Conditions

Lesson gesture	Interview gesture	Number of participants		
		Males	Females	Total
Observed	Instructed	8	6	14
	Allowed	8	6	14
	Restricted	7	7	14
Not-observed	Instructed	7	7	14
	Allowed	9	5	14
	Restricted	8	6	14
Total		47	37	84

Procedure

Parents completed the parental rating scales described in Study 1 (Chapter 6). These ratings, as well as children's age, were taken into account when assigning children to the conditions. A factorial ANOVA revealed there were no significant differences in parental ratings of the children's interest in the solar system[12] or children's knowledge about the solar system[13] across the conditions. Children's age also did not differ across either the learning ($F(1, 78) = 0.02$, $p = .881$, $\eta p^2 < .01$), or interview conditions ($F(2, 78) = 0.17$, $p = .848$, $\eta p^2 = .004$).

Solar system lesson. Children participated in the same solar system lesson described in Study 2, with half the children observing gesture during the lesson and the other children simply hearing the same verbal information.

Interviews. Children were interviewed at two time points; the next day (one child was interviewed 2 days later due to absence) and again 2 weeks later ($M = 14.26$ days, *SD*

---

[12] Interest scores did not differ across the learning conditions ($F(1, 78) = 0.56$, $p = .457$, $\eta_p^2 = .01$) or interview conditions ($F(2, 78) = 0.51$, $p = .604$, $\eta_p^2 = .01$).
[13] Knowledge scores did not differ across the learning conditions ($F(1, 78) = 0.71$, $p = .401$, $\eta_p^2 = .01$) or interview conditions ($F(2, 78) = 0.37$, $p = .693$, $\eta_p^2 = .01$).

= 1.32). Different research assistants carried out the interviews at the two time points, using the same standardized protocol (described in Study 2, Chapter 7). Children were asked to freely recall what they learnt and were then asked cued recall questions, with a number of follow up prompts.

When children were interviewed at the 1-day delay, children were either instructed to gesture, allowed to gesture, or restricted from gesturing. When children were instructed to gesture, they were asked to use their hands and their words while recalling what they learnt (e.g. "I heard you learnt about what makes something a planet. Tell me all about that using your hands and your words."). Children who were allowed to gesture were given no instructions about moving their hands, and participated in exactly the same interview as conducted in Study 2 (Chapter 7). Children who were restricted from gesturing wore an apron tied around their waist, which held their hands flat within tight pockets (see Figure 8.1).

Two weeks later, children were interviewed again, but children were given no instructions about gesturing, and all children were free to move their hands. Children received a small gift for participation, and each classroom received a $20 voucher.

Figure 8.1. Picture of the memory apron used
in the gesture-restricted interview condition

Coding.

Verbal coding. The information conveyed in the lesson was divided into 159 items. For the three main lesson topics (i.e. requirements for a planet, gravity, and surfaces of the planets), each of the 117 pieces of information was categorized as a name, fact, or concept. Only children's recall of facts and concepts were analyzed, however[14].

---

[14] While the children's recall of names (i.e. scientific terms) were coded, the results are not reported in the main body of the text. See Appendix B, Supplementary Analysis S3.1 for the analysis of names (pp. 179 – 180).

A research assistant, unaware of the children's condition or the hypotheses of the study, coded approximately 24% of the interviews at both Time 1 and Time 2 to establish reliability. Coders identified whether each unit of information was *present, absent* or *developing*. Children received a score of 2-points for present information, 1-point for developing information, and 0-points for absent information, resulting in a total possible score of 318 (see Study 2 for more information). Inter-rater reliability (Cohen's Kappa) was calculated when information was rated as present or absent at Time 1 ($k = .87$) and at Time 2 ($k = .89$), and as present, absent, or developing, both at Time 1 ($k = .83$) and Time 2 ($k = .86$).

Error coding was conducted the same way as in Study 2. Reliability was established using the same transcripts as were used for verbal recall (described above). Cohen's kappa was calculated both at Time 1 ($k = .71$) and Time 2 ($k = .76$). Children's accuracy was calculated from their error scores (i.e. total recall/total recall+errors *100).

Gestural coding. Children's gesture production was coded for the children in the allowed and encouraged conditions at Time 1 (the restricted condition was not able to move their hands), and all children's gesture production was coded at Time 2. Gestures were coded without the associated audio content in order to obtain a pure measure of the gestures produced without the verbal information biasing their interpretation. Apart from the removal of the audio content, gestures were coded the same way as in Study 2. A research assistant, unaware of the children's condition, coded approximately 25% of the interviews (27% at Time 1 and 25% at Time 2) to establish reliability. Coders first established the presence or absence of representational gesture. *Representational* gestures were only coded if they were iconic, metaphoric, or deictic. Gestures were defined as absent when 10-30 seconds passed with neither coder identifying a gesture as having occurred. Percentage agreement at Time 1 was 83% ($k = .66$) and was 84% at Time 2 ($k = .69$).

Coders also identified whether gestures were event-relevant or novel. A representational gesture was coded as event-relevant if it was the same, or similar, in form to one of the 27 scripted gestures produced during the learning experience[15]. A representational gesture was coded as novel if it was qualitatively different from the scripted gestures. Of the gestures that both coders agreed were representational, inter-rater

---

[15] See Appendix D for more information about the gesture coding manual used in Study 2 and Study 3 (pp. 207 – 211).

agreement was calculated for whether the gesture was event-relevant or novel. At Time 1 percentage agreement was 90% ($k = .72$) and was 87% at Time 2 ($k = .63$).

Results

The 1-day delay

The individual and interactive effects of manipulating both the gesture children observed during learning and performed during retrieval were examined in terms of children's total verbal recall and accuracy (i.e. total/(total recall+errors)). Two 2 (Learning: gesture-observed vs gesture-not observed) × 3 (Interview: Gesture-instructed, gesture-allowed vs gesture-restricted) factorial ANOVAs were conducted. When children's total recall was assessed, both children's age and knowledge scores were included as a covariate (knowledge and interest were highly correlated so only knowledge was included). When accuracy was examined, only children's age was covaried, see Table 8.2. Prior to analyses, verbal recall scores which were significant outliers were recoded to maintain rank order (Field, 2013). One total recall outlier and four error outliers were recoded. Accuracy scores were calculated with these corrected totals.

The nature of the information that children recalled was also examined by submitting the percentage of facts and concepts recalled into two 2 (Learning: gesture-observed vs gesture-not observed) × 3 (Interview: Gesture-instructed, gesture-allowed vs gesture-restricted) factorial ANOVAs. Age and knowledge were included as covariates for children's recall of facts, but not their recall of concepts (see Table 8.2). The recall of concepts was corrected with a square root transformation[16], but raw means and standard deviations are reported.

Table 8.2

Correlations Between Age, Parental Ratings of Child's Interest, and Parental Ratings of Child's Knowledge and Children's Verbal Recall and Gesture Rates at Time 1

	Interest	Knowledge	Total Recall	Accuracy	Facts	Concepts	Rep gesture	ER gesture
Age	.01	.13	.25*	.35**	.31**	.08	.02	-.01
Interest		.60**	.30**	.23*	.26*	.10	.15	.17
Knowledge			.30**	.21	.34**	.19	.02	.04

* $p < .05$ ** $p < .01$

---

[16] See Appendix B, Supplementary Analysis S3.2 for Kolmogorov-Smirnov tests of normality for the Time 1 verbal recall measures (p. 181).

Descriptive statistics. On average, children were reporting about one-third of the content, both when examined as a score out of 318 ($M = 103.61$, $SD = 29.06$) or as the number items reported out of 159 (regardless of whether scores were developing or present; $M = 57.10$, $SD = 16.20$). While children's recall was generally at the lower end, what they recalled was predominantly correct (88% accurate), see Table 8.3. Given the size of the standard deviations for children's total recall, it is unlikely that floor effects were minimizing the chances of finding experimental effects.

Learning condition effects at the short delay. To examine the hypothesis that the *gesture-observed* and *gesture-not observed* learning conditions would not differ in their total recall or accuracy, the main effects of learning condition were examined. No significant main effects of learning condition were found for either children's total recall ($F(1, 76) = 0.08$, $p = .784$, $\eta_p^2 = .001$), or accuracy ($F(1, 77) = 0.73$, $p = .395$, $\eta_p^2 = .01$). When children's the percentage recall of the facts and concepts taught were examined across the learning conditions, again, no significant main effects were found (Facts: $F(1, 76) = 0.32$, $p = .572$, $\eta_p^2 = .004$; Concepts: $F(1, 78) = 2.39$, $p = .126$, $\eta_p^2 = .03$), see Table 8.3.

Interview condition effects at the short delay. Before assessing the main effects of interview condition on children's verbal recall, the effectiveness of the gesture manipulation in altering children's gesture production was examined. Again, both children's representational (Rep) gestures and event-relevant (ER) representational gestures were assessed[17]. While the representational gesture category captured children's production of semantic gesture, the event-relevant gesture captured the extent to which children were producing similar gesture to those used throughout the lesson. Given the significant correlations between the number of words spoken and children's production of both representational ($r(54) = .54$, $p < .001$), and event-relevant gestures ($r(54) = .51$, $p < .001$), gesture production *rates* were calculated to control for verbosity (i.e. (gesture count/ word count)*100). The distribution of children's event-relevant gesture rate was significantly different from normal, and was corrected with a square root transformation[18].

Independent samples *t*-tests were conducted to assess the hypothesis that children in the *gesture-instructed* condition would gesture at a greater rate than the *gesture-allowed*

---

[17] The variety of different event-relevant representational gestures children produced were coded and analysed, though no significant results were found. See Appendix B, Supplementary Analysis S3.3 for event-relevant gestural variety results (pp. 181 – 183).
[18] See Appendix B, S3.2 for Kolmogorov-Smirnov results for Time 1 gesture rates (p.181).

condition. Unexpectedly, the *gesture-instructed* and *gesture-allowed* conditions did not significantly differ in either their rate of representational gesture production ($t(54) = -1.08$, $p = .286$, $d = .29$), or event-relevant gesture production ($t(54) = -1.49$, $p = .143$, $d = .40$), see Table 8.4[19].

Next, it was hypothesised that children in the *gesture-instructed* condition would recall more information compared to the *gesture-allowed* condition, and that both gesture conditions would show superior recall to the *gesture-restricted* condition. The two main effects of interview condition revealed that whether children were instructed, allowed or restricted from gesturing had no significant effect on either the amount ($F(2, 76) = 2.10$, $p = .130$, $\eta_p^2 = .05$), or accuracy ($F(2, 77) = 1.98$, $p = .145$, $\eta_p^2 = .05$) of children's verbal recall.

The main effects of interview condition were also examined separately for children's percentage recall of facts and concepts. While there was no significant effect of interview condition on children's recall of concepts ($F(2, 78) = 2.04$, $p = .138$, $\eta_p^2 = .05$), interview condition did significantly influence the recall of facts ($F(2, 76) = 3.23$, $p = .045$, $\eta_p^2 = .08$). Helmert contrasts revealed that those who were restricted from gesturing reported significantly fewer facts than those who could gesture, either instructed or allowed ($p = .015$). The gesture allowed and instructed conditions did not significantly differ ($p = .603$). This supported the hypothesis that restricting gesture would hinder children's verbal recall during the interview.

Interactions between learning and interview conditions at the short delay.
Finally, it was hypothesized that children who observed gesture at learning *and* gestured at retrieval would recall more information with superior accuracy compared to the children who had either not observed gesture or had not gestured at retrieval[20]. To examine this hypothesis, interactions between the learning conditions and interview conditions were examined. There was no significant interaction for the amount of information children recalled ($F(2, 76) = 0.10$, $p = .906$, $\eta_p^2 = .003$), but there was a significant interaction for the accuracy of children's verbal report ($F(2, 77) = 3.99$, $p = .022$, $\eta_p^2 = .09$).

Post hoc analyses were conducted with one-way ANOVAs to examine children's accuracy across the conditions. It was found that children in the *gesture-not observed*

---

[19] Gesture counts were also analysed across the interview conditions. Results indicated that the *gesture-instructed* condition produced significantly more gestures than the *gesture-allowed* condition when analysed as a raw count. See Appendix B, Supplementary Analysis S3.4 (p. 183).

[20] The interaction between learning condition and interview condition was also examined in terms of children's gesture production rates. See Appendix B, Supplementary analysis S3.5 for results (p. 184).

conditions did not differ in their recall accuracy across the gesture-instructed, gesture-allowed and gesture-restricted conditions ($F(2, 38) = 0.65$, $p = .526$, $\eta_p^2 = .03$). In contrast, children in the *gesture-observed* condition, did significantly differ in their recall accuracy across the interview conditions ($F(2, 38) = 4.33$, $p = .020$, $\eta_p^2 = .19$). Helmert contrasts indicated that while children in the *gesture-restricted* condition did not significantly differ from those who were able to gesture (*allowed* or *instructed*; $p = .794$), the *gesture-instructed* condition verbally reported information that was significantly less accurate compared to the *gesture-allowed* condition ($p = .006$). The hypothesis that children who observed gesture and were instructed to gesture would provide the most accurate recall was not supported.

Interactions were also examined separately for the recall of facts and concepts. Interactions between the learning and interview conditions revealed no significant interactions for either the recall of facts ($F(2, 76) = 0.17$, $p = .845$, $\eta_p^2 = .004$) or concepts ($F(2, 78) = 0.11$, $p = .897$, $\eta_p^2 = .003$), see Table 8.3.

One-day delay summary. First, as hypothesised, children in the *gesture-observed* and *gesture-not observed* conditions did not significantly differ in their verbal recall. Second, and unexpectedly, children in the *gesture-instructed* and *gesture-allowed* conditions also did not differ in either their gesture rates or the amount and accuracy of their verbal recall. When the recall of facts and concepts was examined, however, children in the *gesture-restricted* condition recalled fewer facts than the other gesture conditions (instructed and allowed). Finally, the hypothesised interactions between the learning and interview conditions were also not found. Instead of observing and producing gesture improving the amount and accuracy of children's verbal reports, children in the *gesture-observed*, *gesture-instructed* condition were less accurate than the other *gesture-observed* interview conditions.

Table 8.3

Raw Means (Standard Deviations) for Verbal Recall at Time 1 and Time 2

Lesson gesture	Interview gesture	Total recall		Accuracy		Facts		Concepts	
		Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Observed	Instructed	105.71 (33.21)	87.64 (38.87)	81.60 (13.01)	82.98 (13.53)	36.52 (14.75)	29.52 (16.22)	16.15 (10.16)	11.45 (9.46)
	Allowed	108.93 (19.63)	105.00 (31.67)	91.23 (5.74)	88.88 (7.66)	37.39 (8.60)	35.74 (11.27)	20.38 (11.40)	15.75 (10.38)
	Restricted	96.14 (43.94)	83.29 (31.11)	87.64 (9.69)	89.42 (8.07)	31.14 (14.61)	27.26 (12.92)	14.86 (9.87)	12.40 (8.28)
	Total	103.60 (33.38)	91.98 (34.56)	86.83 (10.49)	87.09 (10.30)	35.02 (12.96)	30.84 (13.78)	17.13 (10.51)	13.20 (9.37)
Not-observed	Instructed	106.07 (29.98)	85.57 (29.19)	90.00 (5.97)	90.53 (4.33)	33.96 (9.94)	28.87 (10.45)	14.38 (9.72)	11.64 (8.10)
	Allowed	111.29 (22.80)	90.93 (22.74)	87.79 (5.88)	86.30 (10.68)	36.40 (7.32)	29.24 (8.96)	15.65 (8.39)	12.45 (6.05)
	Restricted	93.50 (16.71)	86.36 (20.82)	86.82 (8.55)	87.55 (8.44)	28.62 (11.05)	28.27 (11.45)	10.92 (7.76)	10.66 (6.56)
	Total	103.62 (24.40)	87.62 (24.03)	88.21 (6.87)	88.13 (8.24)	33.00 (9.89)	28.79 (10.09)	13.65 (8.69)	11.59 (6.83)
Total	Instructed	105.89 (31.05)	86.61 (33.75)	85.80 (10.81)	86.75 (10.58)	35.24 (12.41)	29.19 (13.39)	15.27 (9.80)	11.55 (8.64)
	Allowed	110.11 (20.91)	97.96 (27.99)	89.51 (5.96)	87.59 (9.22)	36.90 (7.85)	32.49 (10.53)	18.01 (10.11)	14.10 (8.51)
	Restricted	94.82 (32.65)	84.82 (26.02)	87.23 (8.98)	88.49 (8.16)	29.88 (12.78)	27.77 (11.99)	12.89 (8.94)	11.53 (7.38)
	Total	103.61 (29.06)	89.80 (29.66)	87.52 (8.84)	87.61 (9.29)	34.01 (11.50)	29.82 (12.05)	15.39 (9.74)	12.39 (8.19)

Table 8.4

Means (Standard Deviations) for Children's Gesture Rates at Time 1 and Time 2

Lesson gesture	Interview gesture	Represent Rate		Event-relevant rate	
		Time 1	Time 2	Time 1	Time 2
Observed	Instructed	5.68 (2.50)	4.57 (3.98)	4.31 (2.18)	3.66 (3.28)
	Allowed	4.85 (4.14)	5.51 (4.47)	3.70 (3.52)	4.51 (3.53)
	Restricted		2.73 (1.86)		2.08 (1.51)
	Total	5.27 (3.38)	4.27 (3.72)	4.01 (2.89)	3.42 (3.02)
Not-observed	Instructed	5.00 (3.41)	3.61 (4.24)	3.47 (2.27)	2.75 (3.50)
	Allowed	3.90 (3.26)	3.16 (3.34)	2.80 (2.43)	2.50 (2.55)
	Restricted		2.13 (2.11)		1.56 (1.72)
	Total	4.45 (3.32)	2.97 (3.32)	3.14 (2.33)	2.27 (2.67)
Total	Instructed	5.34 (2.95)	4.09 (4.06)	3.89 (2.22)	3.20 (3.36)
	Allowed	4.38 (3.69)	4.33 (4.06)	3.25 (3.00)	3.51 (3.19)
	Restricted		2.43 (1.97)		1.82 (1.61)
	Total	4.86 (3.35)	3.62 (3.56	3.57 (2.64)	2.84 (2.89)

Effects across the delay

Changes in children's verbal recall between the 1-day delay interview and the 2-week delay interview were also examined[21]. It was of primary interest to examine changes in the amount (total recall) and accuracy (i.e. total recall/ (total recall +errors)) of children's verbal reports, but children's separate percentage recall of the facts and concepts were also assessed.

At the long delay, children were again reporting about one-third of the content, both when examined as a score out of 318 points ($M = 89.80$, $SD = 29.66$), or the number of items reported out of 159 ($M = 49.27$, $SD = 15.82$). What the children reported was also predominantly accurate (88% correct). Compared to children's recall seven months after learning (in Study 2, Chapter 7), children's verbal recall was more detailed two weeks after encoding.

To examine the hypothesis that observing or producing gesture would differentially effect changes in children's total recall and accuracy across the delay, two mixed ANOVAs were conducted. The between-subjects variables were learning condition (gesture-observed versus gesture-not observed) and interview condition (gesture-instructed, -allowed, and -restricted), while the within-subjects variables were the amount and accuracy of the information reported across the delay (Time 1 versus Time 2). Significant outliers were again recoded, resulting in the recoding of one outlier for total recall, and four for errors (accuracy scores calculated using the corrected totals). The distribution of total recall was not normal and could not be corrected, so raw scores were analysed (Schmider et al., 2010).

Across the two time points there was a main effect of delay for children's total recall, ($F(1, 78) = 32.63$, $p < .001$, $\eta_p^2 = .30$), with children recalling less in the second interview. Children's accuracy scores did not significantly differ across the time points, however ($F(1, 78) = 0.02$, $p = .888$, $\eta_p^2 < .001$). The hypothesis that children would show reduced forgetting in the *gesture-observed* versus *gesture-not observed* condition was not supported, as there were non-significant interactions between learning condition and delay for both total recall and recall accuracy (p's > .05). The non-significant interaction between interview condition and delay (p's > .05) also did not support the hypothesis that performing gesture during an earlier interview would reduce forgetting. Finally, it was

---

[21] Analyses examining children's verbal recall and gesture production rates across the experimental conditions were also conducted specifically at the 2-week delay. See Appendix B, S3.6 for results and Kolmogorov-Smirnov tests of normality at the long delay (pp. 184 – 186).

hypothesised that the potential benefits of observing gesture would differ depending on whether or not children had previously performed gesture. There was, however, no significant interaction between delay, learning condition, and interview condition for total recall or recall accuracy (all p's > .05). See Table 8.3 for means and standard deviations.

To examine whether observing or producing gesture would differentially effect changes in children's recall of facts and concepts across the delay, two additional mixed ANOVAs were conducted. Prior to analyses, the distribution of concepts was corrected using a square root transformation[22]. Findings indicated that there was a significant main effect of delay, with children recalling fewer facts ($F(1, 78) = 38.65$, $p < .001$, $\eta_p^2 = .33$), and fewer concepts ($F(1, 78) = 10.65$, $p = .002$, $\eta_p^2 = .12$) in the second interview. Interactions between delay and both the learning and interview conditions were again separately examined, but no significant effects were found for either the recall of facts or concepts (p's > .05). Finally, the interactions between delay, learning condition, and interview condition were examined. There was a significant interaction for the recall of facts ($F(2, 78) = 4.27$, $p = .017$, $\eta_p^2 = .10$), but not concepts ($p > .05$), see Table 8.3.

After applying a Bonferroni correction (.05/6 = .008), a series of post-hoc paired samples t-tests examined the verbal recall of facts at Time 1 (1-day delay) compared to Time 2 (2-week delay). The results are presented in Figure 8.2. For the *gesture-restricted* interview condition, the recall of facts did not differ between Time 1 or Time 2, either when children observed gesture ($t(13) = 1.98$, $p = .069$, $d = .28$) or did not observe gesture at learning ($t(13) = 0.39$, $p = .706$, $d = .03$). When gesture was restricted, low levels of recall were maintained across both time points. For the *gesture-allowed* interview condition, those who had seen gesture at learning maintained a relatively high level of responding ($t(13) = 0.78$, $p = .452$, $d = .17$), but those who had not seen gesture showed significant forgetting across the delay ($t(13) = 5.11$, $p < .001$, $d = .88$). Children in the *gesture-instructed* condition showed significant forgetting, with their high level of recall reducing, both when they had observed gesture during learning ($t(13) = 3.55$, $p = .004$, $d = .45$) or had not observed gesture ($t(13) = 4.38$, $p = .001$, $d = .50$).

Together, these findings suggest that when children were restricted from gesturing, they maintained their low levels of fact recall, regardless of whether or not they saw gesture at encoding. When children were able to gesture (either instructed or allowed) however, their recall was initially higher at Time 1, but decreased at Time 2. The only

---

[22] See Appendix B, S3.6 for Kolmogorov-Smirnov tests of normality at the long delay (p. 186)

exception was that children, who were allowed to gesture and had seen gesture at learning, maintained a high level of fact recall at both time points.
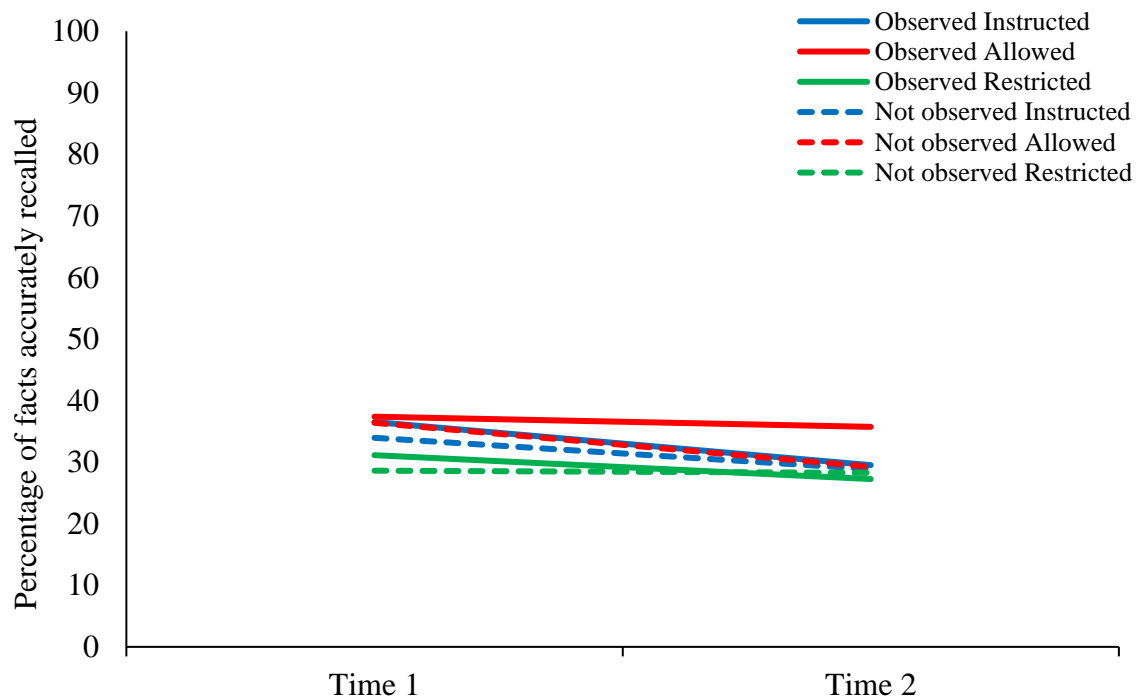
Figure 8.2. Children's recall of facts at Time 1 (1-day delay) and Time 2 (2-week delay) across the six experimental conditions.

Post-hoc analyses

Given that instructing children to gesture was not effective for enhancing children's gesture production, it was assessed whether children's gesture (regardless of interview condition) might mediate the relationship between observing gesture at learning and verbal recall during the interview. This was also done to examine the replicability of the mediational effects in Study 2 with a larger sample size.

Possible mediation at the short delay. For a mediation to occur, there first has to be a significant relationship between the predictor variable (learning condition) and the mediator (gesture production). Independent samples *t*-test indicated, however, that the learning condition had no significant effect on either children's representational gesture production rate ($t(54) = -0.91$, $p = .365$, $d = .24$), or event-relevant gesture production rate ($t(54) = -1.29$, $p = .203$, $d = .34$) during the interview next day. These non-significant effects ruled out the possibility that children's gesture production might mediate the relationship between observing gesture at learning and improved verbal recall, in the same

1-day delayed interview. These results did not replicate the findings of Study 2. The absent effect of observing gesture on children's gesture production also ruled out the possibility that Time 1 gesture production might mediate the relationship between observing gesture at learning and verbal recall two weeks later[23].

Possible mediation at the long delay. It was also examined whether children's gesture production during the Time 2 interview might mediate the relationship between observing gesture at learning and children's verbal recall during the same interview. Gesture rates (rather than raw counts) were again examined due to significant correlations between gesture counts and the number of words spoken by the children[24]. Gesture rate distributions were corrected to normal with log transformations.

First, the relationship between the predictor variable (learning condition) and the mediator (gesture production) was examined. Independent samples t-test revealed that children in the *gesture-observed* condition produced representational gestures ($t(82) = -2.05$, $p = .044$, $d = .45$) and event-relevant gestures ($t(82) = -2.27$, $p = .026$, $d = .50$) at a greater rate than children in the *gesture-not observed* condition, see Table 8.4.

It was then of interest to examine whether children's gesture rate during the 2-week delayed interview (representational and event-relevant) mediated the relationship between learning condition and children's verbal recall in the same interview. Two observed variable path models were created using Analysis of Moment Structures (AMOS Ver.22). The first model included learning condition as the predictor, children's representational gesture rate during the Time 2 (2-week delay) interview as the mediator, and children's Time 2 verbal recall (total, facts, concepts, and accuracy) as the outcome measure. The model was created with an embedded bootstrap for the mediational analysis. The size of the indirect effect was estimated with a 95% bias corrected confidence interval computed on the basis of 2,000 bootstrapped iterations with an estimand.

Children's representational gesture rate during the Time 2 interview significantly mediated the relationship between learning condition and Time 2 total recall ($a*b = 2.93$, $se = 2.22$, 95% CI [0.08, 9.48], $p = .042$), and learning condition and Time 2 recall of facts ($a*b = 1.21$, $se = 0.91$, 95% CI [0.03, 3.92], $p = .043$). No other mediational effects were significant ($p > .05$). When children produced representational gestures at a higher rate

---

[23] Correlations between children's Time 1 gesture production rates and verbal recall both at Time 1 and Time 2 were also examined. See Appendix B, Supplementary Analysis S3.7 for correlation table (pp. 187 − 188).
[24] Significant correlations were found between representational gesture count and words spoken ($r(82) = .74$, $p < .001$), and event-relevant gestures and word count ($r(82) = .68$, $p < .001$)

throughout the 2-week delayed interview, their total recall and recall of facts during that same interview was significantly improved (see Figure 8.3).

Figure 8.3. Model of representational gesture rate at Time 2 mediating the relationship between learning condition and verbal recall at Time 2. Standardised regression coefficients are reported.

$^*p < .05$

The same mediational model was run a second time, including children's event-relevant representational gesture rate during the Time 2 interview as the mediator, see Figure 8.4. Children's event-relevant gesture rate significantly mediated the relationship between learning condition and Time 2 total recall ($a*b = 3.59$, $se = 2.42$, 95% CI [0.33, 10.52], $p = .026$), learning condition and Time 2 recall of facts ($a*b = 1.46$, $se = 0.99$, 95% CI [0.14, 4.34], $p = .030$), and learning condition and Time 2 recall of concepts ($a*b = 0.13$, $se = 0.09$, 95% CI [0.001, 0.38], $p = .047$). Children's Time 2 representational gesture rate did not significantly mediate the relationship between learning condition and Time 2 recall accuracy ($p > .05$). When children observed gesture during learning, they produced event-relevant gestures at a greater rate two weeks later and in doing so improved their total recall, recall of facts, and recall of concepts, during the same interview.
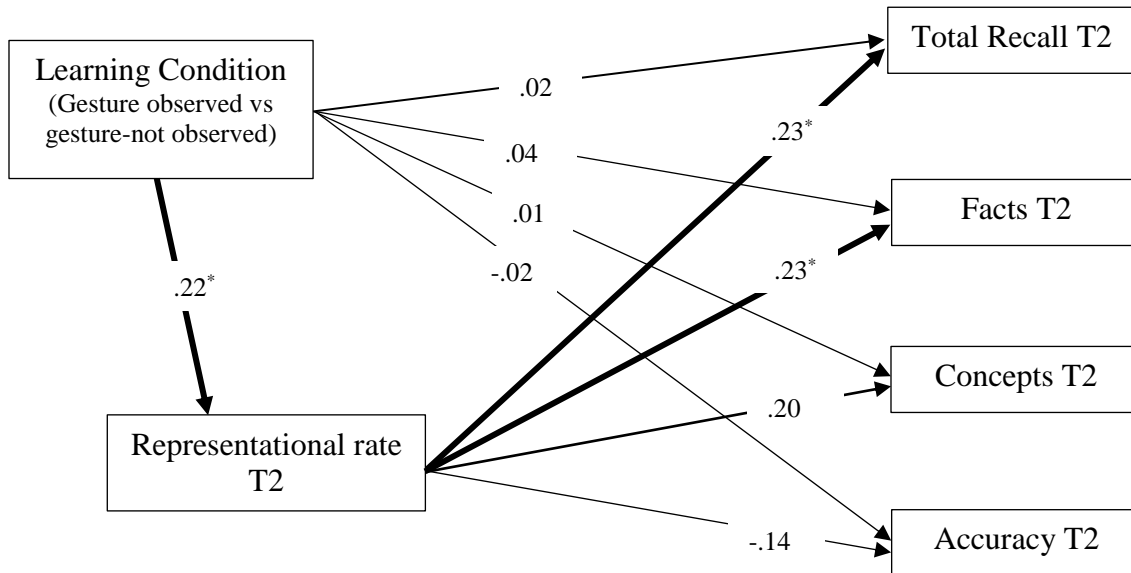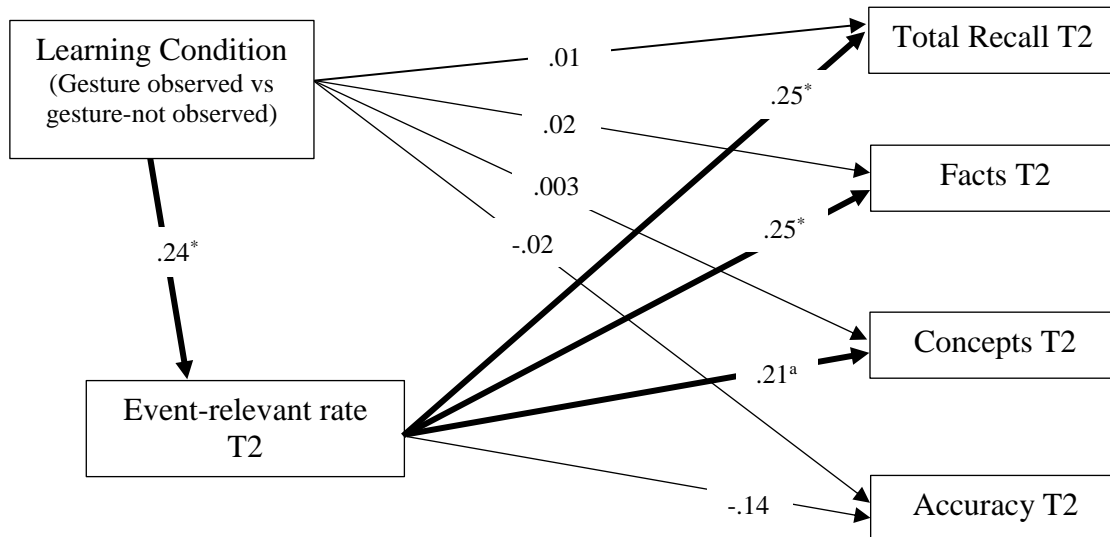
Figure 8.4. Model of event-relevant gesture rate at Time 2 mediating the relationship between learning condition and verbal recall at Time 2. Standardised regression coefficients are reported.

$^{*}p < .05$ $^{a}p < .06$

While the mediations were conducted with cross-sectional data, the absence of an association between learning condition and verbal recall ruled out the possibility that verbal recall may mediate the relationship between learning condition and children's gesture production rate. This is also supported by the findings that gesture typically slightly precedes speech (Chu & Hagoort, 2014; McNeill, 2005; Morrel-Samuels & Krauss, 1992) and plays an important role in influencing how speech is produced (Kita & Özyürek, 2003; Kita et al., 2007; Krauss et al., 1996; Krauss et al., 2000).

Summary. These results suggest that, like in Study 2, children's gesture production rates mediated the relationship between learning condition and verbal recall within the same interview. Interestingly, however, this effect only occurred two weeks after learning when children's gesture production was not manipulated.

Discussion

The aim of this study was to examine the separate and interactive effects of observing gesture at learning and producing gesture at recall, on children's verbal recall of a complex scientific lesson. Of primary interest, however, was whether children's verbal recall could be enhanced by instructing children to gesture or reduced by restricting children from gesturing, compared to when children were simply allowed to move their

hands. The separate and interactive effects of gesture were examined one day after learning, and again two weeks later.

The influence of observing or producing gesture on children's verbal recall

It was first examined whether children who observed gesture at learning would verbally recall more information than those who did not. As hypothesised, the learning conditions did not differ in the amount or accuracy of their verbal recall. When children's recall of facts and concepts were also examined, again, no differences between the conditions were found. These results replicate the findings from Study 2 and suggest that the fleeting nature of both the gestural and verbal information taught, combined with the developmental stage of the participating children (see Church et al., 2000) may have made integrating information across the modalities more challenging. As a result, children may have simply relied on recalling the verbal information they encoded.

The current study also manipulated whether children were instructed to gesture, allowed to gesture, or restricted from gesturing during the 1-day delay interview. It was hypothesised that instructing children to gesture would increase children's gesture rate and verbal recall compared to when children were simply allowed to gesture. Children restricted from gesturing were expected to exhibit the worst verbal recall of all the interview conditions. Contrary to hypotheses, however, the *gesture-instructed* and *gesture-allowed* interview conditions did not differ in either their gesture rate or verbal recall (amount, accuracy, facts, or concepts). Children restricted from gesturing also recalled similar amounts of information, with similar accuracy, compared to the other interview conditions. When children's recall of facts was examined, however, children in the *gesture-restricted* condition did report significantly fewer facts than the other gesture interview conditions.

The limited effect of instructing children to gesture is puzzling given that the instructions were almost identical to those given by Stevanoni and Salmon (2005). In their study, they found that children not only gestured more when they were asked to, but also recalled more information about a personal experience that had occurred two weeks beforehand. Other studies have also found that specifically asking children to gesture influenced children's gesture production, and improved their story recall (Cameron & Xu, 2011) or ability to solve mathematical equivalence problems (Broaders et al., 2007; Cook et al., 2008; Goldin-Meadow et al., 2009). Why then did the instructions in the current study show limited beneficial effects? The differences may be due to the variety of gestures observed at learning or the complexity of the verbal recall.

114

In many of the studies, which have manipulated children's gesture production, children have observed a select number of gestures. These gestures were associated with a particular problem solving method or limited verbal content (e.g. Broaders et al., 2007; Cook et al., 2008; Goldin-Meadow et al., 2012). In those cases, when children were then instructed to gesture it may have been easier for children to associate particular verbal content with the particular gestures they were expected to produce. Other studies, which find beneficial effects of instructing children to gesture, have also specifically told the children how to move their hands (Brooks & Goldin-Meadow, 2016; Goldin-Meadow et al., 2009). In the current study, children heard a large amount of verbal content accompanied by a variety of different gestures. When children were then asked to gesture during the interview they were given no guidance about the types of movements they should produce. As a result, it may not have been as apparent to the children how they should gesture.

Across extant research, the complexity of the information children were required to recall has also varied considerably. When children were asked to recall a staged event they participated in (Stevanoni & Salmon, 2005) or a story (Cameron & Xu, 2011) the information itself was unlikely to be conceptually difficult. Take, for example, the study where children were required to recall a spatial layout (Sauter et al., 2012). The children knew the location of the animals when assessed via a recognition test, but what they found difficult was reporting this information. In such studies, the base gesture rates were typically low. When describing a personally experienced staged event children were only gesturing between 2-4% of the time (Stevanoni & Salmon, 2005) and when recalling a story only 8 of the 15 children instructed to gesture actually did so (Cameron & Xu, 2011). This makes sense given that gesturing is more likely to occur when information is spatially complex or difficult to conceptualise (Melinger & Kita, 2007). When the base gesture rate is low, instruction to gesture may have a greater effect.

In the current study, however, the lesson information that children were required to recall was designed to be challenging. This likely increased the children's natural propensity to gesture to the point where instructing gesture may not have been as effective. Instead, the complexity of the to-be-recalled information may have meant that adding an additional request for the children to move their hands introduced additional cognitive demands (Sweller et al., 2011). In the phonological loop, children had to process the request to gesture as well as retrieve the solar system information. The act of producing gesture was still effective though, as children's gesture rates were associated with

improved verbal recall in the mediational analysis. This suggests that gesturing was beneficial but the instructions to gesture may have thrown a spanner in the works, confusing and taxing the participating children.

Restricting children's gesture did not affect children's total recall or recall accuracy, but it did impair children's recall of facts compared to those who were allowed to gesture. This partially supports the assertion that restricting children's gesture is likely harmful for recall. It is likely that the gestures the children were producing aided their recall of the important solar system information and that restricting such movements limited children's ability to access lexical items (Krauss et al., 1995) and externalise any implicit understandings (Broaders et al., 2007). While instructing the children to gesture had limited beneficial effect, restricting children's gesture did diminish recall, though in subtle ways.

Interactive effects of observing gesture and producing gesture on verbal recall

This study also investigated the interactive effects that observing gesture at learning and manipulating gesture at retrieval would have on children's verbal recall. It was expected that conditions during which children saw gesture and were able to perform gesture would exhibit superior recall compared to conditions during which children had not seen gesture and/or were restricted from gesturing. Contrary to expectations, however, there were no significant interactions between observing and performing gesture at the short delay, for either total recall, recall of facts, or recall of concepts. There was, however, a significant interaction for recall accuracy. Children in the *gesture-not observed* condition did not significantly differ across interview conditions in the accuracy of their verbal reports. Children in the *gesture-observed*, *gesture-instructed* condition, however, were significantly less accurate than children who had seen gesture and were allowed to gesture. As previously discussed, it was hypothesised that the instruction to gesture added cognitive load (Sweller et al., 2011). The reason why this only reduced accuracy when children had observed gesture is difficult to explain. From the current perspective, seeing gesture during learning should have reduced cognitive load during encoding and provided the children with accurate and informative ways to move their hands at retrieval (Singer & Goldin-Meadow, 2005).

Forgetting across the delay

When forgetting was examined, the majority of interactions were non-significant. There was, however, a significant interaction between the conditions for the recall of facts. Across the six experimental conditions, only three did not exhibit significant forgetting.

116

Children in the *gesture-observed, gesture-restricted* and *gesture-not observed, gesture-restricted* conditions maintained their low-level of responding across the delay, while children in the *gesture-observed, gesture-allowed* condition maintained their high-level of recall. In these three cases, children were able to boost their retrieval at the two-week delay.

The maintenance in children's high-level of recall in the *gesture-observed, gesture-allowed* conditions may be due to the mediational effect found at the long delay. Observing gesture at encoding was found to increase children's gesture production two weeks later, which in turn improved children's verbal recall. Children in the *gesture-observed, gesture-allowed* condition may have benefited from the mediational effect. Comparatively, children in the *gesture-observed, gesture-instructed* condition may not have gleaned the same benefits, possibly due to the increased cognitive load created by the gestural instructions in the first interview. As retrieval is a constructive and reconstructive process (Koriat, Goldsmith, & Pansky, 2000), increasing cognitive load during retrieval may have made it more difficult to subsequently retrieve the memory. When children were simply allowed to gesture, however, they were able to improve their recall without externally imposed demands.

The maintained low levels of recall in the two *gesture-restricted* conditions likely occurred due to children being allowed to gesture in the second interview. While their responding in the first interview was comparatively lower than the other conditions, children's ability to move their hands in the second interview may have improved lexical access (Rauscher et al., 1996) and reduced working memory load (Cook et al., 2012; Ping & Goldin-Meadow, 2010). As a result, children may have enhanced their recall of accurate information, and reduced their forgetting. These findings highlight the importance of gesture production in enhancing children's verbal recall.

Separate recall of facts and concepts

When the recall of facts and concepts were separately examined, it was hypothesised that children's recall of concepts would be particularly enhanced by the presence of gesture. Observing or performing gesture was expected to help children create a cohesive understanding of the information as a whole (Cutica & Bucciarelli, 2011, 2013). Instead, neither observing nor producing gesture appeared to benefit children's recall of concepts. There was an indication that children's production of event-relevant gestures mediated the relationship between observing gesture and the recall of concepts, but given how close the confidence interval was to overlapping with zero, this finding

must be interpreted cautiously. Alas, current findings suggest that gesture may be most beneficial for children's verbal recall of isolated facts rather than integrating and understanding the lesson as a whole.

Gesture production as a mediator

Post-hoc analyses were also conducted to examine whether the mediational effect found in Study 2 was replicated in the current study. Children who observed gesture at learning were expected to replicate those gestures during the 1-day delay interview and, in doing so, enhance the amount and accuracy of their recall. Instead, it was found that observing gesture during learning had no significant effect upon children's gesture production. This ruled out the possibility that children's gesture at the 1-day delay would mediate the relationship between observing gesture at learning and enhanced verbal recall in the same interview or across interviews. An indirect effect of observing gesture on recall was found at the long delay, however. Children who observed gesture during learning produced representational and event-relevant gestures at a greater rate two weeks later and recalled more information, and specifically more facts and concepts, during that same interview. Recall accuracy was not associated with gesture production.

The learning materials and interviews were virtually identical across Studies 2 and 3. This raises the question as to why the mediational effect was present at the short delay in Study 2, but not the current study. The influence of observing gesture on children's gesture production appears to be subtle and, as a result, sample differences may have had an effect. Specifically, participant numbers varied across the two studies. In Study 2, there were only two conditions, each with 21 participants, while in Study 3, there were six conditions, each with 14 participants. Study 3 had a larger sample size for the gesture-observed and gesture-not observed conditions (42 participants each), but there was also more variation within the conditions due to the gesture production manipulation. Small samples can leave open the possibility of individual differences having a greater bearing on results. For example, variations in children's everyday exposure to gesture can influence children's gesture production and verbal language (Rowe & Goldin-Meadow, 2009) and children's developmental stage can affect their ability to integrate gestural and verbal information (Church et al., 2000; Kelly & Church, 1998). While this study controlled for differences in age and knowledge across the conditions, other individual difference factors, combined with small samples, may have shaped the differential effects across the two studies.

Another slight difference across the two studies was the gestural coding method used. In Study 2, gestures were coded with audio information present and hand movements were only included as event-relevant if they both matched the scripted gesture and accompanied relevant verbal content (i.e. gesture-speech matches). In the current study, a more conservative approach was used and the gestures were coded without sound. This meant that all gestures, which were visually similar to the scripted gestures, were coded as event-relevant, and could include both gesture-speech matches and mismatches. For example, a child may produce an orbit gesture while saying, "The sun goes around the planets." In Study 2, this would have been coded as a novel gesture due to the incorrect verbal information. In the current study, however, the orbit gesture would have been coded as event-relevant. It could be theorised that observing gesture at learning not only affects the hand movements children produce but also the verbal information these movements accompany. It is often found that gesture and speech are semantically linked (De Ruiter et al., 2012; Kita & Özyürek, 2003; So et al., 2009) unless children are in a state of transitional knowledge (Broaders et al., 2007; Garber & Goldin-Meadow, 2002). By including gesture-speech mismatches within the event-relevant category, associations between observing and producing gesture may have been muddied. Further research should examine how different gestural coding methods may affect associations between gesture production and verbal recall.

At the long delay, the indirect effect of observing gesture on children's recall was found when children's gesture production was included as a mediator. It is not an uncommon finding that the effects of gesture become greater with time (Cook et al., 2008). Once forgetting occurred in this study, it seems likely that children may have relied more heavily on the visual imagery presented to them in gesture during encoding, as images typically create a stronger memory trace than verbal information (Church et al., 2007; Paivio, 1971). Due to the visuospatial nature of the pictorial information, however, it may have been easier for the children to retrieve this information through their own gesture production which is also visuospatial (see Church et al., 2000). They could then use the proprioceptive feedback, and visual information depicted in front of themselves in their gesture, in order to chunk the information into portions that were easy to verbalise (Alibali et al., 2000; Cooperrider et al., 2015; Kita, 2000; Wagner et al., 2004).

The different style of gestural coding may also explain why the mediational effect found in Study 2 was only replicated at the long delay. One could speculate that, during the 1-day delay, children were producing event-relevant gestures but with mismatching

verbal content. For example, a child may produce gesture where they pull their hands in towards their body while saying, "Gravity pushes things down." It is often found that speech-gesture mismatches are typically a precursor to more advanced understanding (Alibali & Goldin-Meadow, 1993; Roth, 2003). It may have taken time for children to develop their understanding to the point where their gestures aided the production of accurate and relevant speech (e.g. verbally conveying gesture as an inward pull). At the two-week delay, children's earlier gesture may have enhanced the verbal recall.

Summarising the mediational effects it appears that at the short delay producing gesture did not positively influence children's verbal recall. Two week later, however, what is most important is not whether or not the child was instructed to gesture but whether or not they initially observed gesture. This is similar to what Cook and Goldin-Meadow (2006) found when children were taught how to solve mathematical equivalence problems. When children observed gesture, they mimicked the gesture they saw, regardless of whether or not they were instructed to. Within middle childhood, observing gesture at encoding appears to play a greater role in enhancing children's gesture production than the instruction for the children to move their hands as they talk.

Conclusions

Both observing and performing gesture can enhance children's verbal recall. When children perform gesture, regardless of whether or not they are instructed to, they appear to recall more information and in particular more facts. In order to enhance children's propensity to gesture, it appears that adults should not specifically ask children to gesture, as this might increase cognitive load. Equally, restricting children from gesturing is also likely to be detrimental for verbal recall, though if the child is later able to gesture they can boost their recall to reduce forgetting. Instead, adults should model richly informative gestures at encoding and then leave children to gesture as they naturally would while they recall their learning.

Given the negative effects which occurred when children were instructed to gesture, further research is clearly needed to examine when gestural instructions are and are not effective. In the current study, it was proposed that the gestural instructions were unhelpful because the children were confused about how they should move their hands. Future research should conduct a similarly complex and verbally rich lesson, but reduce the number of gestures to a select few, which are repeated and are consistently tied to the same words. If children are finding it difficult to identify how to move their hands when they are instructed to gesture, exposing children to a set of repeated gestures may make it

easier for children to recognise when and how they can appropriately move their hands. From there, research could further examine whether interactive effects between observing gesture at learning and producing gesture at retrieval influence children's verbal recall.

The results of the current study highlight the importance of examining gesture in naturalistic learning settings, especially those that teach a variety of facts and concepts accompanied by a range of gestures. Unlike the results found by previous studies, the current study found that specifically asking the children to gesture was not helpful. Instead, modelling gesture and allowing the child to gesture as they recall their learning appears to be the most appropriate means of enhancing recall, though the effects are likely to be subtle.

Chapter 9

Study 4: Gesture and children's reporting of spatial terms

Throughout this thesis, two major limitations of the extant research were raised in terms of the learning experiences children were engaged in, and the length of the delay between encoding and recall. This thesis aimed to address these limitations by having children participate in a verbally rich and complex scientific learning experience. Moreover, children's recall of the science lesson was assessed with comprehensive interviews, at both short (1-day) and long (2-week and 7-month) delays. While Studies 1, 2, and 3 all examined the quantitative aspects of children's verbal recall (i.e., amount, accuracy, and separate recall of facts and concepts), there was also a unique opportunity to examine the qualitative elements of children's verbal reports. Using the data from Studies 1 (Chapter 6), 2 (Chapter 7), and 3 (Chapter 8), this study examined how gesture, either observed or performed, may have influenced children's use of spatial language.

As described throughout this thesis, gesture plays an important role in communicating information (McNeill, 1992, 2005). Compared to speech, the visuospatial nature of gesture also makes it particularly adept at conveying spatial content (Alibali, 2005; Beattie & Shovelton, 1999). Take, for example, describing the layout of a house. To verbalise this content, each piece of spatial information has to be spoken sequentially (e.g. "The lounge is the largest room and it is in the centre, to the right is the bathroom…"). Such verbal descriptions are taxing for both the speaker to produce and for the listener to comprehend. The speaker has to hold all the visuospatial information in mind and convert it into speech, while also keeping track of what information they have reported and what they are yet to report. Equally, the listener is also required to keep each piece of the visuospatial information in mind in order to create a mental representation of the content as a whole (Rapp & Kurby, 2008; So, Shum, & Wong, 2015). Information pertaining to size, shape, and location may be more easily presented with gestural images. It is therefore unsurprising that representational gestures more often accompany spatial information, compared to non-spatial content (Alibali et al., 2001).

Observing gesture and spatial content

The gestures that accompany verbal spatial content are typically meant to be communicative. Studies have found that when individuals can see their conversational partner, they use more representational gestures than when they cannot see the listener

(Alibali et al., 2001; Emmorey & Casey, 2002). When individuals do observe gesture during interactions, their comprehension and recall of spatial information appears to improve. For example, when adult listeners heard spatial terms described with accompanying gesture, they better identified target words than when gesture was not present (Driskell & Radtke, 2003). Similarly, children who observed gesture while learning a directional route were also found to more accurately recreate the route, compared to children who did not observe gesture (Wermeskerken et al., 2016).

While observing gesture may improve individuals' non-verbal spatial recall, or the recall of isolated terms, studies rarely examine the influence of observing gesture on children's ability to describe narrative spatial content. One study did find, however, that children (3 – 4 years old) who observed gesture at encoding described more spatial information as they recalled a narrative story, compared to children who did not observe gesture (Austin & Sweller, 2014). Building upon this research, the current study examined whether observing gesture might also improve the amount of spatial content reported by older (7- to 9-year-old) children as they recalled a complex science lesson. Throughout the current thesis, observing gesture had little influence on children's verbal recall of the *lesson* content compared to when gesture was not observed. This study examined whether the influence of observing gesture may have more specifically improved children's reporting of *spatial* information, at both short (1-day) and long (2-week, 7-month) delays.

Performing gesture and spatial content

Research has also examined the influence that *producing* gesture may exert on children's, and adults', verbal recall of spatial information. Studies with adults have found that participants who gestured more while describing their work activities also reported more spatial information (Trafton et al., 2006). Adult participants who were able to gesture while explaining spatial puzzles also described more orientation information compared to those who could not gesture (Emmorey & Casey, 2002). The more adults produce gesture, the more spatial information they appear to report.

Improved spatial descriptions have also been found when children perform gesture. For example, the more 3- to 4-year-old children gestured during recall, the more spatial information they verbally reported about a character's activities (Austin & Sweller, 2014). Older children (8 – 10 years old), who were required to verbally recall a spatial layout, have also been found to report more spatial information the more they gestured (Sauter et al., 2012). In fact, the study by Sauter et al. (2012) found that spatial information was rarely described unless it was first conveyed in gesture. The current study investigated

124

whether children's gesturing also improved the reporting of spatial information present within a broad and complex lesson. The influence of performing gesture was again examined at both short and long delays after learning.

While individuals generally appear to report more spatial information the more they gesture, restricting gesture has had variable effects on the verbalisation of spatial content. Some studies have found that restricting gesture increases adults' use of spatial terms, compared to when gesture is permitted (Emmorey & Casey, 2002). In contrast, other studies have found that gesture restriction has no effect on adults' spatial descriptions (Hoetjes, Krahmer, & Swerts, 2014; Hostetter, Alibali, & Kita, 2007a). During communication, it has been proposed that the message as a whole is presented across the verbal and gestural modalities, with gesture representing the visuospatial information (McNeill, 1992). From this perspective, restricting gesture would mean that the spatial content, typically conveyed in gesture, would have to be converted into speech. When this idea has been tested, however, studies typically find that the speech and gestural modalities do not compensate for each other (De Ruiter et al., 2012; So et al., 2009). Given the conflicting findings, further research is needed to examine the effect restricting gesture may have on the verbalisation of spatial content. The current study also examined the effect that inhibiting children's gesture production would have on the amount of spatial information subsequently reported.

How might gesture improve spatial recall?

When gesture is observed or performed, spatial descriptions are generally found to improve. This raises the question, why? First, when individuals perform gesture, lexical access may improve as individuals have both a verbal and a spatial-motoric method of accessing spatial words from the lexicon (Krauss et al., 1996; Krauss et al., 2000; Krauss et al., 1995). Secondly, by gesturing, individuals may focus more on the spatial information they could report. For example, within a conversation, adult participants have been found to verbalise more spatial information when gesture was permitted compared to when it was restricted (Rimé, Schiaratura, Hupet, & Ghysselinckx, 1984). Finally, by gesturing, individuals may be able to produce a holistic, imagistic representation of the information, which could then help organise and chunk the spatial content in preparation for verbalisation (Alibali et al., 2000; Kita, 2000; Wagner et al., 2004). When organising the spatial content, gesturing may also enable the rehearsal and maintenance of this imagistic representation in the visuospatial sketchpad (Wesp, Hesse, Keutmann, & Wheaton, 2001).

Similar mechanisms may also be at work when individuals observe gesture. First, gesture may highlight the spatial information represented both gesturally and in speech. Second, the additional gestural representation may disambiguate any confusing verbal spatial content (Kelly, 2001; Thompson & Massaro, 1986) and possibly enhance the depth of encoding and ease of retrieval (Paivio, 1971, 1991). Cognitive load is also likely to be reduced when gesture is present, as the visuospatial and the verbal information can be processed in different working memory subsystems (Sweller et al., 2011). Finally, during retrieval, the visuospatial memory trace might help frame the spatial information in a way that makes it easier to verbally report the spatial content (Alibali et al., 2000; Kita, 2000).

The current study

In the three studies previously conducted throughout this thesis, children were taught about the solar system. Children's recall of the solar system information was assessed with comprehensive interviews at both short (1-day) and long (2-week to 7-month) delays. The current study examined the data from Studies 1, 2, and 3. The aim of this study was to investigate whether the presence of gestures, both those observed at learning and produced by the children at recall, would specifically improve children's reporting of spatial content when recalling the solar system lesson.

At the short (1-day) delay, it was hypothesised that when children observed gesture during learning they would recall more spatial terms compared to children who did not observe gesture. It was also hypothesised, that children who produced more gesture at recall would also report more verbal spatial content. The effect of restricting gesture was uncertain. Either children would report more spatial content in order to compensate for the absence of gestural information, or the gestural restriction would have no effect on the reporting of spatial terms.

While studies have not specifically examined children's recall of spatial information at longer delays, hypotheses were made based on the verbal recall results found previously in this thesis. First, children in the *gesture-observed* condition were expected to report more spatial information than the *gesture-not observed* condition two weeks after learning, but not seven months later. Second, at both the 2-week and 7-month delays, children who produced more gestures were expected to recall more spatial information within the same interview. Whether producing gesture during the Time 1 (1-day delay) interview would increase children's reporting of spatial content at Time 2 (2-week and 7-month delay) was also explored, but no specific hypotheses were made.

Method

Across all three studies, the interview transcripts were analysed for the types of words children used when describing their learning. Five points of data collection were analysed across Study 1 (1-day delay), Study 2 (1-day and 7-month delay), and Study 3 (1-day and 2-week delay). All content spoken by the interviewer was removed from the interview transcripts, as well as non-fluencies (e.g. umm, ahh), and incomplete words (e.g. "we w-, we went") produced by the child. Data were entered into the Linguistic Inquiry and Word Count (LIWC) software. LIWC calculates the percentage of words within each transcript which fit in each of the designated categories (Tausczik & Pennebaker, 2010). A single word can fall within more than one category.

In terms of children's use of spatial language, the percentage of prepositions and quantifiers, as well as motion, space, and time words were all combined into the category of "spatial terms."[25] While time is an abstract concept, it is typically conceptualised spatially in a mental time line (Lakoff & Johnson, 1999). As such, time was included in the spatial definition. See Table 9.1 for examples of the types of words included within each category (examples taken from the LIWC Manual, see Pennebaker, Chung, Ireland, Gonzales, & Booth, 2014).

Table 9.1

Example of Terms Within Each Word Category

Category	Example
Spatial terms	
Prepositions	To, with, above
Quantifiers	Few, many, much
Motion	Arrive, car, go
Space	Down, in, thin
Time	End, until, season

Children's representational (Rep) and event-relevant (ER) gesture rates were again examined. The representational gesture category assessed children's total production of informative gesture, while the event-relevant gesture category captured children's

---

[25] See Appendix B, S4.1 for additional analyses when the spatial information was divided into subcategories (pp. 188 – 192).

production of representational gestures that mimicked those used throughout the lesson. Children's gesture rates per 100 words were assessed in order to control for children's verbosity and therefore their opportunities to gesture. The coding of children's gesture rates was carried out in Study 2 and Study 3, but was not of primary interest in Study 1. To examine children's gesture production in Study 1, additional gesture coding was conducted on the Study 1 interview data following the same procedure as was used in Study 2 (see Chapter 7). A research assistant, unaware of the children's condition, coded 20% of the interviews to establish reliability. Coder agreement was calculated using both percentage agreement and Cohen's Kappa. Agreement as to whether a representational gesture was present or absent was 93% ($k = .75$). When coders agreed that the gesture was representational, percentage agreement as to whether the gesture was event-relevant or novel was 82% ($k = .62$).

Results

Across the studies, approximately 30% of the verbal content children reported conveyed spatial information. This study examined: 1) whether children's recounting of spatial terms differed across the *gesture-observed* and *gesture-not observed* conditions, and 2) whether there was an association between children's gesture production rates and the percentage of spatial information children verbally reported. Children who answered *wh*-questions in Study 1 were excluded from all analyses in order to reduce variance unexplained by the gestural manipulations.

Observed gesture and spatial content

To investigate the hypothesis that children who observed gesture would report more spatial information than those who did not, independent samples *t*-tests were conducted. At the short (1-day) delay, the findings were mixed. In Study 1, the *gesture-observed* and *gesture-not observed* conditions did not significantly differ in the percentage of spatial terms reported, ($t(37) = 1.77$, $p = .084$, $d = .57$). In Study 2, a marginally significant effect of learning condition was found, with children who observed gesture reporting a greater percentage of spatial information than children who did not, ($t(40) = -1.97$, $p = .056$, $d = .61$). In Study 3, children who observed gesture reported significantly more spatial terms than children who did not observe gesture ($t(82) = -2.65$, $p = .010$, $d = .58$). Results indicate that children who observed gesture were generally reporting a greater percentage of spatial content the next day compared to those who did not observe gesture, but the mean differences were typically only small (i.e., a 2-3% difference; see Table 9.2).

Table 9.2

Means (Standard Deviations) for the Percentage of Spatial Information Reported Across the Learning Conditions

	Study 1 (1 day)		Study 2 (1 day)		Study 2 (7 months)		Study 3 (1 day)		Study 3 (2 weeks)	
	Gesture	No Gesture	Gesture	No Gesture	Gesture	No Gesture	Gesture	No Gesture	Gesture	No Gesture
Spatial	29.43	31.52	32.09	29.39	23.28	22.60	27.13	24.66	33.53	30.14
	(3.92)	(3.40)	(4.58)	(4.31)	(5.61)	(5.14)	$(4.15)_a$	$(4.37)_b$	$(6.05)_a$	$(5.92)_b$

Note. Different subscripts represent significant differences across the groups

Table 9.3

Correlations Between Children's Reporting of Spatial Information and Gesture Production Rates During the Same Interview

	Study 1 (1 day)		Study 2 (1 day)		Study 2 (7 months)		Study 3 (1 day)		Study 3 (2 weeks)	
	Rep rate	ER rate	Rep rate	ER rate	Rep rate	ER rate	Rep rate	ER rate	Rep rate	ER rate
Spatial	.25	.16	.08	.03	.53[**]	.44[*]	.36[**]	.35[**]	.13	.16

[*] $p < .05$ [**] $p < .01$

At the longer delays, it was expected that observing gesture would increase the percentage of spatial terms reported two weeks after learning, but not seven months later. As hypothesised an independent samples *t*-tests revealed that children who observed gesture during learning reported a larger percentage of spatial terms two weeks later (Study 3), compared to children who had not seen gesture ($t(82) = -2.59$, $p = .011$, $d = .57$). Seven months after learning (Study 2), the *gesture-observed* and *gesture-not observed* conditions did not significantly differ in the percentage of spatial terms reported ($t(28) = -0.35$, $p = .733$, $d = .13$). Observing gesture may result in small increases in the reporting of spatial information after a delay of a couple of weeks, but not several months (see Table 9.2).

Performed gesture and spatial content

It was also hypothesised that the more children gestured, the more spatial information they would verbally report. To test this hypothesis, bivariate Pearson's correlations were conducted examining associations between children's gesture production rates and their percentage recall of spatial terms, both within and across interviews. Children's representational (Rep) gesture rates and event-relevant (ER) representational gesture rates were both examined. All gesture production rates were normally distributed or were corrected to be so using square root or log transformation[26]. Only data from the long delay in Study 2 could not be corrected.

Within the 1-day delay interview, no significant correlations were found between children's gesture rates and the percentage of spatial terms reported in either Study 1 or Study 2 (see Table 9.3). In Study 3, however, those who gestured at a greater rate, either with representational or event-relevant gestures, verbally reported a significantly greater percentage of spatial content in the same interview. The only difference between Study 3 and the other two studies was the instruction to gesture. Post-hoc analyses were conducted to examine whether the gestural manipulation in Study 3 may explain the different results across the studies.

Correlations revealed that, when children were *instructed* to gesture, there was a significant positive relationship between the percentage recall of spatial terms and both children's representational gesture rate ($r(26) = .51$, $p = .005$) and event-relevant gesture rate ($r(26) = .50$, $p = .007$). When children were *allowed* to gesture, however, correlations between gesture rates and spatial language were non-significant (Representational gesture:

---

[26] See the respective studies for information on the transforms used.

$r(26) = .24, p = .223$; Event-relevant gesture: $r(26) = .25, p = .200$). The instruction to gesture seemed to be driving the significance of the correlation between children's Time 1 gesture production and Time 1 spatial recall in Study 3.

Within the long delay interviews, correlations between children's gesture production rates and the percentage of spatial terms reported were also examined. Two weeks after learning (Study 3), children's gesture production rates were not significantly associated with the percentage recall of spatial information in the same interview. Seven months later (Study 2), however, children who produced representational and event-relevant representational gestures at a greater rate also verbally reported significantly more spatial content in the same interview (see Table 9.3).

Finally, it was examined whether children's gesture production rates during the 1-day delay interview were associated with children's reporting of spatial terms two weeks (Study 3) and seven months (Study 2) later. There were no significant correlations between early gesture rates and the percentage of spatial terms later recalled (see Table 9.4).

Table 9.4

Correlations Between Children's Gesture Production Rates during the Time 1 (1-day delay) Interview and the Percentage of Spatial Information Reported in the Time 2 Interview

	Study 2		Study 3	
	T1 Rep rate	T1 ER rate	T1 Rep rate	T1 ER rate
T2 Spatial information	.15	.02	.11	.10

Note. Spatial information was recalled seven months later in Study 2, and two weeks later in Study 3
$^*p < .05$ $^{**}p < .01$

Manipulating children's gesture production

The influence of restricting children's gesture production on the percentage of spatial information reported was also assessed. In Study 3, children were either instructed to gesture, allowed to gesture, or restricted from gesturing during the 1-day delay interview. A one-way ANOVA was conducted examining children's reporting of spatial terms across the three interview conditions at the 1-day delay. There was no significant effect of interview condition on the amount of spatial language recalled at the 1-day delay

$(F(2, 81) = 0.11$, $p = .898$, $\eta_p^2 = .003)^{27}$. Children who were either instructed to gesture ($M = 25.71$, $SD = 4.96$), allowed to gesture ($M = 25.76$, $SD = 4.28$), or restricted from gesturing ($M = 26.21$, $SD = 4.10$), did not appear to significantly differ in the percentage of spatial terms reported. Children do not seem to verbally compensate for the inability to gesture.

Discussion

Collating the results, the findings were subtle and mixed. Children, who observed gesture during learning, typically reported more spatial information than those who did not. Whether this difference was significant or not varied across the studies and across the delay periods. When children's gesture production and use of spatial language were examined, mostly non-significant correlations were found. Significant correlations between children's gesture production rates and the percentage of spatial terms reported were only found when: 1) children were instructed to gesture or, 2) children's memory of the original lesson had substantially decayed.

Observed gesture.

When recall was assessed one day after the learning experience, children were hypothesised to report more spatial information if they observed gesture during learning compared to when they did not. This hypothesis was generally supported, but findings differed slightly across the studies. In Study 2 and Study 3, children who observed gesture reported more spatial content than those who did not. The difference was significant in Study 3, and marginally significant in Study 2. Comparatively, at the short delay in Study 1, no significant difference in the use of spatial language was found across the learning conditions. Additionally, as hypothesised, the effect of learning condition on the use of spatial language also differed across the delays. Two weeks after learning, children who observed gesture were still recalling more spatial information than those who did not, but seven months later there was no significant difference between the learning conditions.

Generally, observing gesture increased children's use of spatial language. A number of factors may explain why this effect occurred. First, observing gesture may have directed children to attend to and encode the spatial information the gestures conveyed. If the spatial information was encoded both verbally and visually, the depth of encoding and ease of retrieving the spatial content may also have been enhanced (Paivio, 1971, 1991).

---

[27] There was also no effect of interview condition on children's reporting of spatial content at the long delay. See Appendix B, Supplementary analysis S4.2 (p. 192).

Second, the analogue nature of gesture also makes it ideal for conveying spatial information (Alibali, 2005; Goldin-Meadow, 2015; So et al., 2015). As a result, the presence of observed gesture may have cleared up any ambiguities in the verbal information (Kelly, 2001; Thompson & Massaro, 1986). Additionally, the observed gesture may have provided the children with a visuospatial representation which they could draw upon in order to help organise and chunk the spatial content before verbalising it (Kita, 2000). These factors may explain why observing gesture improved the amount of spatial information reported, both one day and two weeks after learning. Furthermore, the maintained benefit two weeks later may also be explained by the finding that visual images decay more slowly in memory (Church et al., 2007; Paivio, 1971). If children recalled the more robustly stored gestural images after the delay, the visuospatial information they contained may have been more easily converted into speech.

At both the 1-day delay in Study 1 and the 7-month delay in Study 2, the *gesture-observed* and *gesture-not observed* learning conditions did not significantly differ in the percentage spatial terms reported. The null effect found in Study 1 was unexpected, but might be explained by the inclusion of visual aids within the learning experience. As previously mentioned, Study 1 included diagrams and interactive elements like swinging a ball around. When visual aids and gesture are both present, children likely have to integrate all the visual information together, which may increase cognitive load (Sweller et al., 2011). As a result, the children in Study 1 may have found it difficult to create a cohesive visuospatial representation of all the information they saw. Having a piecemeal visuospatial representation of the lesson content was potentially inefficient for scaffolding children's conversion of the spatial content into speech (see Kita, 2000). At the long delay, seven months after learning, children who did and did not observe gesture also did not differ in their reporting of spatial information. This effect supported the hypothesis. It is likely that, after seven months, both the visuospatial and verbal representations stored in memory had decayed to the point where limited solar system content was accessible for retrieval (Jones & Pipe, 2002).

Together, these findings suggest that observing gesture may support children in reporting more spatial information, both the next day and two weeks after learning. When multiple visuospatial representations are present or the original memory trace has substantially decayed, however, the presence or absence of observed gesture seems unlikely to affect the amount of spatial information children report.

Performed gesture.

As spatial information can be difficult to verbally describe, it was expected that children's gesture production would support the organisation of spatial content in preparation for verbalisation (Kita, 2000). Contrary to the hypotheses, however, children's gesture production rates were not significantly associated with the amount of spatial information children reported within the same interview. Given the non-significant correlations found *within* the interviews, it makes sense that performing gesture during an earlier interview also did not improve children's reporting of spatial information in a subsequent interview. Children's gesture production rates were only positively associated with the use of spatial language either when children were specifically instructed to gesture, or when seven months had passed since the original lesson.

While significant correlations were found between children's gesture production rates and their use of spatial language in the *gesture-instructed* condition, no significant correlations were found in the *gesture-allowed* condition. This is interesting given that spatial reporting as a whole did not differ across the two interview conditions, and neither did children's gesture production rates (see Study 3, Chapter 8). Within each interview condition, there are likely to be individual differences in terms of children's propensity to gesture (Chu et al., 2014). For example, if a child typically does not gesture as she speaks, instructing her to move her hands might have been confusing. While some children may benefit from the instruction to gesture, others may not. The correlation in the *gesture-instructed* condition indicates that instructing children to gesture may be effective for enhancing spatial language, but only if children understand what is required of them and follow through with the request to gesture. Additionally, given the non-significant correlation in the *gesture-allowed* condition, findings also suggest that making children explicitly aware of their gesture might be important. When children overtly notice the visuospatial information in their hands, they may be more likely to convert the spatial content into speech.

At the long, 7-month delay, gesture production was also correlated with children's reporting of spatial content. At the shorter, 1-day and 2-week delays, however, no significant correlations were found within the interviews. The main difference between the long delay and the other time points was the extent to which the memory had decayed. At the 1-day and 2-week delays, children were reporting around one-third of the solar system

content[28]. At the long delay in Study 2, however, children were reporting only about one-seventh of the information (see Chapter 7). When children were able to access the original memory trace, as they were at the shorter delays, the richness of the original memory may be most important for retrieval. This could explain why children recalled more spatial information when they observed gesture at learning, as they had a more comprehensive visuospatial memory trace to draw from. In comparison, when the original memory trace was sparse, as it was 7-months later, the environmental cues may be more important. When children produce gesture they might cue themselves in to additional information they could report (Frick-Horbury, 2002), and because gestures convey visuospatial content this information may be more likely to be described.

Restricting children's gesture production

It was also of interest to examine the effect of restricting gesture on children's reporting of spatial information. While some literature suggests that people compensate for not being able to gesture by verbally reporting more of the spatial content (Emmorey & Casey, 2002; McNeill, 1992), other studies suggest this is not the case (De Ruiter et al., 2012; So et al., 2009). In the current study, whether or not children were able to gesture had no significant effect on children's reporting of spatial information, suggesting children do not use their speech to compensate for not being able to gesture.

Conclusions

The influence of gesture on children's recall of spatial content was unclear. While the majority of data collection points showed more spatial language being used by children when gesture was observed, the effects were typically small. Additionally, positive associations between performing gesture and reporting spatial language only occurred under specific circumstances. Gesture may aid the verbalisation of spatial content, but whether the benefits are meaningful has yet to be conclusively established. Implications, applications, and future research are discussed further in the general discussion (Chapter 11).

---

[28] See descriptive statistics in Study 1 (Chapter 6), Study 2 (Chapter 7), Study 3 (Chapter 8)

Chapter 10

Meta-Analysis

Across the three studies carried out in this thesis, the effects of observing and producing gesture were varied. While observing gesture typically did not influence children's verbal recall, there were indications it did influence children's gesture production. The effects of children's gesture production on verbal recall were also inconsistent, however, which was likely due to the large variation in children's gesture rates within and across the samples. In order to examine overall effects, across the three studies, meta-analyses were conducted across the five points of data collection (i.e. Study 1, Study 2-Time 1, Study 2-Time 2, Study 3-Time 1, Study 3-Time 2).

In general, meta-analyses examine the 95% confidence intervals across studies and average the point estimates (i.e. the value which is the best guess of the true population value and is represented at the centre of the confidence interval; Cumming & Calin-Jageman, 2016). Within the meta-analyses, the average point estimates are weighted. Studies with shorter confidence intervals and larger sample sizes have more weight than studies with longer confidence intervals or smaller sample sizes. If the results are fairly consistent across the studies, then the confidence interval of the meta-analysis is typically shorter and provides a more accurate measure of the effect than any one study alone (Cumming & Calin-Jageman, 2016).

In this chapter, the data from the three studies were analysed using the Exploratory Software for Confidence Intervals (ESCI). Output from the analyses was represented in forest plots, with the green boxes representing the effect size of each individual study (see figures below). The bars attached to each box represent the 95% confidence interval for the difference between the means or the association between the two variables. The red diamond represents the result of the meta-analysis and the zero-line represents the null hypothesis. If the confidence interval of the red diamond does not cross the zero-line then there is evidence to reject the null hypothesis (i.e. $p < .05$). The more the confidence interval (i.e. red diamond) overlaps with the zero line, the more confident one can be in accepting the null hypothesis (Cumming & Calin-Jageman, 2016). The current meta-analyses examined both the effects of observing gesture and producing gesture on children's verbal recall, as well as the influence of observing gesture on children's own gesture production.

Observed gesture and verbal recall.

In this thesis, it was hypothesised that children in the *gesture-observed* condition would recall more information, more accurately, than children in the *gesture-not observed* learning condition. While none of the individual studies found a significant effect of learning condition on verbal recall, meta-analyses were conducted to examine the combined effect across all the studies (i.e. Study 1, Study 2-Time 1, Study 2-Time 2, Study 3-Time 1, and Study 3-Time 2). Due to the statistical model assuming data are normally distributed (Cummin & Calin-Jageman, 2017); any variables which were not normally distributed were corrected using square root transforms or, if they could not be corrected, were excluded from analyses. To deal with the fact that some outcome measures were raw scores while others were transformed, the meta-analyses were carried out using the *standardised mean difference* or Cohen's *d*, and random effects were implemented (Cumming & Calin-Jageman, 2017). Children from Study 1 were excluded from analyses if they were asked *wh*-questions throughout the lesson. This was done to avoid explaining sample variation unrelated to the gestural effects[29].

The meta-analyses first examined whether the *gesture-observed* and *gesture-not observed* conditions differed in their total recall and recall accuracy. The meta-analysis mean for total recall was $d_{\text{unbiased}} = 0.06$ [-0.17, 0.29] and, given that the confidence interval of the meta-analysis overlapped with zero (see the red diamond in Figure 10.1), observing gesture did not appear to significantly affect the amount children recalled. Similarly, whether or not children observed gesture during learning also had no significant effect on the accuracy of children's recall, $d_{\text{unbiased}} = 0.04$ [-0.24, 0.32], see Figure 10.2.

Meta-analyses also examined children's percentage recall of the solar system facts and concepts, taught across the two learning conditions, again using *Cohen's d*. The recall of facts and concepts were separately examined in order to assess whether observing gesture differentially aided the recall of isolated facts versus the recall of explanations and connections between the facts (i.e. concepts). The recall of concepts in Study 2-Time 2 was excluded from analyses as the distribution significantly differed from normal, and could not be corrected. Again, whether or not children observed gesture during learning

---

[29] When the children who answered *wh*-questions were included in the analyses, results did not differ. See Appendix B, Supplementary Analysis S5.1 for analyses with the *wh*-question conditions included (pp. 192 – 193).

had no significant effect on either the recall of facts ($d_{unbiased} = 0.14$ [-0.09, 0.37]), or concepts ($d_{unbiased} = 0.14$ [-0.10, 0.37])[30].
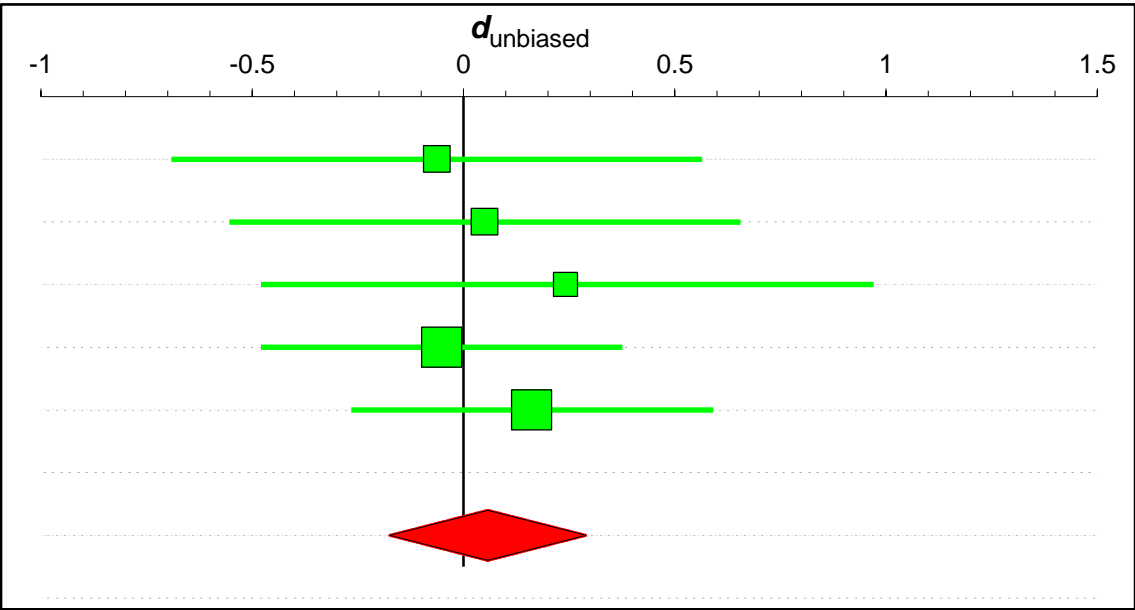

Figure 10.1. **Total recall** across the gesture-observed and gesture-not observed learning conditions (Study 1, Study 2-Time 1, Study 2-Time 2, Study 3-Time 1, Study 3-Time 2).

Figure 10.2. Recall **accuracy** across the gesture-observed and gesture-not observed learning conditions (Study 1, Study 2-Time 1, Study 2-Time 2, Study 3-Time 1, Study 3-Time 2)

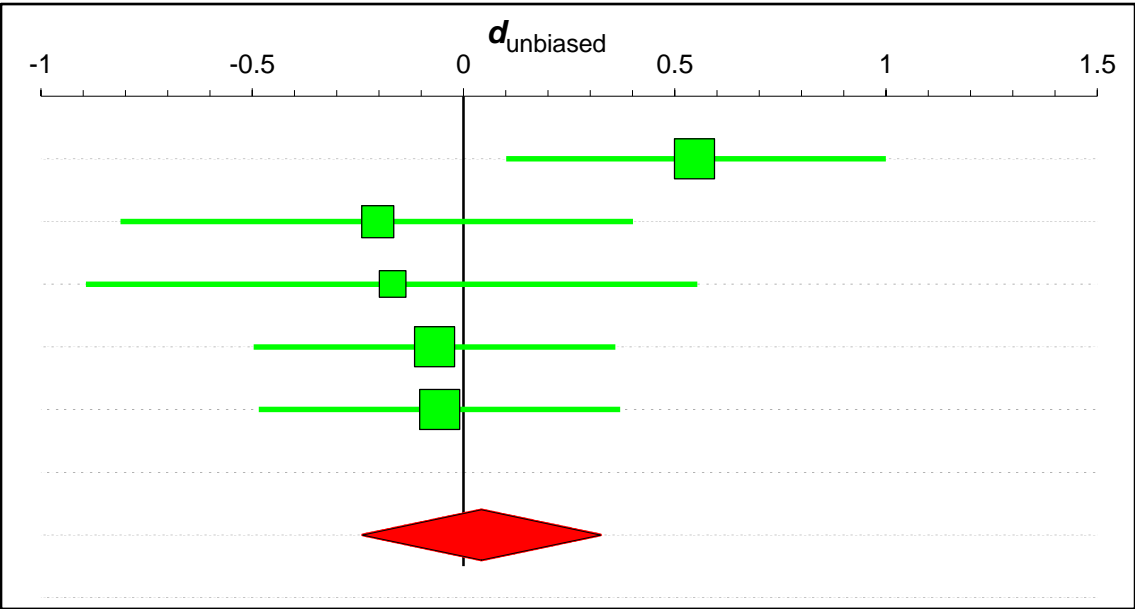---

[30] Figures for the recall of facts and concepts are in Appendix B, Supplementary analysis S5.2 (p. 193).

Gesture production as a mediator

From the results outlined so far, observing gesture during learning had no significant effect upon children's verbal recall. It was also of interest, however, to examine whether children's gesture production might mediate the relationship between observing gesture at learning and children's verbal recall at the short and long delays. In assessing the potential mediational effects of children's gesture production, both children's representational (Rep) and event-relevant (ER) gesture rates were examined. The representational gesture category assessed children's production of informative gestures, and the event-relevant category captured children's production of representational gestures that matched those observed at learning.

Children's gesture production was examined across all three studies. While Study 2 (Chapter 7) and Study 3 (Chapter 8) specifically assessed children's gesture production, Study 1 (Chapter 6) did not, but gesture coding was subsequently conducted (see Study 4, Chapter 9). It is important to note that the gesture coding scheme in Study 3 was different to the coding scheme used in Study 1 and 2. In Studies 1 and 2, gesture was coded with the audio content present to more closely represent the way gesture would be interpreted in the real world. In Study 3, however, gesture was coded without the audio content for a purer measure of gestural content. While combining results across two types of coding schemes is not ideal, it is common for meta-analyses to combine results across slightly differing methodologies to examine the same or similar research questions (Cumming & Calin-Jageman, 2016). In combining the data across the two studies, the potential biases of either coding scheme will likely be reduced.

In order to determine whether a mediation was occurring, three relationships needed to be established: 1) the relationship between the independent variable (IV) and the dependent variable (DV), 2) the relationship between the IV and the mediator, and 3) the relationship between the mediator and the DV.

The relationship between the IV (learning condition), and the DV (verbal recall) is depicted by Line A in Figure 10.3. As presented previously, the meta-analysis found no significant effect of learning condition (i.e. *gesture-observed* versus *gesture-not observed*) on the amount or accuracy of children's verbal recall. The same null effects were also found when facts and concepts were each examined. Additional analyses investigated the

effect of learning condition on children's recall, separately at both the short (1-day) and long (2-weeks and 7-months) delays. The same null effects were again found[31].

When there is no significant relationship between the IV and the DV, but both are associated with the mediating variable, then a mediation is likely occurring (Warner, 2012). Given the non-significant effect of learning condition (IV) on verbal recall (DV), three potential mediational relationships were examined: 1) whether children's Time 1 gesture production mediated the relationship between learning condition and Time 1 verbal recall (see Figure 10.3, Lines B and C); 2) whether children's Time 1 gesture production mediated the relationship between learning condition and children's Time 2 verbal recall (Figure 10.3, Lines B and D); and 3) whether children's Time 2 gesture production mediated the relationship between learning condition and children's Time 2 verbal recall (Figure 10.3, Lines E and F).
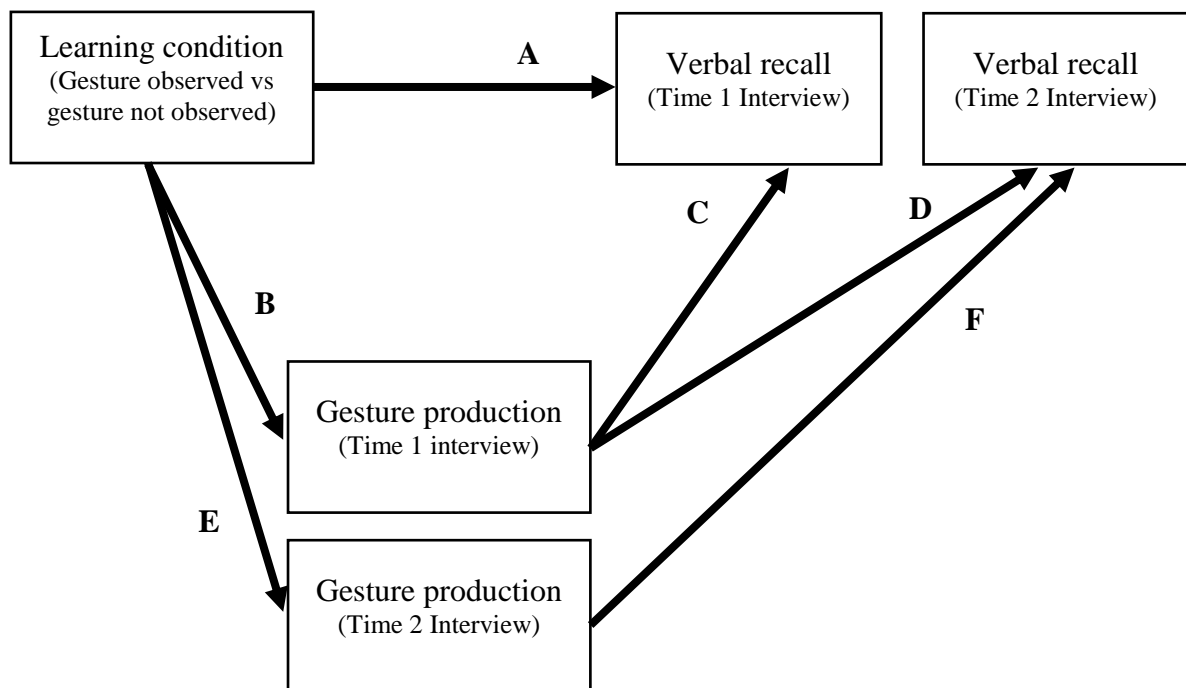

Figure 10.3. Depiction of the mediation analysis, with learning condition as the independent variable, gesture production rates as the mediator, and interview verbal recall as the dependent variable.

Children's Time 1 gesture production mediating the relationship between learning condition and verbal recall (Lines B and C/D). The first and second

---

[31] See Appendix B, S5.3 for supplementary analyses examining children's verbal recall across the learning conditions, separately at the short and long delays (pp. 194 – 198).

mediational effects involved establishing whether children's Time 1 gesture production mediated the relationship between learning condition and either children's Time 1 verbal recall or Time 2 verbal recall.

Learning condition and Time 1 gesture production (Line B). It was first examined whether children's Time 1 gesture production rates differed across the *gesture-observed* and *gesture-not observed* conditions (Figure 10.3, Line B). This would establish the relationship between the IV and the mediator. Both children's representational and event-relevant gesture rates were examined. Prior to analyses, any gesture rate variables that were not normally distributed were corrected with square root transforms; thus Cohen's *d* was analysed, with random effects (Cumming & Calin-Jageman, 2017). Again, children in Study 1 who were asked *wh*-questions were excluded from analyses.

Results indicated that, compared to children in the *gesture-not observed* condition, children in the *gesture-observed* condition produced both representational gestures ($d_{unbiased}$ = 0.35 [0.02, 0.68]) and event-relevant gestures ($d_{unbiased}$ = 0.48 [0.15, 0.82]) at a greater rate during the Time 1 interview, see Figure 10.4. This established a significant association between the IV (learning condition) and the mediator (gesture production rate).
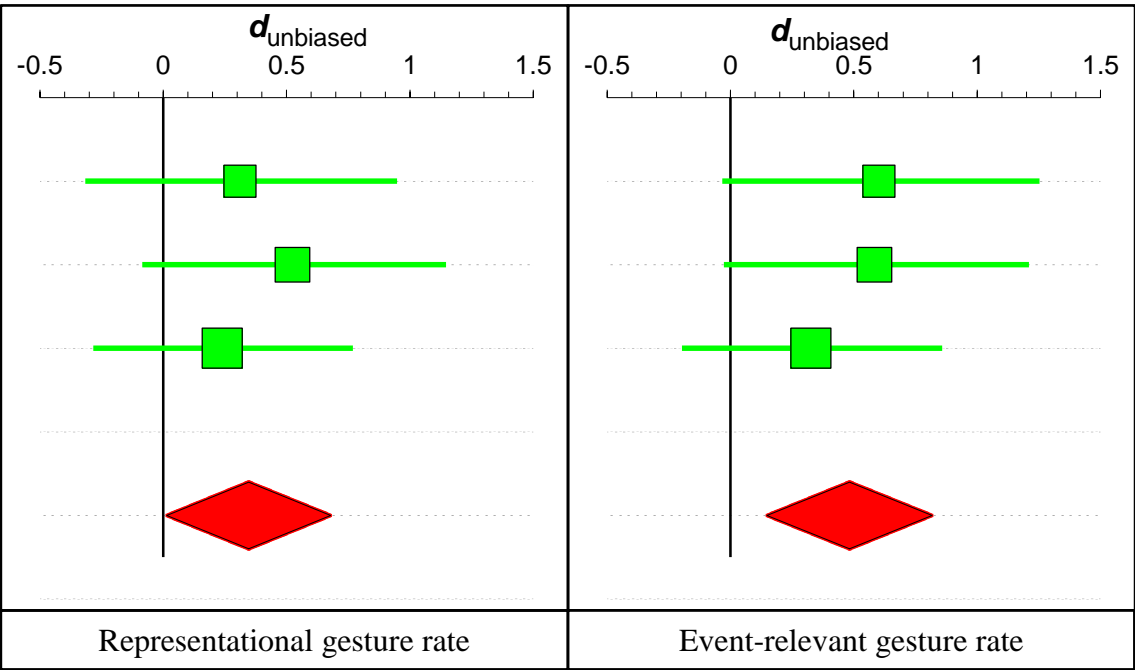

Figure 10.4. Children's Time 1 representational gesture rate (left panel) and Time 1 event-relevant gesture rate (right panel) across the learning conditions (Study 1, Study 2-Time 1, Study 3-Time 1).

***Time 1 gesture production mediating the relationship between learning condition
and Time 1 verbal recall (Line B and C).*** Given the significant association between
learning condition (IV) and children's Time 1 gesture production (Line B, Figure 10.3), it
was then established whether there was an association between children's Time 1 gesture
production (mediator) and Time 1 verbal recall (DV; Line C, Figure 10.3). It was
hypothesised that children's Time 1 gesture production would mediate the relationship
between learning condition and verbal recall the next day. Correlations were examined
between children's Time 1 gesture rates and Time 1 verbal recall. All variables were
normally distributed or corrected to normal with square root transforms.

It was found that children's Time 1 total recall was not associated with children's
Time 1 representational gesture rate ($r = .17$ [-.01, 0.34]), but was associated with
children's Time 1 event-relevant gesture rate ($r = 0.18$ [0.01, 0.35]), see Figure 10.5. Given
how close the confidence interval for the event-relevant gesture rate sits to the zero-line,
however, results must be interpreted cautiously. Comparatively, children's Time 1
accuracy was not significantly associated with either children's Time 1 representational
gesture rate ($r = -.01$ [-.18, .16]) or event-relevant gesture rate ($r = 0.10$ [-0.11, 0.31]), see
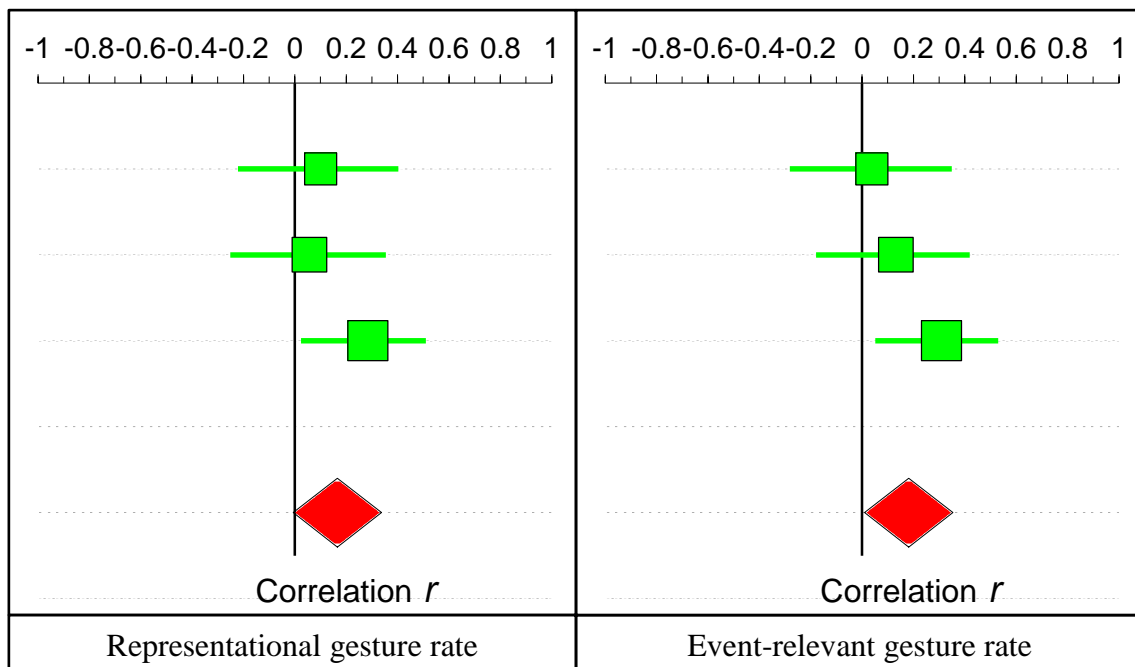Figures 10.6.

Figure 10.5. Correlations between children's Time 1 **total recall** and both children's Time
1 representational (left panel) and Time 1 event-relevant (right panel) gesture rates.
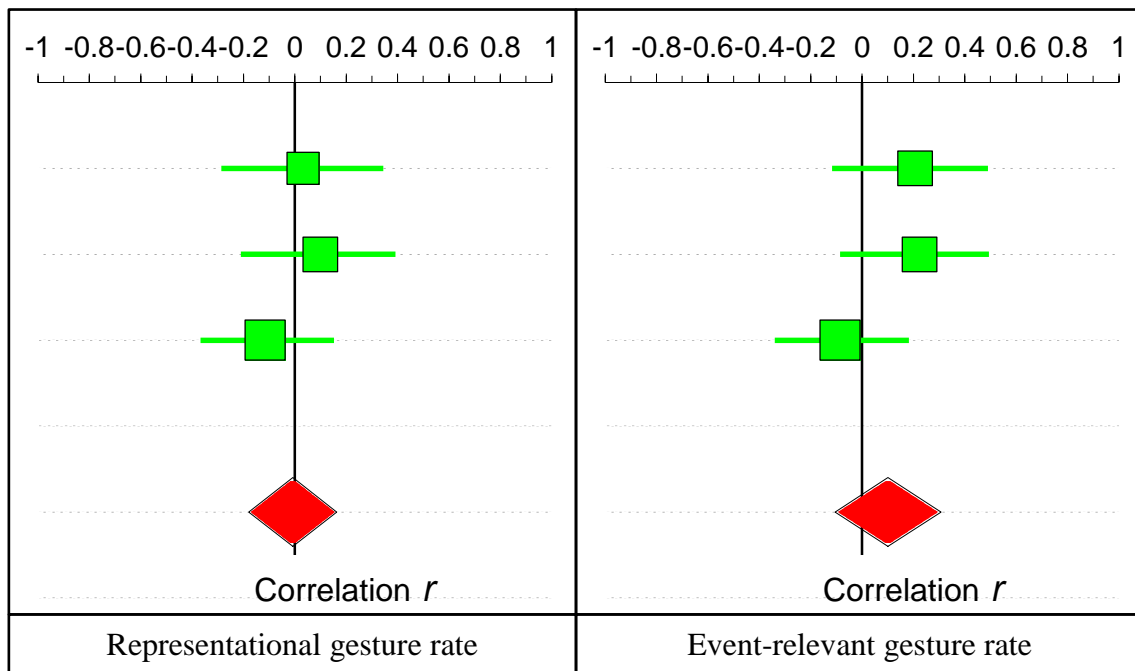
Figure 10.6. Correlations between children's Time 1 **accuracy** and both children's Time 1 representational (left panel) and Time 1 event-relevant (right panel) gesture rates.

When children's percentage recall of facts and concepts were separately examined, children's Time 1 recall of facts was not significantly associated with children's Time 1 representational gesture rate ($r = .17$ [-.001, .34]), but was significantly associated with the event-relevant gesture rate ($r = 0.21$ [0.04, 0.38]), see Figure 10.7. Children's Time 1 recall of concepts, however, was significantly associated with both children's Time 1 representational rate ($r = .21$ [.04, .38]) and event-relevant gesture rate ($r = .24$ [0.07, 0.41]), see Figures 10.8.

Together, results indicated that children's event-relevant gesture production might mediate the relationship between observing gesture at learning and improved verbal recall the next day. Children's representational gesture rate, however, did not appear to exert the same beneficial effect. When children observed gesture at learning, they produced both representational and event-relevant gestures at a greater rate during the interview the next day. While children's event-relevant gesture was associated with improved verbal recall (i.e. amount, facts, and concepts, but not accuracy) within the same 1-day delayed interview, children's representational gesture typically was not.
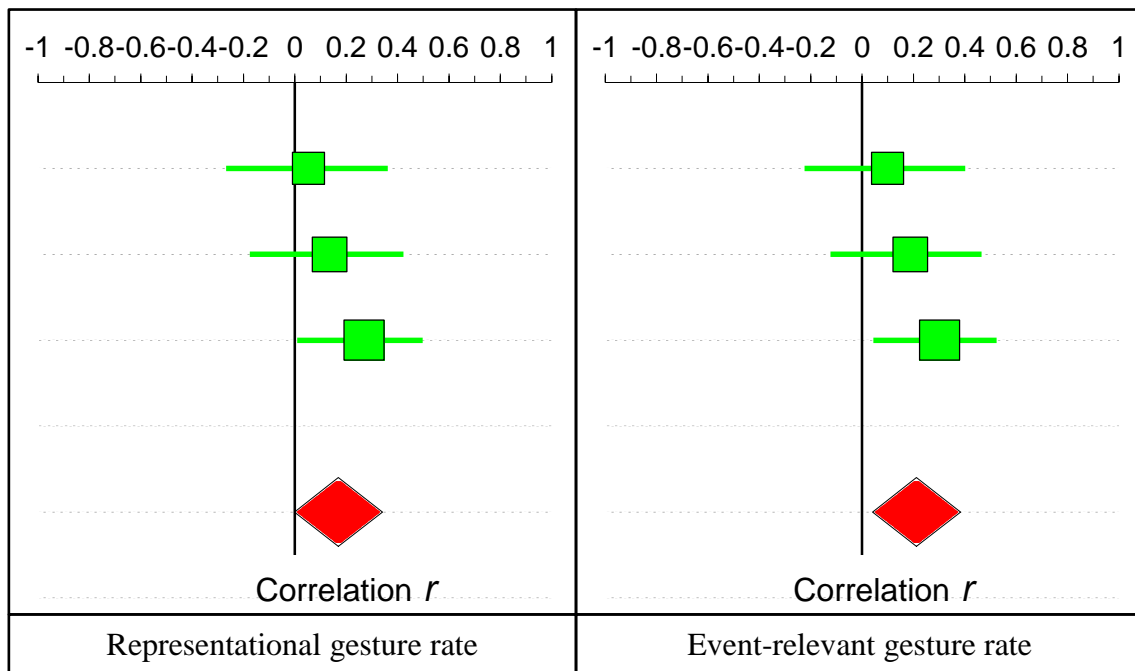
Figure 10.7. Correlations between children's Time 1 recall of **facts** and both children's Time 1 representational (left panel) and Time 1 event-relevant (right panel) gesture rates.
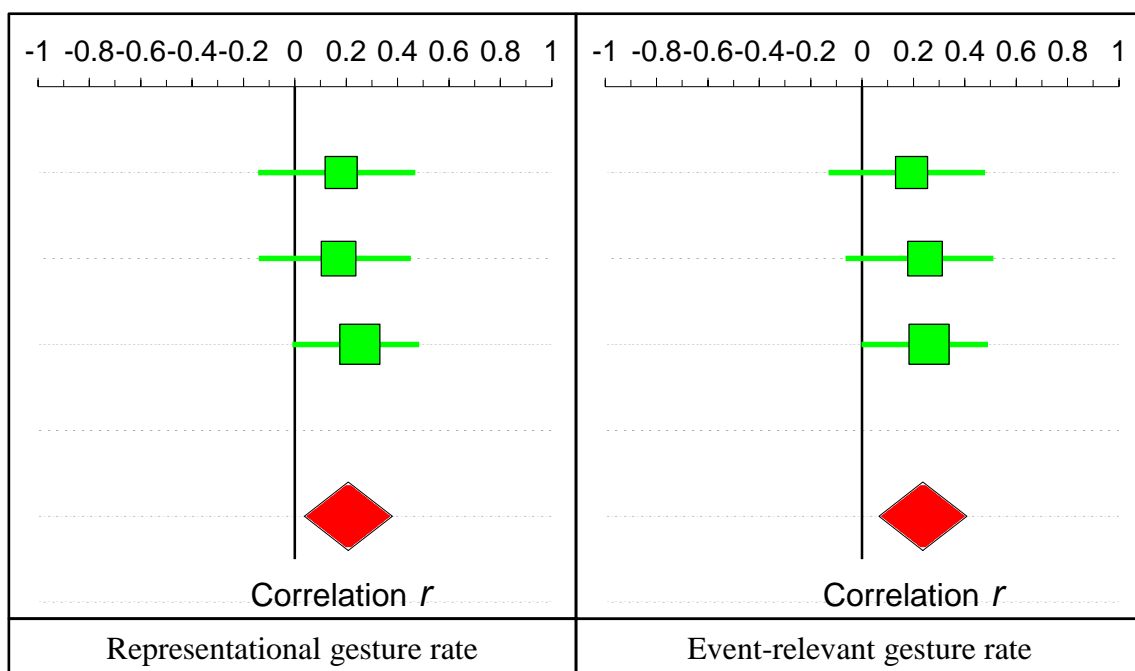

Figure 10.8. Correlations between children's Time 1 recall of **concepts** and, both children's Time 1 representational (left panel), and Time 1 event-relevant (right panel) gesture rate.

Time 1 gesture production mediating the relationship between learning condition and Time 2 verbal recall (Line B and D). Given the significant effect learning condition (IV) had on children's Time 1 gesture production rates (mediator; Line B, Figure 10.3), the association between Time 1 gesture and children's Time 2 verbal recall was also examined

145

(both 2 weeks and 7 months after learning; Line D, Figure 10.3). It was hypothesised that children's Time 1 gesture production rate would mediate the relationship between learning condition and children's Time 2 verbal recall. Correlations between children's Time 1 gesture and Time 2 verbal recall were conducted with random effects. All variables were normally distributed or corrected to be so with square root or log transformations. If variables could not be corrected, they were excluded from analyses.

First, correlations between children's Time 1 gesture and the amount and accuracy of children's Time 2 verbal recall were assessed. There was no significant association between children's Time 1 representational gesture rate and Time 2 total recall ($r = .18$ [-.03, .40]), but there was a significant, positive association between Time 1 event-relevant gesture and Time 2 total recall ($r = .22$ [.01, .44]). This significant result must be interpreted cautiously; as the confidence interval only just missed crossing the zero-line, see Figure 10.9. In terms of children's Time 2 accuracy, there were no significant associations with either children's Time 1 representational gesture rate ($r = -.09$, [-.31, .12]), or Time 1 event-relevant gesture rate ($r = -.05$ [-.26, .17]), see Figure 10.10.
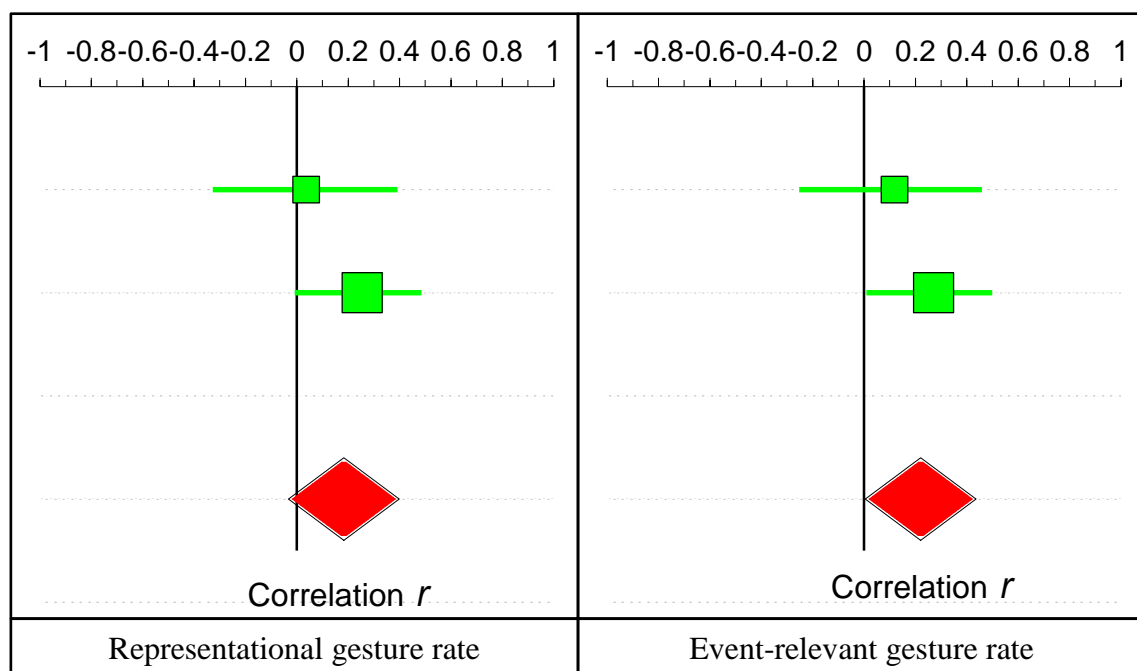

Figure 10.9: Correlations between children's Time 2 **total recall** and both children's Time 1 representational (left panel) and Time 1 event-relevant (right panel) gesture rates**.**
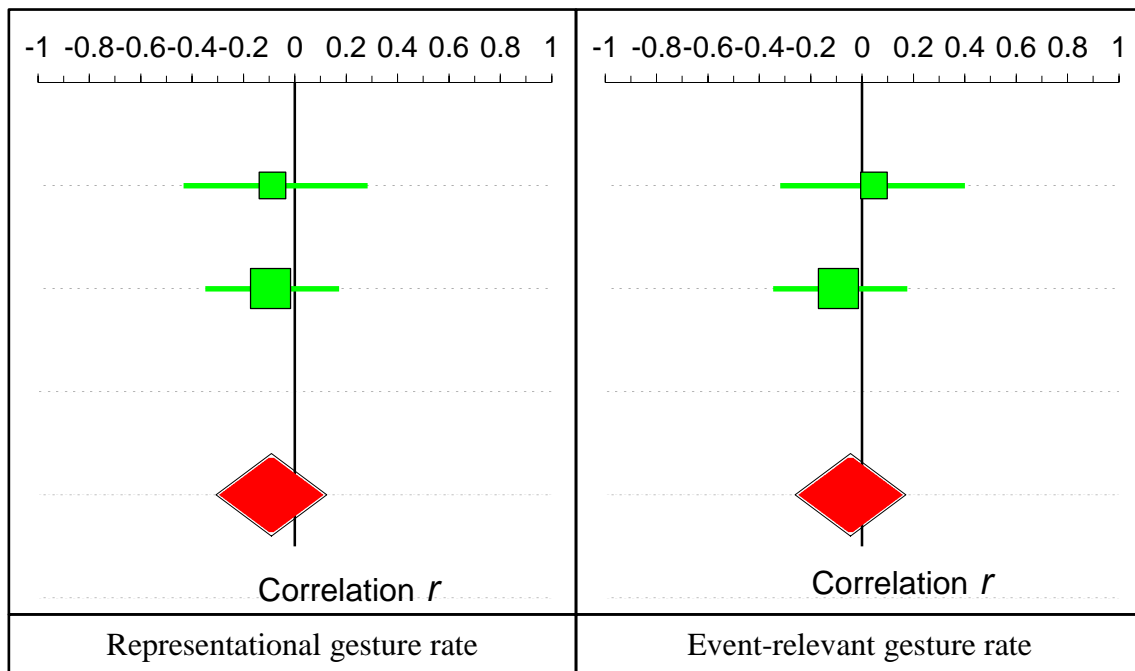
Figure 10.10: Correlations between children's Time 2 **accuracy** and both children's Time 1 representational (left panel) and Time 1 event-relevant (right panel) gesture rates.

The association between children's Time 1 gesture production and Time 2 recall of facts was also examined. Due to issues with normality, children's recall of concepts was not assessed. Children's Time 2 recall of facts was not predicted by either children's Time 1 representational rate ($r = .16$ [-.15, .46]) or Time 1 event-relevant gesture rate ($r = .23$ [-.02, .47])[32].

Together, results indicated that while children who observed gesture during learning typically produced gesture at a greater rate the next day, children's gesture during the Time 1 interview was rarely associated with improved recall at the longer (2-week and 7-month) delays. It appeared that only children's event-relevant gesture rate at Time 1 mediated the relationship between learning condition and the amount of correct information (i.e. total recall) children recalled following long delays.

Children's Time 2 gesture production mediating the relationship between learning condition and Time 2 verbal recall (Lines E and F). The third mediation examined whether children's Time 2 gesture production mediated the relationship between learning condition and Time 2 verbal recall (both two weeks and seven months after learning).

---

[32] See Appendix B, Supplementary analysis S5.4 for figure of the association between Time 1 gesture production and Time 2 recall of facts (p. 198).

147

First, the effect of learning condition (IV) on children's Time 2 gesture production (DV) was examined (Figure 10.3, Line E). Only children's event-relevant gesture rate was assessed, due to the distribution of the representational gesture rate not being normally distributed. It was found that learning condition (i.e. *gesture-observed* versus *gesture-not observed*) had no significant effect on children's Time 2 event-relevant gesture rate ($d_{unbiased}$ = 0.06 [-0.86, 0.99]), see Figure 10.11. The null effect of the IV on the mediator ruled out any potential mediational effects.

It was also of interest to examine whether children's Time 2 gesture rates were associated with Time 2 verbal recall (Line F). This association could not be examined, however, due to children's gesture rates in Study 2-Time 2 not being normally distributed.
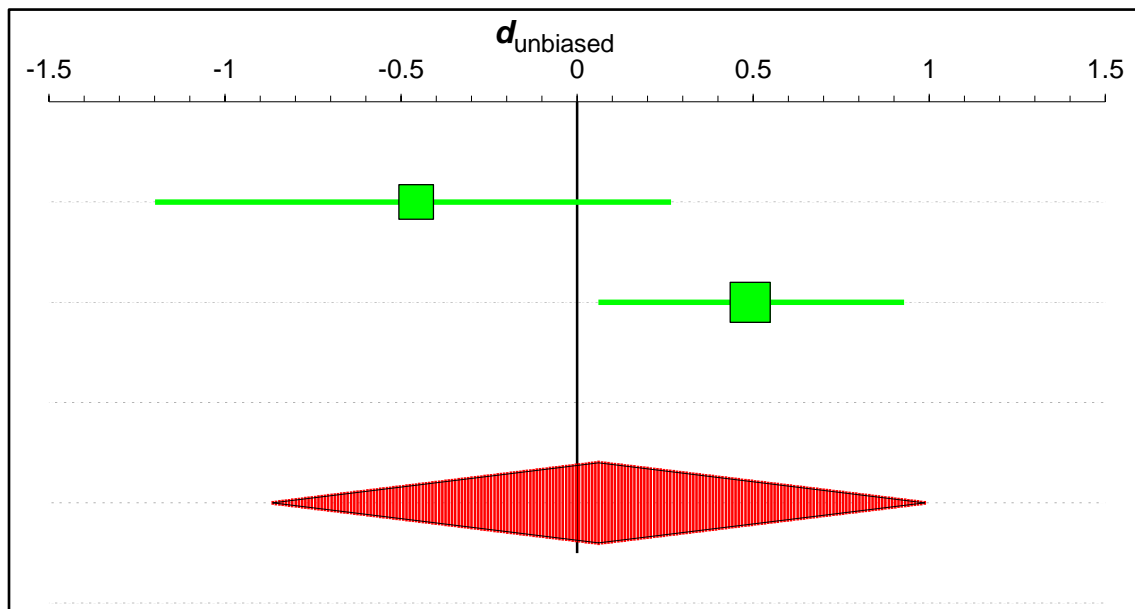

Figure 10.11. Children's Time 2 event-relevant gesture rates across the two learning conditions (2 weeks and 7 months later).

Summary. Three mediational analyses were carried out to examine whether children's gesture production rates mediated the relationship between observing gesture at learning and children's later verbal recall. Children's Time 1 event-relevant gesture seemed to mediate the relationship between learning condition and both children's Time 1 verbal recall (see Figure 10.3, Lines B and C) and Time 2 verbal recall (Figure 10.3, Lines B and D). Children's Time 2 gesture production, however, did not appear to mediate the relationship between learning condition and children's Time 2 verbal recall (Figure 10.3, Lines E and F). In many cases, however, the meta-analysis confidence intervals were close to the zero-line, so results must be interpreted cautiously.

Gesture and children's spatial recall

As described in Study 4, the amount of spatial language children included in their descriptions of the solar system lesson was variable across the studies. The overarching effects of observing and producing gesture on children's reporting of spatial terms were examined via meta-analyses. Again, children who answered *wh*-questions in Study 1 were excluded.

First, the influence of observing gesture on children's reporting of spatial content was examined. Cohen's *d* was entered into the ESCI and analysed using random effects. There was no significant difference across the *gesture-observed* and *gesture-not observed* conditions in terms of the amount of spatial terms reported ($d_{unbiased} = 0.30$ [-0.11, 0.70]). Any potentially significant effects, however, appeared to have been washed out by the results from Study 1, displayed in the top box plot in the left panel of Figure 10.12. In Study 4, it was highlighted that one of the main differences between Study 1 and the other studies was that visual information, other than gesture, was available. Presenting additional information in a similar modality may increase cognitive load (Sweller et al., 2011), and reduce the attention paid to either the gesture or the visual aids individually. When Study 1 was excluded from the meta-analysis, the *gesture-observed* condition was found to recall a greater percentage of spatial terms than the *gesture-not observed* condition ($d_{unbiased} = 0.51$ [0.26, 0.77]).

Second, the relationship between children's gesture production rates and their recall of spatial terms was assessed via correlation coefficients. Study 2-Time 2 gesture rates were excluded from analyses due to issues with normality. Children's reporting of spatial terms was significantly positively associated with both children representational gesture rates ($r = .20$ [0.07, 0.34]) and event-relevant gesture rates ($r = .19$ [0.05, 0.32]), see Figure 10.13.

Overall, children who observed gesture during learning appeared to report more spatial terms, if alternative visual information was not present. Then, when children were producing gesture at greater rates during the interview, they also reported more spatial information verbally within that interview.
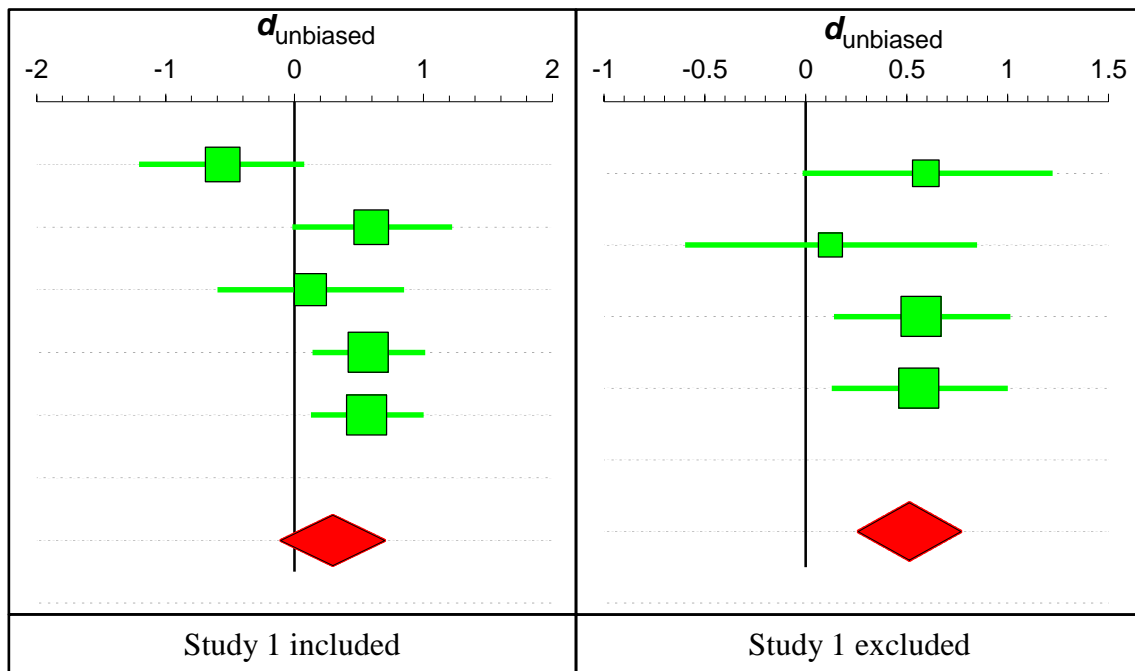
Figure 10.12. Meta-analysis comparing children's recall of **spatial terms** across the learning conditions (all five points of data included). The left panel displays results with Study 1 included, while the right panel excluded Study 1 from analyses.
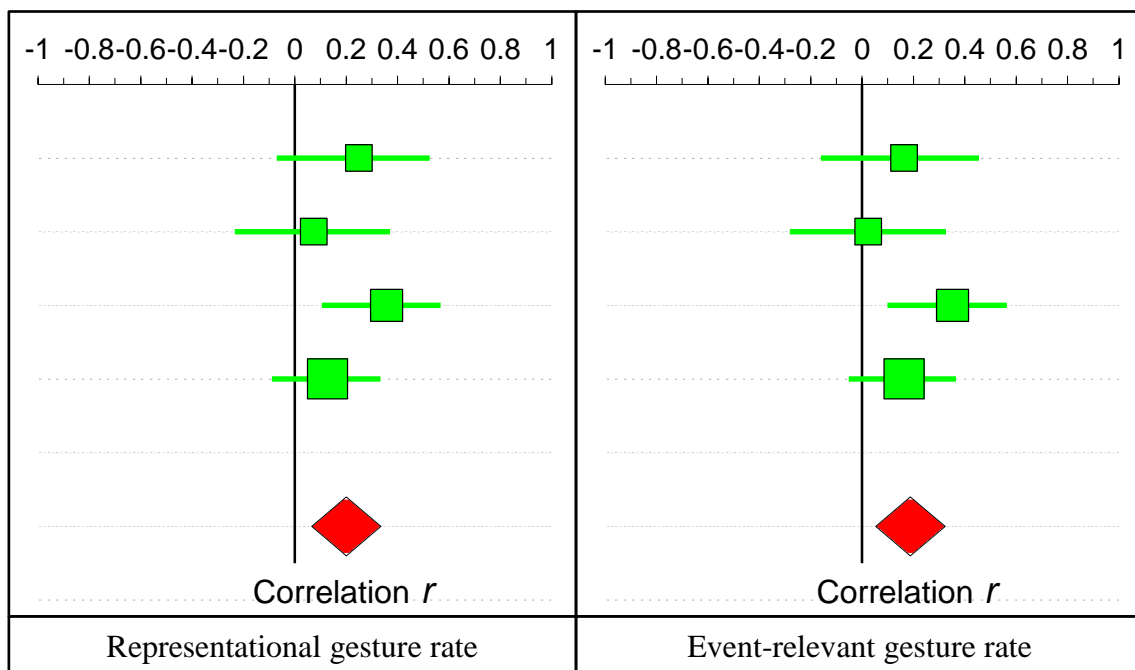
Figure 10.13. Correlations between children's recall of **spatial terms** and children's representational (left panel) and event-relevant (right panel) gesture rates (Study 1, Study 2-Time 1, Study 2-Time 1, Study 3-Time 2).

Summary

Children's event-relevant gesture rate may mediate the relationship between observing gesture at learning and children's verbal recall, both within and across interviews. When children observed gesture during learning, they reproduced the gestures they had observed. This process was likely implicit, but when children produced such gesture their verbal recall also appeared to improve. Representational gesture rates were comparatively less robust as a mediator.

Gesture was also beneficial for enhancing children's reporting of spatial content during the interview. When children observed gesture at encoding, they went on to report more spatial content in the interviews, at least when alternative visuospatial aids were not present. Children's own gesture production rates also increased children's reporting of spatial terms within the interview. These findings will be discussed further in the general discussion.

Despite the significant indirect effect observing gesture exerted on children's verbal recall, through children's own gesture production, the effect sizes were small and the confidence intervals were long and typically close to the zero-line (representing the null hypothesis). There was also substantial variation within and across studies. Results suggest that, even when gesture is effective at enhancing children's recall, the actual benefits created are likely to be small and improvements may vary depending on environmental or individual factors. The influence of observing and producing gesture on children's learning and recall, and the limitations, implications, and applications of these effects will be further discussed in the next chapter.

Chapter 11

General Discussion

The thesis examined the role of gesture, both observed and performed, in scaffolding 7- to 9-year-old children's learning and recall of a discursive scientific lesson. Previous research has found that when children observe gesture at encoding they better understand isolated concepts and are better able to recall simple narrative content compared to when gesture was not observed. Extending upon previous work, this thesis examined whether observing gesture, while encoding a discursive science lesson, would improve children's learning and recall of the information taught.

This thesis examined the potential benefits of observing gesture during learning on children's recall, both independently and when combined with other gestural and verbal scaffolding techniques. The major questions addressed were:

1. Does observing gesture during encoding improve children's verbal recall of a complex learning experience, compared to when children have not observed gesture? (Studies 1, 2, and 3)
   a. Does seeing gesture at encoding improve the amount (total recall) and accuracy (percentage accurate) of children's verbal recall the next day, compared to when children have not seen gesture?
   b. Does observing gesture differentially influence children's recall of facts and concepts?
   c. Does observing gesturing influence children's reporting of spatial information?
   d. Are any effects maintained two weeks and seven months later?
2. Do children's own gestures scaffold and enhance their verbal report? (Studies 2 and 3)
   a. Do children who gesture at greater rates report more information about the solar system, with greater accuracy, compared to children who gesture less?
   b. Do children who gesture at greater rates recall a greater percentage of the facts and concepts taught?
   c. Do children who gesture at greater rates also report more spatial information?
   d. Do children who gesture at greater rates verbal recall more information both within the same interview and across a delay?

3. Do children who have observed a variety of representational gestures accompanying interrelated facts and concepts produce more gestures similar to those seen, compared to those who have not observed gesture? (Studies 2 and 3).

Together these questions examined both the direct effect of observing gesture upon children's verbal recall, but also the indirect effect observing gesture may have on verbal recall by influencing children's gesture production. In answering these main questions, a number of additional issues were also addressed.

1. Whether the presence of visual aids within the learning experience decreased the effectiveness of gesture? (Study 1 versus Study 2 and 3).
2. Whether additive improvements in children's verbal recall were found when observed gesture was used in conjunction with asking *wh*-questions (Study 1).
3. Whether instructing gesture or restricting gesture further influenced children's verbal recall (Study 3).

The results, relating to these questions, will be discussed throughout this section and integrated with the meta-analysis findings.

Observing gesture and verbal recall

Studies 1, 2, and 3 all examined whether observing gesture would improve the amount and accuracy of children's verbal recall of the complex scientific information taught, compared to those who did not observe gesture. Additionally, children's separate recall of facts and concepts were also examined across the learning conditions. No significant improvements in verbal recall were found at any of the time points (i.e. one day, two weeks, or seven months later) and the same null effects were also established when the data was combined in the meta-analysis. When Study 4 (and the meta-analysis) examined children's recall of spatial terms, however, those who observed gesture recalled more terms than those who did not. This effect was only present when visual aids and props were not present during learning.

Observed gesture and recall of the solar system content

It was expected that the recall of the solar system content would improve when children observed gesture at learning, compared to when they did not. Methodological and developmental differences between previous research and the current studies may explain the null effects.

154

First, the information children were taught in the solar system lesson was designed to be unfamiliar and conceptually challenging. The fact that children typically recalled only about one-third of the content suggests this aim was met. Comparatively, previous studies required children to recall simple narrative information, such as a story (Macoun & Sweller, 2016) or directional route (Austin & Sweller, 2014). While gesture observation has been found to improve children's recall of conceptually simple information, it was not effective in the current studies.

To understand why lesson complexity may have influenced children's recall and the effectiveness of observed gesture, it is important to step back and examine the constructive process of memory. When individuals encode information, content must be linked together in a cohesive representation. At retrieval, the stored information then has to be appropriately recombined (Schacter, Norman, & Koutstaal, 1998). For example, in retrieving information about the pull of the sun's gravity, children might need to retrieve information about the sun in combination with information about gravity. When particular content nodes are activated, the activation also spreads to related ideas (Schacter, 2001). For example, recalling that the sun's gravity pulls the planets may activate the related idea that planets orbit the sun.

How well information is organised and stored at encoding will likely influence the ease with which the content is recombined at retrieval. Encoding efficiency may depend upon children's working memory resources and prior knowledge. When incoming information is simple and familiar, children may have the required resources to effectively implement memory strategies (Bjorklund et al., 1997). Additionally, when children have relevant prior knowledge they can use this information to inform how to chunk and store the information in order to improve later recall (Chi, 1978; Schneider & Sodian, 1997). While studies in this thesis did not include a measure of working memory capacity, it was found that children with higher knowledge scores showed superior recall of the lesson.

When the information taught was familiar, as in previous research, children may have been able to use their knowledge and cognitive resources to appropriately store an integrated verbal representation of the information. The associated gesture then may have enhanced the depth of encoding, providing additional retrieval cues for the children to draw upon. During the solar system lesson, however, it is possible that the novel and complex content may have taxed children's cognitive resources and left them with a less coherent framework for organising their encoding and retrieval. As a result, the piecemeal and disjointed lesson representation was then likely to be difficult to recombine at

155

retrieval. Observed gesture was hypothesised to circumvent some of these issues, by clarifying ambiguous verbal content and providing children with additional visual information to process and store in order to fill any knowledge gaps (Sekine et al., 2015; Thompson & Massaro, 1986).

The null effect of observing gesture during the solar system lesson may have been a result of the developmental stage of the participating children. Church et al. (2000) found that when children were around nine years old they were at a dip in a U-shaped developmental curve. The children in their study could separately encode and recognise the verbal and gestural content imparted, but had difficulty combining the information into an integrated message. The complexity involved in recalling the information taught may have revealed subtle developmental deficits in children's ability to integrate speech and gesture. Given the weak links forged across the modalities and the sociocultural emphasis placed upon understanding verbal information (McNeill, 1992; Sekine et al., 2015), children may have primarily relied upon recalling the verbal information stored. If the gestural content was activated at retrieval, verbal recall may not have been enhanced due to children finding it difficult to convert the visuospatial gestural information into speech. Alternatively, the weak links forged between gesture and speech may have meant that the gestural content could not spread activation to related verbal content for the children to report. Instead of gesture clarifying and filling knowledge gaps to enhance recall, children may have been left with a disjointed verbally encoded event representation.

Observed gesture and reporting spatial information

When children observed gesture during learning, they did appear to report more spatial information compared to when gesture was not observed. This raises the question; why observing gesture at encoding improved children's reporting of spatial content, but not their ability to recall the facts and concepts taught. Spatial terms are isolated words that children are typically very familiar with. Unlike facts and concepts, recalling such terms does not require integrating verbal and gestural content together within a comprehensive understanding. It could be speculated that when children observed gesture, the visuospatial information contained in these gestures was intuitively understood and stored along with a verbal label (Paivio, Rogers, & Smythe, 1968). For example, when children saw an orbit gesture they may also have encoded the term "around." In encoding a closely linked verbal and a visual representation children may have enhanced the depth of encoding and ease of retrieving these spatial terms. This finding also highlights the assertion that the difficulty

of the content taught reduced children's ability to integrate across the gestural and verbal modalities when recalling their learning more broadly.

Observing gesture and recall following a delay

Observing gesture was also expected to support children in better maintaining their learning across the delay. Previous research has found that children's problem solving actually improves following a 24 hour delay when children have observed gesture at encoding versus when they have not (Cook et al., 2013) and that visual information is less prone to decay than verbal content (Church et al., 2007; Paivio, 1971). Across all analyses, children who did and did not observe gesture at learning showed no difference in their recall or forgetting of the solar system facts and concepts at the extended 2-week and 7-month delays. In terms of children's reporting of spatial terms, however, Study 4 found that children who observed gesture at learning reported more spatial terms 2 week later, but there was no difference between the learning conditions at the 7-month delay.

It was previously theorized that children were at a developmental stage where they had difficulty integrating the gestural and verbal content into a comprehensive message. Instead the children may have primarily relied upon accessing and recombining the verbal content they had stored when recalling the lesson. If this was the case, it makes sense that learning condition also had no effect on children's lesson recall at the longer delays. Even if the gestural content was stored in a more robust manner (Paivio, 1971), if it could not be accessed and converted it into speech then no recall benefits would be observed.

When children observed gesture they recalled more spatial terms at the 2-week delay, but not the 7-month delay. There appears to be critical windows during which memories can be reactivated. Jones and Pipe (2002) found that when 5- to 7-year-old children were asked to recall a personal experience, their immediate recall did not significantly differ from their recall one month later. Six months later, however, recall had significantly decayed. In this thesis, children's encoding of gesture may have enhanced their recall of spatial terms at the 2-week delay due to the depth of encoding being enhanced (Paivio, 1971). At the 7-month delay, however, the children's memory of the lesson was likely more skeletal and reactivation of relevant memory-trace information was probably difficult regardless of whether or not gesture was present at encoding.

Summary

It was speculated, that the effectiveness of the observed gesture in improving children's recall was diluted due to both the developmental stage of the children and the difficult lesson content. If children had difficulty accessing the gestural content at recall

and converting it into speech, then it makes sense that the current studies found no improvement in children's recall of the solar system lesson when gesture was observed versus when it was not. One way children may access and test out links between the verbal and gestural content is to produce their own gestures. This is discussed in the following sections.

Performing gesture

Children's gesture production was examined in both Study 2 and Study 3, to assess whether children's own gesture may scaffold their recall of the scientific information taught. Across the two studies the findings were mixed. In Study 2, children's event-relevant gesture, during the short delay interview, predicted children's recall of concepts and recall accuracy within the same interview. In Study 3, however, children's gesture rate predicted children's total recall and the recall of facts. Similarly, variable findings also occurred at the 2-week and 7-month delays. While the effects were consistently in the same direction, with increasing gesture rates associated with improved recall, the large standard deviations found across the studies and relatively small sample sizes potentially created variability in whether or not results were significant across the studies.

To create a holistic understanding of the potential benefits children's gesture rates exerted on their solar system recall, all gesture data (including Study 1) were entered into a meta-analysis. At the short (1-day) delay, it was found that children's event-relevant gesture production was predictive of children's total recall, recall of facts, and recall of concepts. Children's representational gesture as a whole, however, was only significantly associated with the recall of concepts. Children's event-relevant gesture rate within the 1-day delay interview also predicted children's total recall following a longer delay (including both the 2-week and 7-month delay data), while representational gestures did not. These findings indicate that when children produce gestures similar to those seen during learning (i.e. event-relevant gestures), recall improves. Informative representational gestures more generally do not seem to exert the same significant effects. Given the interdependence of the two measures, however (i.e. event-relevant gestures are included in the representational gesture category), all gestural effects were in the same direction.

Study 4 examined associations between children's gesture production and the recall of spatial terms. The associations were typically non-significant. When combined in the meta-analysis, however, children's gesture rates (both representational and event-relevant) were associated with increases in the percentage of spatial information children reported.

Gesture production and recall within the interview

Children who performed gesture at greater rates showed improved recall of the solar system information. In particular, children's production of event-relevant gesture during the 1-day delay significantly predicted children's verbal recall, while increases in representational gestures typically did not (or at least not significantly). In order to outline why these differences may have occurred, the gesture types will be reviewed. Representational gestures were coded when the children produced any iconic, metaphoric or deictic gesture, while gestures coded as event-relevant only included representational hand movements which were similar in form and function to the scripted gestures used throughout the event. To give an example; if a child was talking about gravity and they produced a downwards pulling motion, this would be coded as a representational gesture. If a child produced a pulling motion inwards, however, this gesture would be coded as both representational and event-relevant, as the gesture conveyed correct information taught during the event.

Given that the event-relevant gestures included relevant and accurate information, it makes sense that producing such movements would also be related to improved verbal recall. When producing such gestures, children were likely cueing themselves with appropriate information (Frick-Horbury, 2002) and may have been able to reduce cognitive load while testing out their developing ideas (Cook et al., 2012; Kita, 2000). Alternatively, when children produced *any* representational gesture they could still test out ideas and cue their recall, but the subsequent verbal information they report may be irrelevant or inaccurate.

Interestingly, the meta-analysis showed that producing gesture at a greater rate, representational or event-relevant, was not significantly associated with improved overall accuracy. The lack of significant effect in the meta-analysis was likely driven by the Study 3 findings. There were two main differences between Study 3, and Studies 1 and 2; the first was the manipulation of children's gesture production and the second was the way gesture production was coded. When gesture production was manipulated in Study 3, children instructed to gesture provided less accurate verbal reports compared to children allowed to gesture or restricted from gesturing. By asking children to gesture, cognitive demands may have been increased, leaving children with fewer resources available for source monitoring, and potentially increasing the number of intrusions or distortions reported (Johnson, Hashtroudi, & Lindsay, 1993). Second, in Studies 1 and 2, children's gesture was coded as event-relevant if they produced one of the scripted gestures with

159

relevant verbal content, meaning accuracy was essentially built-in to the measure. In Study 3, however, children's gesture was coded without the associated audio, meaning children could produce relevant gestures without being verbally accurate. A third factor present across Studies 1, 2 *and* 3, was that children who produced gesture at a greater rate were recalling more information. It is likely that this increased recall also came with a few errors[33]. When results were combined in the meta-analysis the larger sample size in Study 3 may have washed out any positive associations between gesture production and accuracy.

Findings from the meta-analysis indicated that children also recalled more spatial information the more they gestured. When gesturing, children may have been more likely to attend to the spatial information conveyed in their hands (Rimé et al., 1984), and their access to spatial lexical terms may also have improved (Rauscher et al., 1996). Children's gesture production may also have enabled them to organise the spatial content in order to facilitate the process of speaking (Kita, 2000). When each study was separately examined in Study 4, however, very few significant correlations were found between children's gesture production and the percentage of spatial information they reported. The only significant correlations occurred either when children were specifically instructed to gesture (Study 3) or when seven months had passed since the original lesson (Study 2). At both these time points, recall may have been especially difficult. When children were instructed to gesture in Study 3, they were likely placed under greater cognitive load, and when children were recalling content 7 months after learning, the lesson memory trace appeared to be severely decayed. While gesturing seems to enhance the recall of spatial terms in general, when recall is more challenging children's gesture production appears to be particularly predictive of recall. This may have been due to children relying more heavily on the spatial cues provided visually and proprioceptively by gesture.

Gesture production across the delay

This thesis also examined the effect of children's gesture production across the interviews. The meta-analyses showed that when children produced event-relevant gestures at a greater rate during an earlier interview they recalled more information in the next, delayed interview. This was the only significant effect, however. There were no significant associations between children's event-relevant gesture rate and their later recall

---

[33] See Appendix B, S6.1 for correlations between error rates and word counts (p. 199)

of facts or recall accuracy[34], or between representational gesture rates in the 1-day delay interview and children's recall in the later interviews.

A number of mechanisms may explain the significant association between children's event-relevant gesture rate in the first interview and the amount the children recalled in the later interviews. First, children's gesture production may strengthen the lesson memory trace. When children embody their knowledge in their hands, they create a motor representation in memory. The more representations stored in memory, the easier a memory is to retrieve (Paivio, 1971, 1991). Additionally, when children gestured at the short delay interview, they likely improved their verbal recall in the moment. The better children's verbal recall in the earlier interview the more retrieval practice children were getting, which again may have enhanced the ease of recalling content following a delay (Karpicke & Blunt, 2011). Finally, children's own gesture production may also have allowed them to test out their burgeoning ideas and hypotheses in the first interview, which over time become realised in their speech (Broaders et al., 2007). Given that the finding was only present for total recall and that the confidence interval barely missed crossing the zero line (representing the null hypothesis), the robustness of this effect must not be overstated.

Summary

When children performed gesture, they recalled more information. Within the interview, children's event-relevant gesture production appeared to support recall. Children's gesture production may have freed cognitive resources and supported the organisation of information in order to facilitate verbalisation. There were also indications that children's event-relevant gesture production during an earlier interview predicted the amount of information children later recalled. This may have been due to children's gesture production strengthening the lesson memory trace through retrieval practice and the addition of a motor representation. Alternatively, producing gesture may have provided children with an opportunity to test out their emerging ideas, which later come be to understood and verbalised.

<div align="center">

Gesture production as a mediator

</div>

This thesis also examined, 1) whether children who observed gesture during learning would mimic the gestures they had seen at retrieval and, 2) whether children's

---

[34] Concepts were not assessed due to problems with the normality of the distribution

gesture production may mediate the relationship between observing gesture at learning and children's verbal recall.

In terms of children's gestural mimicry, the findings across the studies were variable. While Study 2 found that children who observed gesture produced more event-relevant gesture at the short delay, Study 3 did not replicate this finding. Instead, in Study 3, the *gesture-observed* condition only produced more event-relevant gestures than the *gesture-not observed* condition in the interview that occurred two weeks after learning. When results were combined in meta-analyses, results indicated that children who observed gesture produced gesture at a greater rate (both representational and event-relevant) during the 1-day delay interview, but the effect disappeared at the long delay (2 weeks – 7 months later).

Another aim of the thesis was to examine whether children's gesture production mediated the relationship between observing gesture at learning and improving children's verbal recall. When children observed gesture, they were found to produce gestures at a greater rate compared to those who did not observe gesture. Children's event-relevant gesture production at the short delay then predicted improved recall both within the same interview and across a longer delay (2 weeks – 7 months later). Children's representational gesture production, however, typically did not predict later recall (except for the recall of concepts).

Extant research varies as to whether children spontaneously mimic the gestures they observe. The studies which find gestural mimicry, for example Cook and Goldin-Meadow (2006) and the current studies, have included children who may be in a developmental stage where they have difficulty integrating the verbal and gestural content together (Church et al., 2000; Kelly & Church, 1998). Comparatively, studies which have not found gestural mimicry (Austin & Sweller, 2014; Ping & Goldin-Meadow, 2008; Valenzeno et al., 2003) have included younger participants. These children may not be in the same dip in the U-shaped developmental curve (e.g. Macoun & Sweller, 2016; Sekine et al., 2015). These studies have instead found direct effects of observing gesture on children's recall and problem solving.

Why might the ability to integrate gestural and verbal information affect whether children mimic the gesture they observe? Take for example, a child hearing the description that "gravity is a pull," accompanied by a gesture pulling towards the body. When children are able to integrate the verbal and gestural information into a comprehensive understanding, they may be able to combine all the information together as they verbally

report their learning (i.e. "gravity is an inwards pull"). As a result, observing gesture may directly improve children's verbal recall and problem solving without the need to produce gesture. When children have difficulty combining the messages into a comprehensive understanding, however, children may separately store the verbal and gestural information. The only way to retrieve visuospatial gestural content may be non-verbally through the children's own gesture (see Church et al., 2000). By gesturing (e.g. gravity comes inwards), children may then be able to structure the conversion of the spatial information into speech (Kita, 2000). When children have gestured during an earlier interview, improved recall benefits may be maintained as children were likely given the opportunity to combine the gestural and verbal content together during the earlier interview.

Representational gestures, as a broad category, typically did not mediate the relationship between observing gesture at learning and improved verbal recall during the interviews. When children mimicked the gestures they saw at encoding (i.e. event-relevant gesture) they were likely cueing themselves to report relevant and correct information (Frick-Horbury, 2002). Representational gestures more broadly, might not always include accurate or relevant information, however. This potentially reduced their effectiveness in enhancing children's verbal recall. There was, however, a significant mediational effect of children's representational gesture rate for the recall of concepts. This effect may have occurred because children were able to test out new ways of explaining the facts they had learnt.

The effect of visual aids.

Following the null effects of observing gesture in Study 1, it was hypothesised that the presence of the visual aids within the study (i.e. 3D representation of the solar system, diagrams, and interactive experiences) created additional cognitive load and reduced the uniqueness of the information presented in gesture. Studies 2 and 3 were thus conducted without any visual aids present. Children simply sat and listened to the verbal information taught, either with or without hand gestures. Even when the lesson was pared back, however, neither study found any direct effects of observing gesture on children's verbal recall.

Differences between Study 1 and the latter two studies indicate that the presence of visual information may still have had an influence on the results found, however. In Study 1, the accuracy of children's verbal report significantly improved when gesture was observed at encoding compared to when it was not. This finding was not replicated in

163

either Study 2 or Study 3. I speculate that the accuracy of the children's verbal report improved because the observed gesture grounded the information within the specific environment (Novack et al., 2014; Ping & Goldin-Meadow, 2008). Children may then have had a constrained lesson representation that was specific and potentially less prone to interference or distortion.

There were also indications that the presence of visual aids may have disrupted children's encoding and recall of visuospatial information they saw. For example, when Study 1 was excluded from the meta-analysis, the effect of observing gesture upon children's recall of spatial terms was significant. The exclusion of this study likely had a significant effect because children in Study 1 were actually reporting fewer spatial terms if they observed gesture (though the difference was not significant). Additionally, children's event-relevant gesture rate also appeared lower in Study 1 compared to the other two studies[35].

The inclusion of visual aids alongside gesture may have been problematic for a number of reasons. First, when both visual aids and gesture were present at encoding the visuospatial sketchpad may have become overloaded. Children likely had to process redundant information across the representations and integrate information into a single visuospatial representation, creating cognitive load (Sweller et al., 2011). Second, the uniqueness of the content conveyed in gesture was potentially reduced. The movement and visuospatial content conveyed in gesture was likely washed-out by the visuospatial nature of the visual aids and the movement involved in the interactive demonstrations. Finally, children children's attention may have been split between the gestural content and alternate visual information. As such, children may have attended to and encoded less information from either visual representation. When the children then went to recall information, they may only have had a fragmented visuospatial representation to structure their gesture production or verbal recall of spatial content.

Wh-questions and gesture.

In Study 1, the influence of adults asking *wh*-questions throughout the lesson was examined in combination with adult produced gesture. As both forms of scaffolding had previously been found to be effective in extant literature, Study 1 examined whether

---

[35] In Study 1, 56% of the representational gesture produced in the gesture-not observed condition were event-relevant and 67% for gesture-observed condition. The percentages of event-relevant gestures produced were even higher in Study 2 (70% for gesture-not observed, 79% for gesture-observed) and Study 3 (64% for gesture-not observed and 73% for gesture-observed).

additional benefits could be observed when the two forms of scaffolding occurred in tandem. Findings showed, however, that neither the observed gesture nor the *wh*-questions were particularly effective in enhancing children's learning and recall. Just as there are possible boundaries to when observing gesture is and is not effective, there seem to be similar limits to the conditions under which *wh*-questions are effective.

It was hypothesised that the *wh*-questions were not effective in supporting children's learning because they did not provide enough scaffolding to support children in giving correct answers. As a result, children may have strengthened incorrect retrieval paths and been put under additional cognitive load as they tried to update their understanding (Sweller et al., 2011). While the findings from Study 1 did not rule out the possibility that the effects of gesture may be enhanced by the talk adults use during learning, further research is required to examine the boundaries and conditions under which both *wh*-questions and gesture are separately helpful. Once the separate effects are better understood, the additive benefits can then be examined.

Manipulating children's gesture production

Given the significant role children's own gesture production seemed to play in aiding children's verbal recall of the solar system lesson, Study 3 specifically manipulated the gestures children produced. It was of interest to see whether instructing children to gesture could further enhance the beneficial effects of gesture, and whether restricting children's gestures would hinder recall.

When children were instructed to gesture, they were no better at recalling the learning experience compared to those who were either allowed to gesture or restricted from gesturing. In fact, if anything, the instruction to gesture was actually harmful for the accuracy of children's verbal recall, likely because it created additional cognitive load. When children were restricted from gesturing they recalled fewer facts than those able to gesture, indicating that the ability to move one's hands is probably important for the process of recalling information. The most beneficial effect on recall seemed to occur when children were simply allowed to gesture.

Implications and applications

In the introduction, issues were raised around the ecological validity of the research conducted to date. The current program of research taught children a variety of interconnected facts and concepts accompanied by a variety of different representational gestures in order to mirror the type of learning that may occur naturalistically in a

165

classroom. In order to advance the scientific community's understanding of gesture, this thesis aimed to answer a number of questions detailing the boundaries to when gesture is and is not effective in educational settings.

First, this thesis examined *whether observing gesture during encoding would improve children's recall of a complex scientific learning experience.* Based on the research we have conducted, the answer is no, not directly. Typically, children who observed gesture during the complex and comprehensive learning experience did not show superior recall in terms of the amount, or the accuracy of their verbal report, and there was also no difference in the separate recall of facts or concepts. There was an indication, however, that those who observed gesture recalled more spatial content.

Second, this thesis assessed whether *children's own gesture production enhanced their verbal report.* When children gestured while recalling their learning, their gesture rate, and in particular the rate of event-relevant gestures they produced, predicted the amount, but not the accuracy, of their verbal report. Both children's recall of facts and recall of concepts were also found to increase when children produced event-relevant gestures.

Third, *do children who observe a variety of representational gestures accompanying interrelated facts and concepts mimic the gestures they have seen?* The answer is yes. While children who observed gesture at learning generally produced representational gestures at a greater rate, children who had seen gesture at learning were also more likely to mimic the gestures they had previously seen. Those event-relevant gestures were also found to mediate the relationship between observing gesture and recalling more information.

It was also examined whether *observing and producing gesture continued to benefit children's recall across delays longer than 24 hours.* Following a delay of either two weeks or seven months, there were no direct effects of observing gesture upon children's verbal recall. When children observed gesture at learning, they did produce more event-relevant gestures at the 1-day delay, however. Such gestures were then found to predict children's verbal recall both within the interview and following a longer delay.

In addition to the four main questions, this thesis was interested in examining a number of subsidiary issues. The first was *whether additive improvements in children's verbal recall were found when observed gesture was used in conjunction with asking wh-questions.* Study 1 did not find any additional recall benefits when *wh*-questions and

observed gesture occurred in tandem. The absence of improvement was likely due to the way the *wh*-questions were implemented.

The second subsidiary question examined *whether the presence of visual aids within the learning experience decreased the effectiveness of gesture?* While direct comparisons were not made across the studies in terms of the influence of visual aids, no beneficial effect of observing gesture was found, either when visual aids were present or absent. Across the studies there were indications, however, that the presence of visual aids may have disrupted children's processing of the visuospatial information. Further research, directly contrasting the effects of gesture with and without additional visual aids, is required.

Finally, children's gesture production was manipulated in Study 3, with the intention of answering *whether instructing gesture or restricting gesture would further influence children's verbal recall*. Instructing children to gesture did not result in improved recall, and restricting children from gesturing had a negative effect on children's recall of facts. Children's gesture production was most effective when children could move their hands as they naturally would.

Results across the studies suggest that children's own gesture production seems to drive the beneficial effect gesture has on children's verbal recall. Adults should still be aware of the hand movements they are producing when instructing children, however, as these gestures likely set the stage for children's own gesture, and subsequent verbal recall. In order to optimise the effectiveness of the hand gestures children produce, it may be important for educators to plan not only their verbal explanations but also the accompanying information contained in their hand movements. In addition, educators should also ensure that whatever gesture they produce is readily available to children and not obscured by desks or lecterns. This may be especially important when selecting or designing videoed educational content, as hand gestures are often not in the camera shot. Finally, adults should also be aware of the visual aids they use in conjunction with gesture. Unless the visual information is integral to the task at hand, the presence of competing visual information may reduce any beneficial effect observed gesture may exert on either children's gesture production or verbal recall.

If children are to orally recount their learning, teachers and adults should ensure that children are able to move their hands as they do so. Sitting behind a desk which restricts movements or allowing children to talk with their hands in their pockets or arms folded may limit the improvements gesture can exert on children's verbal recall. At the

same time, however, specifically asking the children to gesture while explaining their learning is unlikely to be helpful. Instead, children should be left to gesture how they choose. Given the large standard deviations noted when children's gesture rates were examined, however, there are likely to be individual differences in terms of how effective spontaneous gesture may be for learning. This will be discussed in the forthcoming sections.

<div align="center">

Limitations and future research

</div>

Sample size

The sample sizes within this programme of research were selected by consulting previous literature. Studies have typically found significant effects of gesture with up to 25 participants per condition, with some finding significant effects with as few as 12 or 13 participants. Despite the justification for the sample size chosen, the non-significant results may have been due to the small sample sizes and reduced power to detect subtle effects. When the results were merged in the meta-analyses, however, the results across the studies were typically similar to the combined effects found but some of the trends became significant. This suggest that the results from Studies 1, 2, and 3 were valid, but may potentially underrepresent the effect observing and producing gesture has on learning.

Lack of a pre-test

The current studies did not include a pre-test to measure children's solar system knowledge prior to the lesson. The reason this was done was to avoid priming children into the important content before they participated in the lesson. Prior to developing the lesson, a pilot test was conducted, however[36]. This was done to establish what children 7 – 9 years old generally did and did not know about the solar system and to ensure the learning experience was both novel and developmentally appropriate. By including a pre-test, however, *changes* in knowledge could have been assessed. This may be an interesting area of future research.

Individual differences

Within and across conditions in the studies conducted, there was considerable variation in the rate at which children gestured. It would have been interesting to include individual difference measures to examine whether children's propensity to gesture influenced the effectiveness of the observed and performed gesture manipulations. Studies have found that those who have a lower working memory capacity are more likely to

---

[36] See Appendix A for information about the pilot test (p. 173)

gesture during spatial transformation tasks (Chu et al., 2014; Pouw et al., 2014), and their performance is more likely to be affected by gestural restrictions (Marstaller & Burianová, 2013; Pouw et al., 2014). It may be the case that within this thesis, those who were more prone to gesture showed a greater detriment in their recall performance when restricted from gesturing, compared to those who typically did not gesture. Conversely, those who typically did not produce gesture may have shown poorer recall when they were instructed to move their hands, compared to those who naturally produced gesture. Understanding how individual gestural differences influence gestural manipulations would indicate who to target with gestural scaffolding. Taking individuals differences into account may also clarify some of the disparate findings within the literature.

Coding gesture production during the lesson

While anecdotally it was noted that children rarely gestured during the learning experience, it would be interesting for future research to code this information to see whether there was any correlation between the children who did and did not gesture during the lesson and their later recall. Further research could also look at manipulating the gestures children produced *during learning* within a similar research paradigm.

Type of gestures produced during learning

The solar system lesson was accompanied by a variety of representational gestures which naturally fitted with the content taught. The gestures were all scripted and were fleeting in nature. Some gestures were repeated more than others throughout the lesson due to the verbal content they accompanied. For example, the orbit gesture was repeated a number of times in relation to the orbit of the planets, the orbit of the sun, the distance planets travelled around the sun etc. Comparatively, a criss-cross gesture was only produced once while describing how the orbits do not overlap. Previous research has often repeated the same gestures with the same verbal content (e.g. Cook et al., 2013) or has held the gestures for long amounts of time (Macoun & Sweller, 2016). In these cases, observing gesture was found to be beneficial for learning.

It would be interesting to manipulate the repetition of gesture or the amount of time gesture was held for within the same solar system lesson. It may be the case that if gesture was repeated and held, children would be better able to integrate the verbal and gestural content into a cohesive understanding in order to improve recall. Additionally, reducing the number of gestures and linking the same gesture to specific words or phrases may also have a similar beneficial effect.

Gesture coding

Across Studies 2 and 3, there was a slight difference in the way the gestures were coded. In Study 2, only speech-gesture matches were included in the event-relevant representational gesture category. When coded this way, the event-relevant gesture rate within the short delay interview significantly mediated the relationship between observing gesture at learning and improved recall within the same interview. In Study 3, however, gesture was coded without the audio information. As a result, both speech-gesture matches and speech-gesture mismatches were included as event-relevant. Study 3, only found a mediational effect of producing gesture within the interview conducted two weeks after learning.

Further research should examine the influence observing gesture has on the production of speech-gesture matches and mismatches. For example, it would also be interesting to examine whether children similarly progress from producing mismatching gestures (gesture correct but speech incorrect) before they produce matching representational gesture (where both speech and gesture is accurate) across the delayed interviews.

Developmental research

Further research is also needed to examine the effect of children's developmental stage on the ability to integrate gestural and verbal information. In this thesis, it was proposed that in middle-childhood children descended into a U-shaped developmental curve. As a result, children may have had difficulty integrating the verbal and gestural information at encoding and converting across the modalities at recall. These children were suggested to be particularly likely to benefit from producing gesture, which enabled them to test out connections between ideas and aid the verbalisation of their understanding. Further research is needed to directly test this assertion. Specifically, does the mediational effect found when recalling discursive scientific content occur across all developmental stages, or is 7 – 9 years of age a particularly special developmental period. Cross-sectional or longitudinal research could answer this question.

The influence of visual aids

Throughout the thesis, the potential issues associated with observing gesture in the presence of visual aids were repeatedly raised. Future research should directly test whether attention and working memory are affected when children are required to both observe gesture and process alternative visual information. If visual aids are found to reduce the effectiveness of observed gesture in paradigms where gesture is typically effective for

enhancing learning, further research should be conducted to examine potential mechanisms. Eye tracking tasks could assess whether attention is affected, while dual-task paradigms could examine the load placed on cognitive resources.

Conclusions

Collapsing the results across the studies, findings indicated that when children observed gesture during learning they appeared to pick up on the gestures observed and reproduced them. In doing so, children's gesture production appeared to enhance the verbal recall of the complex, discursive, scientific lesson. Observing gesture was only effective in improving recall because it encouraged children to gesture in appropriate ways. Given the small sample sizes, however, improvements in children's recall were likely negligible.

Previous research has highlighted the fairly robust role gesture plays in supporting children's learning. When examining the individual studies carried out in this thesis, however, the picture painted is one of inconsistency. While results were typically in the same direction across the studies, sometimes effects were significant and at other times they were not. Further research is needed to examine when gesture is and is not effective during learning, especially within complex learning environments. It is important to establish how educators can make the most of moving their hands in order to encourage learning and recall. At this point, however, the benefits gesture may exert on children's learning and recall appear subtle and variable.

Appendix A

Solar system pilot testing

The pilot study was conducted with 44 children, 7 – 9 years of age (M_{months} = 100.42, SD= 6.95). Children were recruited from two local Wellington primary schools and parents completed the parental knowledge and interest rating scales (described in Chapter 6). The pilot test asked children to complete a 29-item solar system test to assess children's knowledge about various topics relating to space and the solar system. Participating children individually answered the questions, with required responses ranging from one word (e.g. "which planet is the biggest?"), to complete sentences (e.g. "how do we get day and night on earth?").

To check the validity of the parental rating scales, associations between children's solar system test score and their knowledge and interest ratings were assessed. Significant positive correlations were found between children's performance on the solar system test and parental ratings of both knowledge ($r(41)$ = .58, p < .001) and interest ($r(41)$ = .33, p = .034). These findings suggest that the parental rating scales were valid, and appropriate for use as an indicator of children's solar system knowledge.

Results from the 29-item test were used to decide which solar system topics should be taught during the science lesson. The aim was to include information that was novel and complex, but still fell within 7- to 9-year-old children's zone of proximal development. It was found that children correctly responded to over half the questions relating to: Properties of the sun (M = .89, SD = .21), properties of the planets (M = .57, SD = .29), day and night (M = .63, SD = .36), and heat in the solar system (M = .72, SD = .29). Given children's level of knowledge about these topics, this information was not taught during the solar system lesson. Children correctly responded to less than half the questions relating to: Planet movements (M = .42, SD = .29), gravity (M = .47, SD = .32) and types of planets (M = .27, SD = .29). These topics were therefore included in the solar system lesson, along with a basic introduction detailing the names, locations, and some properties of the planets.

Appendix B

Supplementary analyses

Study 1

S1.1: Children's recall of names and contextual information

Coding. Children's percentage recall of both names and contextual information were also examined. *Names* included specific scientific terms (e.g. "gravity" or naming the various planets). *Context* included recalling content about the setup of the room and how the lesson played out (e.g. "we learnt about the solar system in the staffroom" or "I stood in a box of breath." Children's percentage recall of names and contextual information were assessed across the three main lesson components. There were 10 codes for names (total score = 20) and 35 codes for context (total score = 70).

Results. Children's recall of names was not normally distributed, and transforming the data did not resolve this (see Table B1). These scores must therefore be interpreted with caution. A 2 (Gesture: Observed vs not observed) $\times$ 2 (*Wh*-questions: present vs absent) factorial ANOVA was run with names as the dependent variable and knowledge ratings included as a covariate[37]. There was no significant main effect of gesture ($F(1, 74)$ = 2.73, p = .103, η_p^2 = .04), or main effect of *wh*-questions ($F(1, 74)$ = 2.48, p = .120, η_p^2 = .03) on the percentage recall of names. There was also no significant interaction between observing gesture and answering *wh*-questions (p's > .05), see Table B1.

A 2 (Gesture: Observed vs not observed) $\times$ 2 (*Wh*-questions: present vs absent) factorial ANOVA was also run, with contextual recall as the dependent variable. There was no significant main effect of gesture on the percentage recall of contextual information ($F(1, 75)$ = 1.54, p = .218, η_p^2 = .02). There was, however, a significant main effect of *wh*-questions ($F(1, 75)$ = 7.85, p = .006, η_p^2 = .10), with children who answered *wh*-questions recalling less contextual information than those who did not answer *wh*-questions, see Table B1. There was also no significant interaction between observing gesture and answering *wh*-questions (p > .05).

---

[37] There was a significant correlation between children's recall of names and their parental knowledge score, $r(77)$ = .25, p = .028

Table B1

Means (Standard Deviations) and Kolmogorov-Smirnov Tests of Normality for the Percentage Recall of Names and Contextual Information Across The Conditions

		Names		Context	
		M(SD)	*D*	*M(SD)*	*D*
Gesture-not observed	No *wh*-questions	63.74 (22.40)	.16	34.25 (13.93)	.10
	Wh-questions	53.89 (28.70)	.12	24.50 (17.13)	.14
Gesture-observed	No *wh*-questions	69.44 (14.28)	.23[*]	30.45 (20.27)	.13
	Wh-questions	63.89 (20.31)	.15	18.82 (15.70)	.19
Total	No *wh*-questions	66.67 (18.66)		32.30 (17.35)	
	Wh-questions	58.89 (25.06)		21.66 (16.47)	
	Gesture-not observed	58.69 (25.97)		29.25 (16.21)	
	Gesture-observed	66.67 (17.56)		24.63 (18.84)	
	Total	62.73 (22.34)		26.91 (17.63)	

[*] *p* < .05

S1.2: Kolmogorov-Smirnov test of normality for the verbal recall measures

Table B2

Kolmogorov-Smirnov Tests of Normality for Verbal Recall Across The Conditions

		Total recall	Facts	Concepts	Accuracy
Gesture-not observed	No *wh*-questions	.22[*]	.11	.12	.19
	Wh-questions	.14	.13	.15	.17
Gesture-observed	No *wh*-questions	.09	.11	.12	.17
	Wh-questions	.13	.12	.16	.22[*]

[*] *p* < .05

Study 2

S2.1: Children's recall of names across the conditions

Children's percentage recall of names (i.e. scientific terms) was also examined across the conditions. *Names* included recalling specific words (e.g. "gravity" or naming the various planets). Across the three main lesson components there were 9 codes

assessing children's recall of names (total score = 18). Kolmogorov-Smirnov tests of normality showed children's distribution of the percentage recall of names was not normal at Time 1, so was corrected with a square root transformation (See Table B3). When analysing variables across the two time points, if a square root correction was applied to a variable at T1 it was also applied at T2, and vice versa.

Children's recall of names was examined across the *gesture-observed* and *gesture-not observed* conditions at both the 1-day and 7-month delay. Independent samples *t*-tests revealed no significant differences between the conditions for the recall of names at Time 1 ($t(40) = 0.70$, $p = .486$, $d = .21$), or Time 2 ($t(28) = -0.24$, $p = .811$, $d = .09$), see Table B3.

A repeated measures ANOVAs was also run with delay (i.e. names at T1 vs T2) as the within-subjects variable and learning condition (i.e. *gesture-observed* versus *gesture-not observed*) as the between-subjects variable. Children recalled fewer names across the delay ($p < .001$), and there was a significant interaction between learning condition and delay ($F(1, 28) = 4.49$, $p = .043$, $\eta_p^2 = .14$). Post hoc *t*-tests, with a Bonferroni correction (.05/2 = .025), were conducted within the sample of children who completed both interviews. While the difference did not reach significance, there was a trend that, at Time 1, children recalled more names if they did not observe gesture ($M = 46.79$, $SD = 14.25$) than if they did observe gesture ($M = 34.80$, $SD = 15.38$), $t(28) = 2.20$, $p = .036$, $d = 0.82$. At Time 2, there was no significant difference in the recall of names between the children who observed gesture ($M = 23.04$, $SD = 15.25$) and those who did not ($M = 21.79$, $SD = 12.05$; $t(28) = -0.12$, $p = .905$, $d = -0.04$).

Table B3

Means (Standard Deviations) and Kolmogorov-Smirnov Tests of Normality for the Percentage Recall of Names at Time 1 and Time 2

	Time 1		Time 2	
	M(SD)	D	*M(SD)*	D
Gesture not observed	41.67 (18.07)	.16	21.79 (12.05)	.14
Gesture observed	35.71 (14.47)	.20[*]	23.04 (15.25)	.19

[*] *p* < .05

S2.2: Kolmogorov-Smirnov tests of normality for the verbal recall data at Time 1 and Time 2

Table B4

Kolmogorov-Smirnov Tests of Normality for Verbal Recall at Time 1 and Time 2 (with outliers removed)

	Total recall		Accuracy		Facts		Concepts	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Gesture not observed	.16	.21	.16	.14	.16	.25[*]	.20[*]	.27[*]
Gesture observed	.10	.20	.16	.13	.11	.16	.16	.30[*]

[*] $p < .05$

S2.3: Kolmogorov-Smirnov tests of normality for gesture production at Time 1 and Time 2

Table B5

Kolmogorov-Smirnov Tests of Normality for Gesture Production at Time 1 and Time 2

	Representational rate		Event-relevant count		Event-relevant rate	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Gesture not observed	.17	.15	.22[*]		.19[*]	.20
Gesture observed	.15	.24[*]	.21[*]		.21[*]	.20

[*] $p < .05$

S2.4: Time 1 correlations between child gesture and verbal recall

      Correlations were also conducted to examine associations between the children's gesture production and verbal recall within the same 1-day delayed interview. There was a significant positive correlation between children's event-relevant gesture count and the recall of concepts. None of the other relationships were significant, though there was a trend that children's event-relevant gesture rate positively correlated with children's recall accuracy ($p = .061$), see Table B6.

Table B6

Correlation Matrix Examining the Associations Between Children's Verbal Recall and

Gesture Production During the Time 1 interview

	Facts	Concepts	Accuracy	ER count	Rep rate	ER rate
Total recall	.94[**]	.83[**]	.47[**]	.21	.06	.09
Facts		.85[**]	.52[**]	.25	.14	.13
Concepts			.39[*]	.35[*]	.17	.22
Accuracy				.24	.11	.29
ER count					.91[**]	.95[**]
Rep rate						.92[**]

[*] $p < .05$ [**] $p < .01$

Study 3

S3.1: Analyses examining children's recall of names

The 1-day delay. The percentage of names (i.e. scientific terms) children recalled was assessed across the three main event phases (i.e. what makes a planet, gravity, and the surfaces of the planets). The distribution of names was not normal and could not be corrected (see Table B7), so analyses were carried out with raw scores (see Schmider et al., 2010). In assessing children's recall of names, a 2 (Learning: gesture-observed vs gesture-not observed) × 3 (Interview: Gesture-instructed vs gesture-allowed vs gesture-restricted) factorial ANOVAs was conducted. Children age and parental knowledge ratings were included as covariates[38]. There was no significant main effect of learning condition ($F(1, 76) = 0.04$, $p = .849$, $\eta_p^2 < .001$) or interview condition ($F(2, 76) = 1.81$, $p = .171$, $\eta_p2 = .05$) on children's verbal recall of names. There was also no significant interaction between the learning and interview conditions ($F(2, 76) = 1.63$, $p = .202$, $\eta_p^2 = .04$). See Table B7 for raw means and standard deviations.

Two-week delay. Children's recall of names was also analysed at the long delay using the same factorial ANOVA. Age and parental knowledge scores were again included as covariates[39]. The distribution of the recall of names at Time 2 was found to be

---

[38] Significant positive correlations were found between children's Time 1 percentage recall of names and both children's age ($r(82) = .30$, $p = .006$) and children's parental knowledge scores ($r(82) = .32$, $p = .003$).
[39] Significant positive correlations were also found between children's Time 2 percentage recall of names and both children's age ($r(82) = .32$, $p = .003$) and parental knowledge scores ($r(82) = .32$, $p = .003$).

significantly different from normal and could not be corrected (see Table B7) so raw scores were analysed. Prior to analyses, one outlying score was recoded. Findings indicated there were no significant main effects of either learning condition ($F(1, 76) = 1.03$, $p = .314$, $\eta_p^2 = .01$), or interview condition ($F(2, 76) = 1.54$, $p = .221$, $\eta_p^2 = .04$), on children's recall of names. The three way interaction between delay, learning condition, and interview condition was also non-significant ($F(2, 76) = 0.84$, $p = .436$, $\eta_p^2 = .02$), see Table B7.

Table B7

Means (Standard Deviations) for the Percentage Recall of Names at Time 1 and Time 2

		Time 1 Names		Time 2 Names	
Event gesture	Interview gesture	*M(SD)*	*D*	*M(SD)*	*D*
Observed	Instructed	54.37 (18.51)	.17	50.79 (16.14)	.22
	Allowed	59.92 (22.66)	.21	54.67 (16.19)	.32[*]
	Restricted	45.63 (18.25)	.21	43.25 (16.90)	.24[*]
	Total	53.31 (20.31)		49.57 (16.71)	
Not-observed	Instructed	48.81 (16.19)	.23[*]	50.79 (11.91)	.30[*]
	Allowed	56.35 (14.76)	.24[*]	53.97 (16.22)	.21
	Restricted	53.57 (16.54)	.19	50.79 (19.72)	.22
	Total	52.91 (15.77)		51.85 (15.94)	
Total	Instructed	51.59 (17.30)		50.79 (13.92)	
	Allowed	58.13 (18.85)		54.32 (15.91)	
	Restricted	49.60 (17.56)		47.02 (18.42)	
	Total	53.11 (18.08)		50.71 (16.27)	

[*] $p < .05$

Effects across the delay. Children's recall of names was also examined across the delay. A mixed ANOVA, with learning condition and interview condition as the between subjects variables and delay (recall of names at Time 1 versus Time 2) as the within subjects variable was conducted. There was no significant main effect of delay ($F(1, 78) = 3.56$, $p = .063$, $\eta_p^2 = .04$) and no significant interaction between learning condition and delay ($p > .05$) or between interview condition and delay ($p > .05$). There three-way interaction between delay, learning condition, and interview condition was also non-significant ($p > .05$), see Table B7.

S3.2: Kolmogorov-Smirnov tests of normality for Time 1 verbal recall and Time 1 gesture rates

Table B8

Kolmogorov-Smirnov Tests of Normality for Children's Verbal Recall and Gesture Rates at Time 1

Event gesture	Interview gesture	Total recall	Accuracy	Facts	Concepts	Rep Rate	ER rate
Observed	Instructed	.22	.18	.16	.14	.15	.13
	Allowed	.14	.14	.12	.25*	.18	.23*
	Restricted	.21	.18	.10	.22		
Not-observed	Instructed	.15	.17	.11	.17	.14	.18
	Allowed	.13	.13	.11	.12	.16	.17
	Restricted	.19	.21	.21	.27*		

* $p < .05$

S3.3: Coding and data analysis of event-relevant gestural variety

It was also of interest to examine whether children differed in the variety of event-relevant (ER) representational gestures they produced when recalling the lesson. Given that children who had observed gesture during learning were exposed to a greater variety of gestures, it was expected that children in the *gesture-observed* condition would produce a greater variety of event-relevant gestures compared to children in the *gesture-not observed* condition. Children in the *gesture-instructed* condition were also expected to produce a greater variety of gestures compared to the *gesture-allowed* conditions at the short delay. At the long delay, both conditions which had previously been able to gesture (instructed or allowed) were expected to produce a greater variety of gestures than children in the *gesture-restricted* condition. It was also hypothesised the increased gestural variety would be associated with improved verbal recall.

Coding. Of the gestures that coders agreed to be representational, Cohen's kappa was also calculated for the type of representational gesture produced. Coders decided whether the gesture was novel or which one of the 27 event-relevant gestures had been enacted. Cohen's Kappa was calculated for the representational gesture type, both at Time 1 ($k = .79$), and at Time 2 ($k = .74$).

181

Results. Preliminary analyses revealed significant correlations between the variety of different gestures children used during recall and the number of words spoken, both at Time 1 ($r(54) = .27$, $p = .043$) and at Time 2 ($r(82) = .52$, $p < .001$). To control for children's verbosity, gesture variety rates were calculated (i.e. (gesture count/word count)*100). Children's gesture variety was normally distributed at Time 1, but was corrected using a log transformation at Time 2. No significant correlations were found between the event-relevant variety rate and children's age, knowledge scores, or interest scores at either the short or the long delay ($p > .05$).

Gesture production variety at the 1-day delay. To examine the influence of both the learning and interview conditions on children's gestural variety, a 2 (Learning: gesture-observed vs gesture-not observed) $\times$ 2 (Interview: Gesture-instructed vs gesture-allowed) factorial ANOVA was conducted. The variety of event-relevant gestures the children produced per 100 words did not differ across the learning conditions ($F(1, 52) = 1.38$, $p = .246$, $\eta_p^2 = .03$) or interview conditions ($F(1, 52) = 0.55$, $p = .460$, $\eta_p^2 = .01$) and there was no significant interaction ($F(1, 52) = 1.58$, $p = .214$, $\eta_p^2 = .03$), see Table B9. Correlations between children's event-relevant gesture variety rate and verbal recall measures were also non-significant (i.e. total recall, accuracy, facts, or concepts; all p's $> .05$).\

Table B9

Means (Standard Deviations) and Kolmogorov-Smirnov Tests of Normality for Children's Event-Relevant (ER) Gesture Variety Rate at Time 1 and Time 2

		Time 1 ER variety rate		Time 2 ER variety rate	
Event gesture	Interview gesture	*M* (*SD*)	*D*	*M* (*SD*)	*D*
Observed	Instructed	1.19 (0.46)	.15	1.04 (0.65)	.12
	Allowed	1.28 (0.78)	.14	1.39 (0.83)	.22
	Restricted			0.89 (0.58)	.11
Not-observed	Instructed	1.21 (0.60)	.12	0.93 (0.69)	.22
	Allowed	0.86 (0.73)	.20	1.00 (1.04)	.21
	Restricted			0.65 (0.54)	.23[*]

Gesture production variety at the two-week delay. To examine the influence of the learning and interview experimental conditions on children's gesture rate at the long delay, a 2 (Learning: gesture-observed vs gesture-not observed) $\times$ 3 (Interview: Gesture-

instructed, gesture-allowed vs gesture-restricted) factorial ANOVA was conducted. There was no significant main effect of learning condition ($F(1, 78) = 3.17$, $p = .079$, $\eta_p^2 = .04$), or interview condition ($F(2, 78) = 1.64$, $p = .201$, $\eta_p^2 = .040$) on children's event-relevant gestural variety two weeks later. There was also no significant interaction between the learning and interview conditions ($F(2, 78) = 0.42$, $p = .657$, $\eta_p^2 = .011$), see Table B9. Correlations also revealed that children's event-relevant gesture variety rate during the Time 2 interview did not significantly correlate with any of the verbal recall measures within the same interview (p's > .05).

Conclusions. The experimental conditions had no significant effect on children's event-relevant gesture variety rate, and equally children's gestural variety also had no significant effect on children's verbal recall. Controlling for children's verbosity may have influenced the non-significant effect. Since children were not producing a wide variety of gestures to begin with, converting the gesture variety count to a gesture rate may have reduced the sensitivity with which differences could be detected across the conditions. Regardless, it appears that what was important for recall was not the variety of gesture produced, but rate of gesture production. When children gesture more often during recall they may be cuing with correct information (Frick-Horbury, 2002) and supporting the ease with which the content could be verbalized (Alibali et al., 2000; Wagner et al., 2004).

S3.4: Raw gesture count across the interview conditions

As well as examining the gesture rate across the three interview conditions, children's raw gesture count was also analysed, after recoding outliers. Independent samples t-tests revealed that children in the *gesture-instructed* condition ($M = 48.75$, $SD = 44.28$) produced significantly more gestures than the *gesture-allowed* condition ($M = 28.04$, $SD = 25.61$), $t(54) = -2.14$, $p = .037$, $d = .57$. The *gesture-instructed* condition also performed more event-relevant gestures ($M = 31.64$, $SD = 21.74$), compared to the *gesture-allowed* condition ($M = 20.43$, $SD = 19.40$), $t(54) = -2.04$, $p = .047$, $d = .54$. While there were indications that the gesture manipulation was successful, the effects appear to be subtle, as no significant differences were found across the conditions when children's gesture rates were measured.

S3.5: Interactions between learning condition and interview condition on children' gesture production rate at the short delay

Interactions between the learning conditions and interview conditions were also examined at the short delay in terms of children's gesture production rates. It was hypothesized that children in the *gesture-instructed* condition would produce gesture at a greater rate than children in the *gesture-allowed* condition, with children who had observed gesture at learning producing more gestures, in particular event-relevant gestures. Two 2 (Learning condition: Gesture observed vs Gesture) × 2 (Interview condition: Gesture instructed, gesture allowed) factorial ANOVAs were conducted with representational and event-relevant gesture rates as the dependent variables. There were, however, no interactions between the learning and interview conditions for either the rate of representational gestures produced ($F(1, 52) = 0.02$, $p = .881$, $\eta_p^2 < .001$), or the rate of event-relevant gestures produced $F(1, 52) < 0.01$, $p = .982$, $\eta_p^2 < .001$), see Table 8.4 (Chapter 8).

S3.6: Effects at the two-week delay

Both children's total verbal recall and accuracy were analysed at the two-week delay. Two 2 (learning condition: gesture-observed vs gesture-not observed) × 3 (interview condition: gesture-instructed, -allowed or -restricted) factorial ANOVAs were conducted. For the analysis of children's total recall, knowledge and age were both included as covariates, while only knowledge was included as a covariate for the accuracy analysis, see Table B10.

Table B10

Correlations Between Age, Parental Ratings of Child's Interest and Parental Ratings of Child's Knowledge and Children's Verbal Recall and Gesture Rate at Time 2

	Interest	Knowledge	Total recall	Accuracy	Facts	Concepts	Rep rate	ER rate
Age	.01	.13	.24*	.38**	.32**	.21	-.09	-.09
Interest		.60**	.28*	.15	.28*	.25*	.15	.15
Knowledge			.41**	.17	.44**	.32**	.08	.09

$* p < .05$ $** p < .01$

The same factorial ANOVAs were also conducted when separately examining children's percentage recall of the facts and concepts taught. When examining children's recall of facts, both children's knowledge and age were included as covariates. For the recall of concepts, however, only knowledge was included as a covariate, see Table B10. Prior to analyses, the distribution of concepts was corrected using a square root transformation, see Table B11.

Learning condition effects at the long delay. The hypothesis that children in the *gesture-observed* and *gesture-not observed* conditions would not differ in the amount or accuracy of their verbal report was assess by examining the main effect of learning condition. There were no significant differences across the *gesture-observed* and *gesture-not observed* learning conditions for either total recall ($F(1, 76) = 0.13$, $p = .717$, $\eta_p^2 = .002$), or the accuracy of children's verbal recall ($F(1, 77) = 0.38$, $p = .540$, $\eta_p^2 = .01$).

Children's recall of facts and concepts were also examined separately. Again, there were no significant main effects of learning condition, either for the recall of facts ($F(1, 76) = 0.22$, $p = .643$, $\eta_p^2 = .003$), or concepts ($F(1, 77) = 0.06$, $p = .812$, $\eta_p^2 = .001$). See Table 8.3 (Chapter 8) for raw means and standard deviations.

Interview condition effects at the long delay. At the long delay, it was hypothesised that children in the *gesture-instructed* condition would continue to gesture the most, followed by those in the *gesture-allowed* condition. Even though the children in the *gesture-restricted* condition were free to move their hands, they were expected to continue to gesture the least. Contrary to hypotheses, there was no main effect of interview condition on children's production of representational gestures ($F(2, 78) = 1.51$, $p = .228$, $\eta_p^2 = .04$) or event-relevant gestures ($F(2, 78) = 2.04$, $p = .137$, $\eta_p^2 = .05$), see Table 8.4 (Chapter 8).

In terms of the amount and accuracy of children's verbal recall, it was hypothesised that children in the *gesture-instructed* condition would recall more information, more accurately than children in the *gesture-allowed* condition, and that children in the *gesture-restricted* condition would show the poorest recall. There were, however, no significant main effects of interview condition on children's total recall ($F(2, 76) = 1.53$, $p = .224$, $\eta_p^2 = .04$), or accuracy ($F(2, 77) = 0.37$, $p = .694$, $\eta_p^2 = .01$). Similarly, when children's recall of facts and concepts were separately examined, there were no significant main effects of interview condition on the recall of facts ($F(2, 76) = 1.14$, $p = .326$, $\eta_p^2 = .03$), or concepts ($F(2, 77) = 0.80$, $p = .454$, $\eta_p^2 = .02$), see Table 8.3 (Chapter 8).

Interactions between learning and interview conditions at the long delay.

Finally, it was hypothesised children in the *gesture-observed* condition would recall more information, more accurately, than children in the *gesture-not observed* condition. Within the respective learning conditions, those who performed gesture (i.e. *instructed* or *allowed*) were hypothesised to recall more information, more accurately than children in the *gesture-restricted* condition. There were, however, no significant interactions for either children's total recall $F(2, 76) = 0.60$, $p = .550$, $\eta_p^2 = .02$, or recall accuracy ($F(2, 77) = 2.77$, $p = .069$, $\eta_p^2 = .07$). Equally, when children's recall of facts and concepts were separately examined, again there were no significant interactions (Facts: $F(2, 76) = 0.78$, $p = .464$, $\eta_p^2 = .02$; Concepts: $F(2, 77) = 0.33$, $p = .720$, $\eta_p^2 = .01$), see Table 8.3 (Chapter 8).

It was also hypothesised children's gestures production may differ across the learning and interview conditions. Two 2 (Learning condition: Gesture observed vs gesture not observed) × 3 (Interview condition: Gesture instructed, allowed, or restricted) factorial ANOVAs were conducted with gesture rates (representational or event-relevant) as the dependent measure. There were, however, no significant interactions for either children's rate of representational gesture production ($F(2, 78) = 0.39$, $p = .677$, $\eta_p^2 = .01$) or event-relevant gesture production ($F(2, 78) = 0.36$, $p = .699$, $\eta_p^2 = .01$), see Table 8.4 (Chapter 8).

Table B11

Kolmogorov-Smirnov Tests of Normality for Children's Verbal Recall and Gesture Rate at Time 2

Event gesture	Interview gesture	Total recall	Accuracy	Facts	Concepts	Rep rate	ER rate
Observed	Instructed	.14	.22	.12	.21	.16	.19
	Allowed	.12	.19	.21	.19	.15	.18
	Restricted	.22	.15	.17	.21	.11	.11
Not-observed	Instructed	.27[*]	.11	.17	.23[*]	.33[*]	.28[*]
	Allowed	.27[*]	.15	.23	.10	.27[*]	.24[*]
	Restricted	.27[*]	.13	.20	.16	.16	.23

* $p < .05$

Table B12

Correlations Between Children's Gesture Production During the Time 1 Interview and Children's Verbal Recall at Time 1 and Time 2

		Time 1					Time 2					
		ER rate	Total recall	Accuracy	Facts	Concept	Rep rate	ER rate	Total recall	Accuracy	Facts	Concept
Time 1	Rep rate	.94**	.29*	-.16	.27*	.25	.72**	.72**	.26	-.09	.29*	.16
	ER rate		.28*	-.16	.24	.23		.73**	.20	-.11	.25	.11
	Total recall			.32**	.89**	.59**			.71**	.26*	.74**	.48**
	Accuracy				.41**	.21				.78**	.38**	.30**
	Facts					.59**					.85**	.54**
	Concept											.61**
Time 2	Rep rate						.98**	.23*	-.15	.24*	.21	
	ER rate							.25*	-.14	.26*	.21	
	Total recall								.19	.92**	.71**	
	Accuracy									.33**	.17	
	Facts										.71*	

* $p < .05$ ** $p < .01$

S3.7 (continued)

It was also examined whether children's gesture production (regardless of interview condition) was associated with verbal recall. This was examined both within the same Time 1 interview, and across interviews. It was found that children's Time 1 event-relevant gesture rate was significantly, positively correlated with children's total recall within the same interview. Children's Time 1 representational gesture rate was also significantly and positively correlated with children's total recall, recall of facts and there was a trend that it was also associated with recall of concepts within the interview ($p = .059$), see Table B12.

It was also examined whether children's gesture production during the Time 1 interview was associated with children's verbal recall two weeks later. While there were no significant associations between children's Time 1 event-relevant gesture rate and children's verbal recall two weeks later, there was a significant, positive relationship between the rate of representational gesture produced and the recall of facts two weeks later, see Table B12.

Study 4

S4.1: Gesture and children's reporting of specific types of spatial language.

As well as examining the reporting of spatial information as a whole, children's separate use of prepositions and quantifiers, as well as motion, space, and time words were also examined. Data were analysed across the three studies and children who were asked *wh*-questions in Study 1 were excluded from analyses. The distributions of children's representational (Rep) and event-relevant (ER) gesture production rates were assessed via Kolmogorov-Smirnov tests of normality. Distributions which were not normal were corrected with transforms.

Observed gesture. A series of independent samples *t*-tests were conducted to examine whether the *gesture-observed* and *gesture-not observed* conditions differed in their percentage use of prepositions and quantifiers (verbs), or the percentage of their recall which included motion, space, or time descriptions.

Use of prepositions and quantifiers. At the short delay (1 day later), there were no significant differences in the use of propositions or quantifiers across the *gesture-observed* and *gesture-not observed* learning conditions, in any of the studies (all *p's* > .05). There were also no significant differences in the reporting of prepositions or quantifiers across

the learning conditions two weeks (Study 3) or 7 months (Study 2) after learning (all p's > .05). See Table B13 for means and standard deviations.

Descriptions of motion, space, and time. At the short delay (1 day later), children's descriptions of motion, space, and time were varied across the studies. The recall of spatial terms significantly differed between the learning conditions across all studies. In Study 1, children who observed gesture reported a smaller percentage of spatial terms than children who did not, $t(37) = 2.04$, $p = .049$, $d = .65$. Comparatively, in both Study 2 ($t(40) = -2.17$, $p = .036$, $d = .67$) and Study 3 ($t(82) = -2.28$, $p = .025$, $d = .50$), children recalled a greater percentage of spatial terms if they had observed gesture at learning. In terms of motion and time terms, there were no significant differences in children's percentage recall across the learning conditions in either Study 1 or Study 2 (all p's > .05). Study 3, however, found that children who observed gesture recalled a significantly greater percentage of motion terms ($t(82) = -2.57$, $p = .012$, $d = .56$) and time words ($t(82) = -2.46$, $p = .016$, $d = .54$), compared to children who did not observe gesture (see Table B13).

Two weeks later, children who observed gesture recalled significantly more motion words ($t(82) = -3.00$, $p = .004$, $d = .66$) and time words ($t(82) = -2.05$, $p = .044$, $d = .45$) than children who did not observe gesture, but there were no significant differences in the reporting of space words ($p > .05$). Seven months later, there were no significant differences across the learning conditions in the reporting of motion, space, or time words (all p's > .05)

Table B13

Means (Standard Deviations) for the Percentage of Spatial Information Reported

	Study 1 (1 day)		Study 2 (1 day)		Study 2 (7 months)		Study 3 (1 day)		Study 3 (2 weeks)	
	Gesture	No Gesture	Gesture	No Gesture	Gesture	No Gesture	Gesture	No Gesture	Gesture	No Gesture
Prepositions	0.65	0.89	6.83	6.57	0.76	0.61	1.03	0.74	6.57	6.06
	(0.45)	(0.67)	(2.84)	(2.14)	(0.52)	(0.71)	(0.84)	(0.69)	(1.76)	(2.28)
Quantifiers	11.22	11.41	2.00	2.42	10.85	9.69	10.30	10.05	2.58	3.27
	(2.02)	(2.08)	(0.85)	(1.21)	(2.51)	(1.88)	(2.00)	(2.12)	(1.40)	(2.08)
Motion	0.01	0.01	13.04	11.61	0.26	0.13	0.13	0.03	13.51	11.28
	(0.04)	(0.02)	(2.48)	(3.12)	(0.32)	(0.22)	$(0.22)_a$	$(0.12)_b$	$(3.24)_a$	$(3.56)_b$
Space	14.46	15.97	2.05	1.44	9.94	10.45	13.28	11.93	2.61	2.23
	$(2.38)_a$	$(2.25)_b$	$(1.02)_a$	$(0.78)_b$	(2.80)	(3.40)	$(2.41)_a$	$(2.95)_b$	(1.38)	(2.12)
Time	3.09	3.23	8.18	7.36	1.47	1.72	2.39	1.92	8.26	7.31
	(1.13)	(1.17)	(2.02)	(2.36)	(0.83)	(1.30)	$(0.83)_a$	$(0.92)_b$	$(2.17)_a$	$(2.09)_b$

Note. Different subscripts represent significant differences between the groups

Performed gesture. Correlations were also conducted to examine the relationship between both children's representational (Rep) and event-relevant (ER) gesture rates and children's reporting of the various spatial words. Correlations were examined within the same interview. See Table B14 for an overview of the significant correlations found.

Use of prepositions and quantifiers. At the short delay (1 day later) in Study 1, children's gesture production rates were not significantly associated with the recall of propositions or quantifiers (p's > .05). In Study 2, however, there was a significant positive correlation between children's representational gesture rate and children's reporting of prepositions, $r(40) = .31$, $p = .047$. No other relationships were significant within Study 2 (all p's > .05). In Study 3, there were significant positive correlations between children's reporting of quantifiers and both children's rate of representational gesture ($r(54) = .27$, $p = .044$) and event-relevant gesture ($r(54) = .26$, $p = .049$). The relationships between children's gesture rates and reporting of prepositions were non-significant (p's > .05).

Correlations were also examined at the longer delays. Two weeks after learning (Study 3), there was a significant positive relationship between children's reporting of prepositions and both their rate of representational gesture production ($r(82) = .22$, $p = .045$) and event-relevant gesture production ($r(82) = .26$, $p = .017$). There were no significant correlations between gesture production and quantifiers two weeks later (p's > .05). Seven months later, all correlations between children's gesture rates and reporting of prepositions and quantifiers were non-significant (p's > .05).

Descriptions of motion, space, and time. At the short delay, significant relationships between gesture production and verbal recall were only found in Study 3, all Study 1 and Study 2 correlations were non-significant (p's > .05). In Study 3, there was a significant association between children's reporting of space terms and both children's representational gesture rate ($r(54) = .29$, $p = .032$) and children event-relevant gesture rate ($r(54) = .28$, $p = .035$). There were no significant correlations between time or motion terms and children's gesture production ($p > .05$).

Two weeks later after the learning experience (Study 3), there was a significant positive correlation between children's event-relevant gesture rate and the percentage recall of motion terms ($r(82) = .23$, $p = .037$), but there were no significant associations between children's representational gesture rate and motion terms, or between any gesture rate and children's percentage recall of space or time terms (p's > .05). Seven months later, children who were producing more representational gestures were also recalling a greater percentage of space terms ($r(28) = .53$, $p = .003$) and time terms ($r(28) = .55$, $p = .002$).

191

Similarly, the event-relevant gesture rate was also significantly, positively associated with recalling space ($r(28) = .53$, $p = .002$) and time words ($r(28) = .46$, $p = .011$).

Table B14

Associations Between Children's Gesture Production and the Spatial Words Spoken During the Same Interview

	Study 1		Study 2 (Time 1)		Study 2 (Time 2)		Study 3 (Time 1)		Study 3 (Time 2)	
	Rep	ER	Rep	ER	Rep	ER	Rep	ER	Rep	ER
Preposition	-	-	↑	-	-	-	-	-	↑	↑
Quantifier	-	-	-	-	-	-	↑	↑	-	-
Motion	-	-	-	-	-	-	-	-	-	↑
Space	-	-	-	-	↑	↑	↑	↑	-	-
Time	-	-	-	-	↑	↑	-	-	-	-

Note. ↑ represents a significant positive correlation, ↓ represents a significant negative correlation.

S4.2: Recall of spatial terms in Study 3 at the long delay

A one-way ANOVA also examined whether children's recall of spatial terms at the long delay varied across the interview conditions (i.e. gesture-instructed, gesture-allowed, or gesture-restricted). During this delayed interview, children were free to move their hands as they wished. No significant main effect of interview condition was found, $F(2, 81) = 0.72$, $p = .492$, $\eta_p^2 = .02$.

Meta-analyses

S5.1: Observed gesture and verbal recall (Study 1 *wh*-question conditions included)

Children's verbal recall across the *gesture-observed* and *gesture-not observed* conditions was also examined when children who answered *wh*-questions in Study 1 were included in analyses. Some of the variables significantly deviating from normal; these variables were corrected using square root transforms or, if they could not be corrected, were excluded from analyses.

First, the amount and accuracy of children's recall was examined across the learning conditions. The *gesture-observed* and *gesture-not observed* learning conditions did not differ in their total recall ($d_{\text{unbiased}} = 0.07$ [-0.15, 0.28]) or recall accuracy ($d_{\text{unbiased}} =$

0.03 [-0.22, 0.29]). Similarly, whether or not the children observed gesture during learning had no significant effect on children's separate percentage recall of either facts ($d_{unbiased} =$ 0.14 [-0.08, 0.35]), or concepts ($d_{unbiased} = 0.22$ [-0.01, 0.45]).

S5.2: Percentage recall of facts and concepts across the learning conditions, including both long and short delay data (figures only).
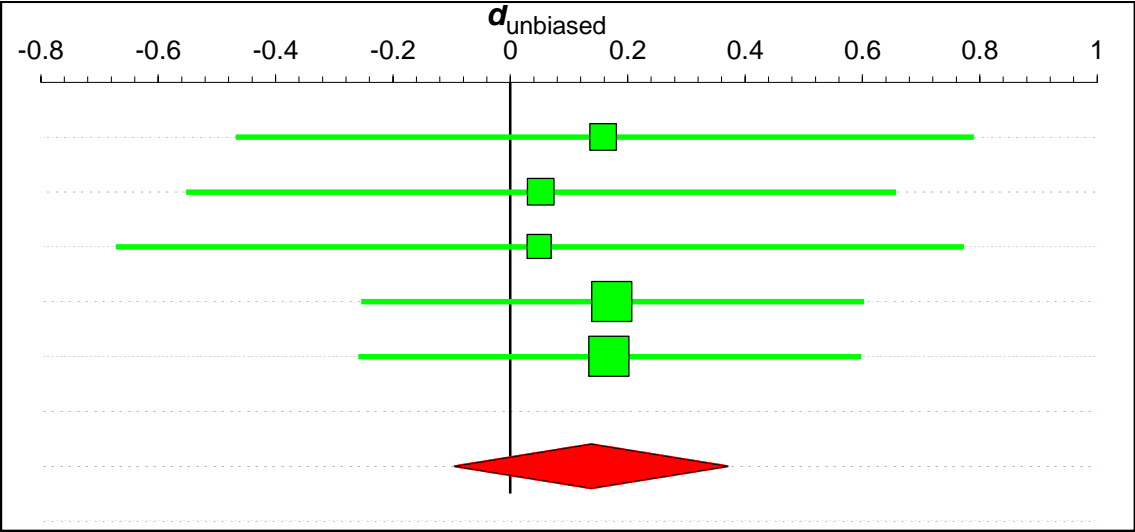
Figure B1. Recall of **facts** across the gesture-observed and gesture-not observed learning conditions (Study 1, Study 2-Time 1, Study 2-Time 2, Study 3-Time 1, Study 3-Time 2).
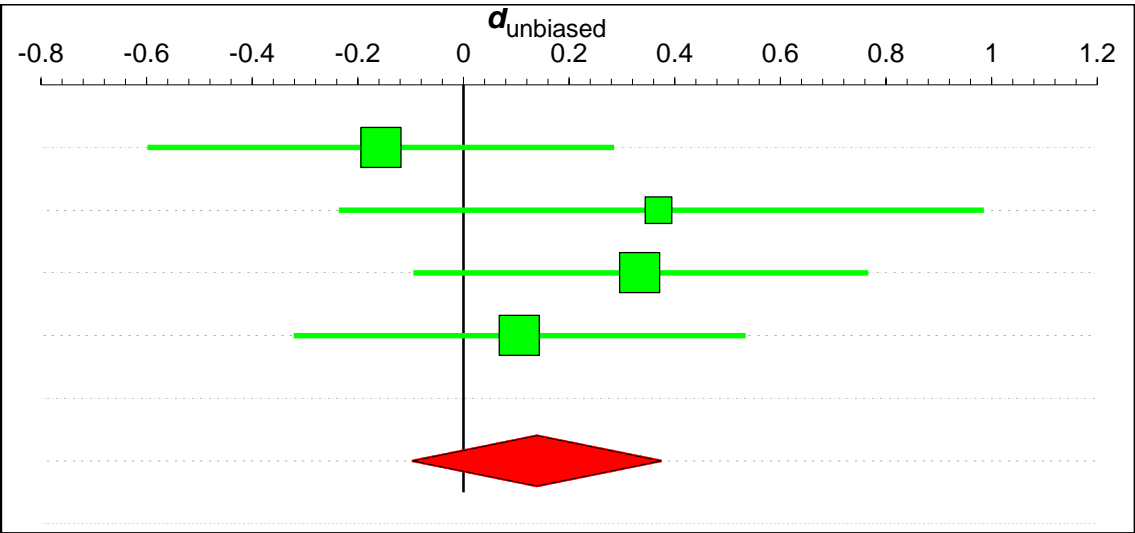
Figure B2. Recall of **concepts** across the gesture-observed and gesture-not observed learning conditions (Study 1, Study 2-Time 1, Study 3-Time 1, Study 3-Time 2).

S5.3: Verbal recall across the learning conditions, examined separately at the short and long delays.

At the short delay. Children's verbal recall across the learning conditions was examined purely at the short (1-day delay) delay. Analyses were conducted with *wh-*questions excluded. The *gesture-observed* and *gesture-not observed* learning conditions did not significantly differ in their total recall ($d_{unbiased}$ = -0.03 [-0.33, 0.27]), see Figure B3, or recall accuracy ($d_{unbiased}$ = 0.11 [-0.35, 0.57]), see Figure B4. When children' recall of facts and concepts were examined separately, neither the recall of facts ($d_{unbiased}$ = 0.14 [-0.16, 0.44]), nor the recall of concepts ($d_{unbiased}$ = 0.16 [-0.18, 0.51]) significantly differed across the learning conditions, see Figures B5 and B6.
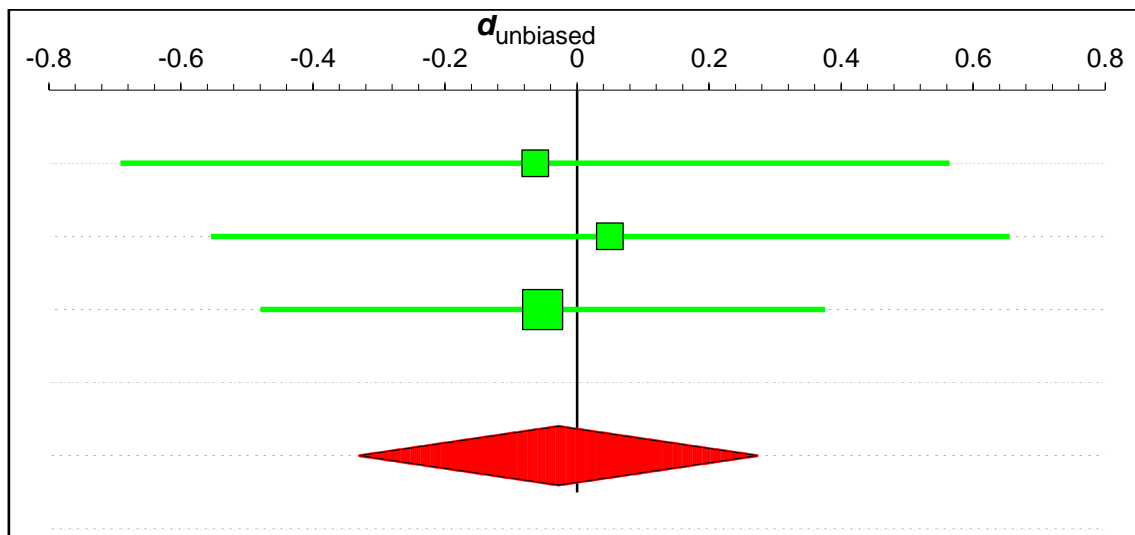
Figure B3. **Total recall** across the gesture-observed and gesture-not observed conditions at the 1-day delay only (Study 1, Study 2-Time 1, Study 3-Time 1).
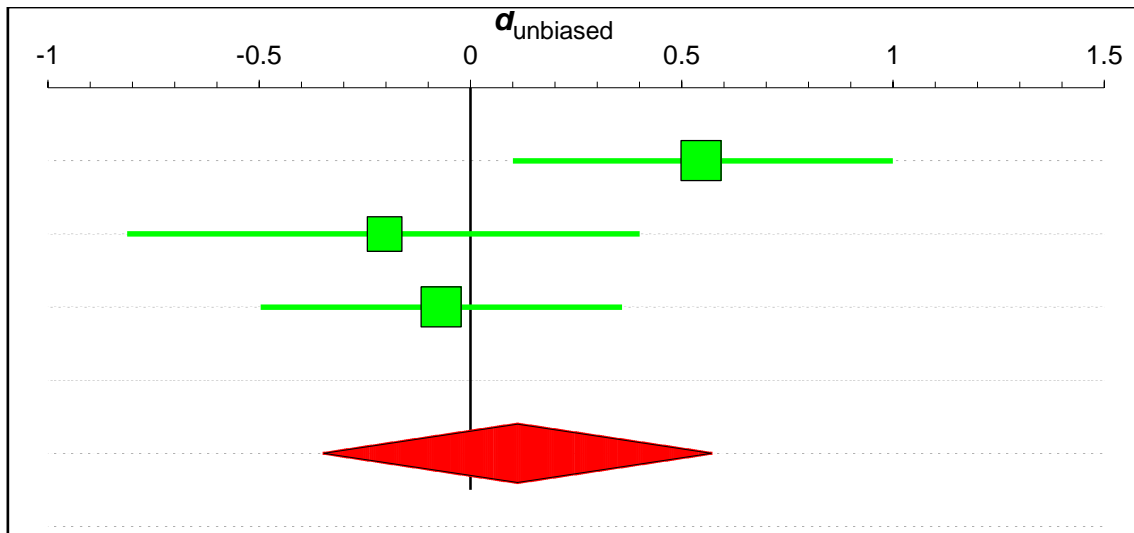
Figure B4. **Accuracy** of children's recall across the gesture-observed and gesture-not observed conditions at the 1-day delay only (Study 1, Study 2-Time 1, Study 3-Time 1)
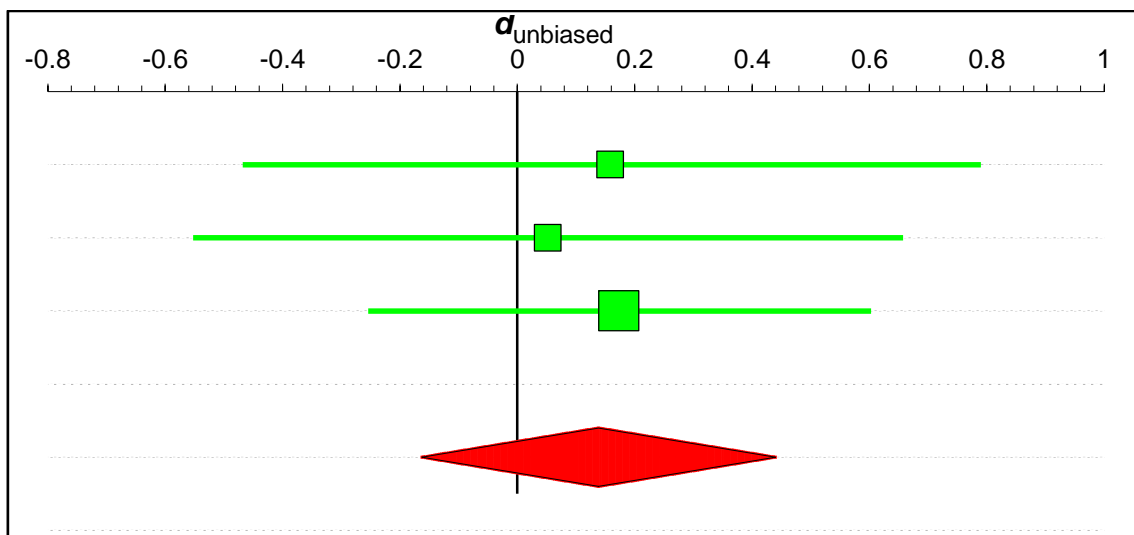
Figure B5. Recall of **facts** across the gesture-observed and gesture-not observed conditions at the 1-day delay only (Study 1, Study 2-Time 1, Study 3-Time 1)
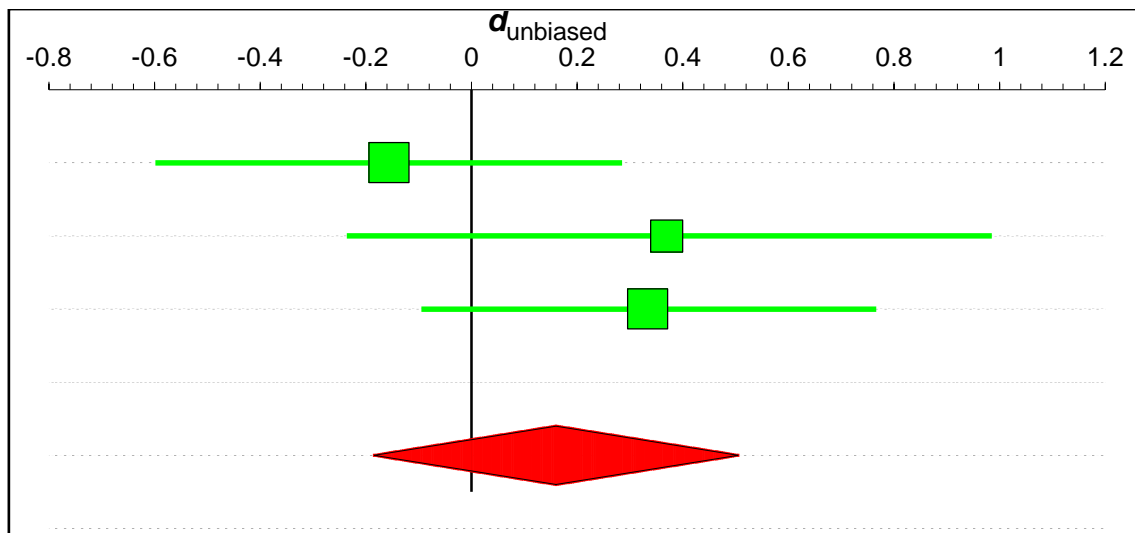
Figure B6. Recall of **concepts** across the gesture-observed and gesture-not observed conditions at the 1-day delay only (Study 1, Study 2-Time 1, Study 3-Time 1)

At the long delay. Children's verbal recall across the learning conditions was also separately examined at the long delays (2 weeks and 7 months after learning). Analyses were again conduced with *wh*-questions excluded. The amount and accuracy of children's recall, as well as children's recall of facts were examined. Due to problems with the normality of the distribution for the percentage recall of concepts, this measure was not analyzed.

Children's total recall did not significantly differ across the *gesture-observed* and *gesture-not observed* learning conditions ($d_{unbiased} = 0.18$ [-0.18, 0.55]), and neither did recall accuracy ($d_{unbiased} = -0.09$ [-0.45, 0.28]), or children's recall of facts ($d_{unbiased} = 0.14$ [-0.23, 0.50]). See Figures B7, B8, and B9.
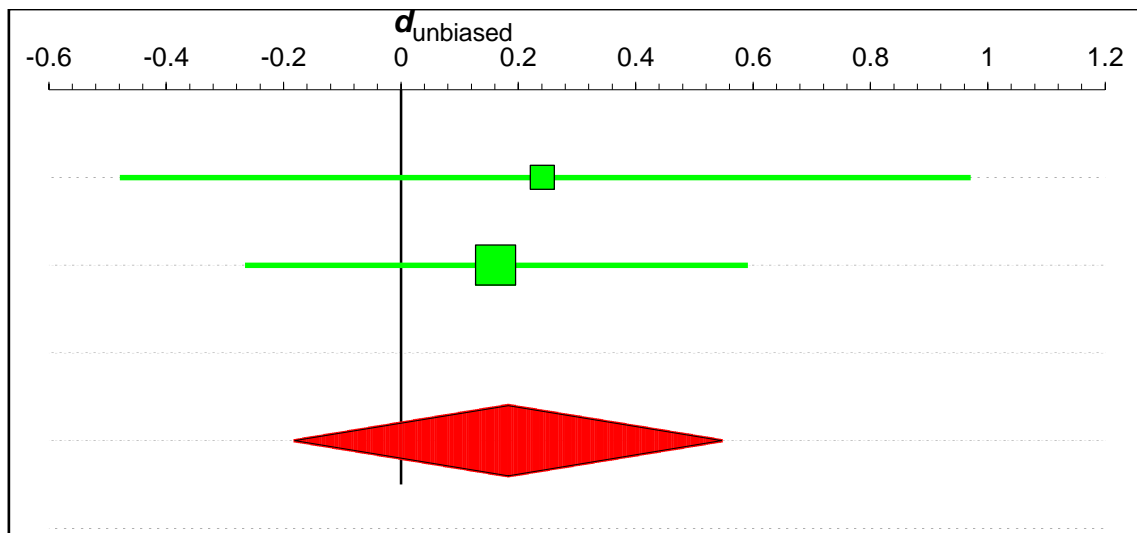
Figure B7. **Total recall** across the gesture-observed and gesture-not observed conditions at the long delay only (Study 2-Time 2, Study 3-Time 2).
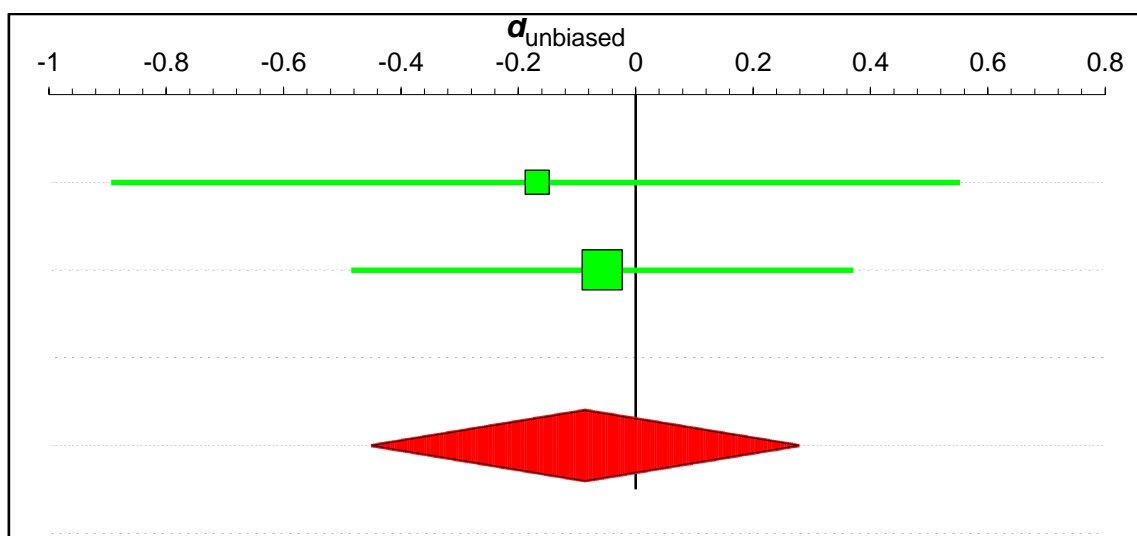
Figure B8. **Accuracy** of children's recall across the gesture-observed and gesture-not observed conditions at the long delay only (Study 2-Time 2, Study 3-Time 2).
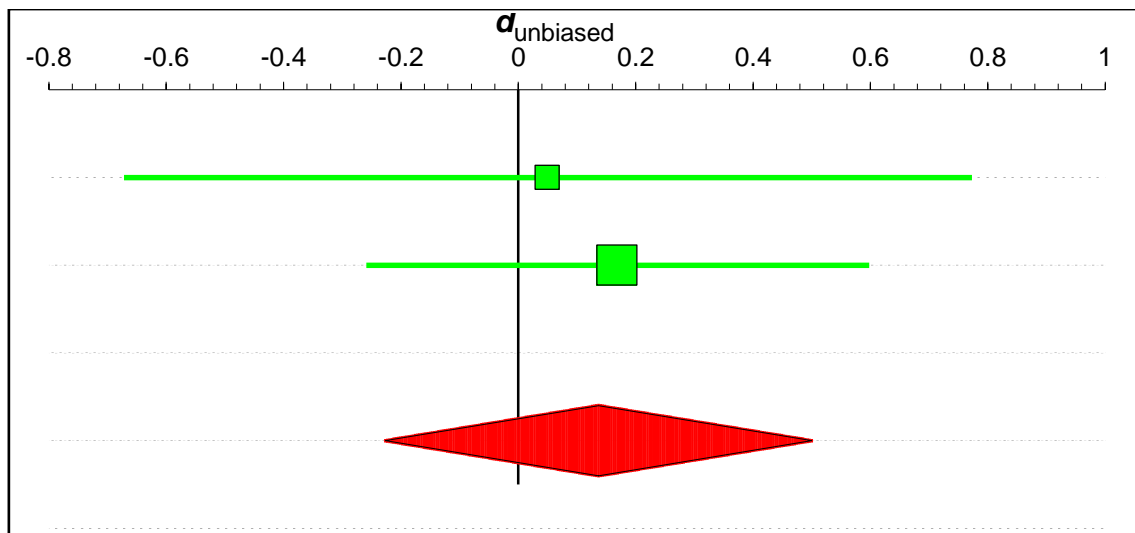
Figure B9. Recall of **facts** across the gesture-observed and gesture-not observed conditions at the long delay only (Study 2-Time 2, Study 3-Time 2).

S5.4: Children's Time 1 gesture production rates and Time 2 recall of facts (figure only)
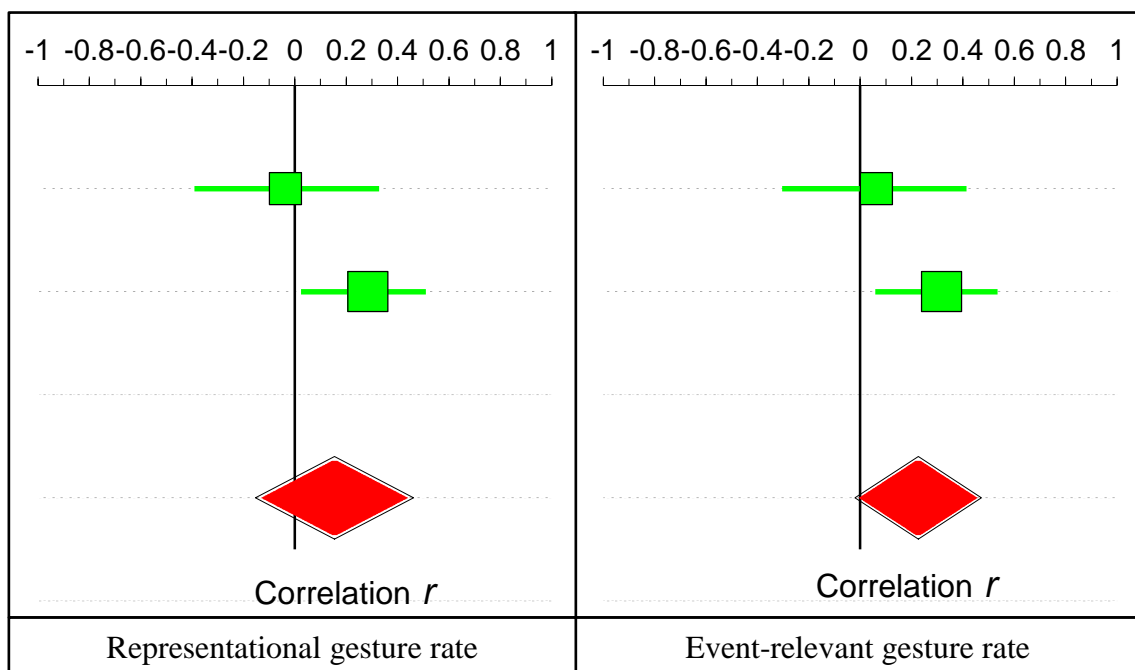
Figure B10. Meta-analysis assessing the relationship between children's Time 1 representational and event-relevant gesture production rates and children's Time 2 recall of **facts**.

General discussion

S6.1: Correlations between errors and word count across the studies

Table B15

Correlations Between Children's Raw Error Counts and the Number of Words Spoken During the Same Interview

	Study 1	Study 2- Time 1	Study 2- Time 2	Study 3- Time 1	Study 3- Time 2
Person Correlation	.70**	.44**	.76**	.88**	.93**

Appendix C

Solar system script

Solar system script used in study 3. The other scripts were similar if with slight alterations.

Introduction

Hi _____ my name is Lynley. Today we are going to be learning about space and the solar system. Are you ok to learn about the solar system with me today? Now you don't have to do this, and if you want to stop at any stage you can, ok? Just let me know. So when we are talking today I'd like you to look at me and I'd like you to do your best listening ok? So look at me and please don't talk unless I ask you a question. That sound ok?

Today we are going to talk about two of the most important things in our solar system and these are; the sun [gesture sun in the middle] and the planets that are around it [point to different positions around sun].

In the middle of the solar system [gesture circular ring] is the sun [point to the middle of the ring], which is a giant ball [gesture bigger ball with both hands] of gas that sends out heat and light [gesture things moving outwards from the sun]. Then, circling around the sun are the 8 planets [gesture orbit with one finger moving around the fist]. The closest planet to the sun is Mercury [point with one finger close to the sun], then Venus, Earth, Mars, Jupiter, Saturn, Uranus and the furthest out is Neptune [gesture fist a little further from the sun for each planet, larger jumps for four planets further out].

The four planets [four fingers] closest to the sun [hold other hand close to central ball representing sun] are Mercury, Venus, Earth and Mars [gesture each of these a little further from the ball] and these 4 planets are small [gesture small with thumb and forefinger], and close together [gesture with palm of two hands coming together to stop with not much distance between].

The four planets [four fingers] furthest from the sun [hold hand far out from central ball representing the sun] are Jupiter, Saturn, Uranus and Neptune [gesture each of these a little further from the ball, but start at point further than where Mars was gestured as being] and these 4 planets are big [gesture hands far from each other] and far apart [gesture with two fingers from different hands moving further apart]. So there is lots of space between them.

So there are two groups [gesture one group with two hands and then move hands to the side to gesture other group] of planets in our solar system. There are the 4 planets closest to the sun [gesture one group like before] which are small [small with two fingers] and close together [palms pressing together], and the 4 planets further from the sun [gesture other group like before] which are bigger [gesture big sphere with both hands] and more spread out [palms far from each other].

What makes a planet?

Now let's talk about what makes something a planet. There are three rules [gesture 3 fingers] that have to be met for something to be counted as a planet. The first [hold up one finger] is that a planet has to have a round shape [gesture a round sphere with both hands]. The second [hold two fingers] is that it has to move around the sun in a circle [gesture orbit around fist, keep going until end of sentence] and this circle the planet makes as it moves around the sun is called an orbit. The third rule [hold up three fingers] that has to be met is that it must have a clear pathway [gesture two hands moving away from each other to signify clear] as it orbits around the sun [gesture orbit].

So some things that seem like planets, like Pluto for example, aren't counted as planets because they don't met this third rule [wiggle third finger] and these planets are called dwarf planets instead. So while dwarf planets are round [gesture touching first finger] and orbit around the sun [gesture touching second finger], they don't have this clear orbit [gesture moving things away to signify clearing] and often bump into things [gesture palm hitting into things and going around in a circle] as they move around the sun. So, because dwarf planets have stuff in their orbits [gesture fists and a line moving through] they don't meet this third rule [hold up three fingers] and aren't counted as planets.

So planets are round [gesture sphere], orbit around the sun [gesture orbit] and have a clear pathway as they orbit [gesture clearing and then orbit]. Now because all the planets have a clear pathway[gesture two hands moving away from each other to signify clear], none of the orbits overlap [gesture two hands criss-crossing] and each planet's orbit is a little bit further from the sun [use fist for sun and indicate a curved hand moving in steps away from the sun]. This means that all the planets take a different amount of time to orbit the sun. Planets closer to the sun [point close to fist], like Mercury, take less time to orbit around the sun [gesture orbit close to fist] because they have shorter orbit lines and less distance to travel [gesture short – one finger from each hand close together, draw line between them as draw apart]. Planets that are far from the sun though [point far from fist],

like Neptune, take much longer to orbit around the sun [gesture far orbit] because they have longer orbit lines and much further to travel [gesture long – one finger from each hand far apart, draw line between them as the fingers move apart].

So for something to be counted as a planet it has to meet 3 rules [three fingers]. It has to be round [gesture sphere], orbit around the sun [gesture orbit], and have a clear pathway as it orbits [gesture clearing motion and orbit]. Because each planet has its own orbit, close to the sun take less time to orbit [gesture close orbit] than planets that are far away [gesture hand far from fist].

Gravity

Now let's talk about why everything stays in the solar system. The reason everything stays in the solar system is because of gravity. Gravity is an invisible pull [gesture pulling with one fist towards the body], which keeps everything in the universe where it's supposed to be [gesture holding down motion with both hands]. Everything has gravity, but bigger things [gesture big with both hands] have more gravity and more pull [gesture pulling harder than before].

The sun is the biggest thing [gesture big circle] in all the solar system [gesture swiping motion] and because the sun is so big [gesture big] it has heaps of gravity to pull [gesture pulling motions in towards the body] on all the planets and stops them from floating away into space [gesture planets floating away].
But, the planets don't get sucked into the sun [gesture both hands sucked into body] because the planets are moving [gesture motion in a line outwards with hand]. The planets move fast enough [gesture fast orbit] that the sun's gravity isn't strong enough [gesture difficulty pulling] to pull the planets all the way in to make them crash into the sun [gesture both hands crashing into body].

Let's use an example. Imagine you had a tennis ball [gesture ball with fist] and you tied it to a long piece of string [indicate a piece of string coming off the ball] and then took then end of the string and used it to spin the ball around in a circle [use one hand to make a spinning motion]. The ball is moving [indicate swishing motion past the fist], but the piece of string [indicate a line coming from fist] holds the ball in place [indicate ball at end of the string] and makes it move around your hand in a circle [indicate ball moving around fist in a circle]. In this example, the ball is like the moving planet [move planet hand up and down] and your hand [gesture line coming off other fist] and the string are like the

sun's gravity. The gravity pulls the moving planet [gesture pulling motion into sun fist] in and makes it orbit around the sun in a circle [gesture orbit].

If you let go of the string while you were spinning the ball around [Gesture swinging the ball around and then opening fist], the ball would go shooting off [gesture shooting off]. This is what would happen if the sun had no gravity to pull the planets in [gesture pulling motions in towards the body]. All the planets would go shoot off into space [gesture each planet shooting off].

So as I said before, bigger things have more gravity and more pull. In our solar system [gesture smaller ring with both hands], the sun is the biggest thing [point to middle of ring], so it has heaps of gravity to pull [gesture pulling] the planets in and makes them orbit around it [gesture orbit]. In our galaxy though [gesture larger ring with both hands], there are lots of suns and solar systems [open fist in different areas to gesture the different suns and solar systems] and in the middle of the galaxy [gesture point in the middle of galaxy] is a humongous black hole [gesture circle/hole with both hands] which keeps everything in the right place [gesture holding suns and solar systems]. This black hole [gesture circle/hole with both hands] is much bigger than the sun's [gesture big], so it has lots of gravity [gesture strong pulling] to pull in all the sun's and make them orbit around the black hole [gesture orbit]. Our sun takes 250 million years to orbit [gesture orbit slowly].

So gravity holds everything in the solar system where it should be. The sun's gravity [gesture sphere for sun] pulls on the planets [gesture pulling] and makes them orbit around the sun [gesture orbit around fist] and the gravity from the black hole [gesture large circle/hole with both hands] pulls on the sun [gesture pulling] and makes it orbit around the black hole [gesture orbit].

Surface of planets

Now let's talk about the types of planets. So in our solar system there are 8 [hold up four fingers on both hands] planets that meet the three rules to be counted as a planet. These 8 planets are divided [split apart the four fingers on each hand] into two types: the rocky planets [move four fingers on one hand] and the gas planets [move four fingers on other hand]. The four planets closest to the sun are rocky planets [point out four planets close to fist], which have a <u>solid</u> ground made of dirt and rock [form a ball with both hands with pressure in fingertips to indicate hard]. The four planets [point out four planets far

from fist] furthest from the sun are gas planets, and have a ground made of <u>gas</u> [gesture ball with fingers moving slightly].

Let's talk about the solid, rocky planets first. Solids are things which have a shape that stays the same [gesture a ball that is hard]. For example, a rock is a solid, so all the particles or little bits that make up the rock [gesture small bits within the sphere that was made], are packed tightly together [push palms of two hands together] so the shape stays the same.

The rocky planets, the four planets closest to the sun [point out four planets close to fist], have a solid surface [motion both hands creating a sphere] made of dirt and rock. The particles, or little bits that make up the rock and dirt, are packed tightly together [gesture two palms pressing together] so they keep their shape [gesture flat oval with both hands] and can support your weight when you stand on them [gesture one hand pushing on the back of the other hand]. So if you visited a rocky planet you would be able to stand on the ground [gesture fingers standing on back of hand and ground resisting], like you can here on earth [point to ground].

Now let's talk about the gas planets. Gases, on the other hand, have no shape [make a ball sort of shape and then move hands all over the place]. For example, the air we breathe is a gas (breathe out) and all the particles or bits that make up a gas are spread far apart [[gesture small bits with both hands where the air was blown all spread out].

The gas planets, the four planets furthest from the sun [point out four planets far from fist], have a surface made of gas [hands surface of a ball]. The particles, or bits that make up the gas, are spread far apart [gesture little particles more spread out than last time] and can't support your weight [gesture movement down between the spread molecules]. So if you were able to visit a gas planet there would be no ground to stand on and you would sink through [gesture fingers walking and falling through back of hand], just like if you tried to stand on air.

But the gas planets aren't gas all the way through because further into the centre of the planet everything gets pressed and squeezed by pressure [gesture sphere moving inwards and squeezing]. The gas particles [gesture spread out molecules], get squeezed closer together start to turn to liquid [gesture both hand collecting the molecules inwards]. So inside of the gas planets is liquid [gesture middle of sphere], and scientists also think there might be something pretty cool in there [keep hands in same place and beat]. Scientists think there might be huge [gesture large with both hands] diamonds and some diamonds might be as big as skyscrapers!

So there are rocky planets and gas planets. Rocky planets are close to the sun [four fingers close to fist] and have solid surfaces [gesture hard sphere] you can stand on. Gas planets are further from the sun [four fingers far from fist] and have surfaces made of gas [gesture movable sphere] which you would fall through [finger move through sphere].
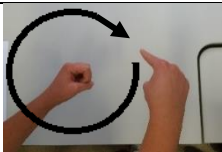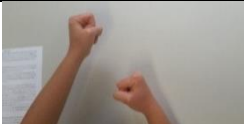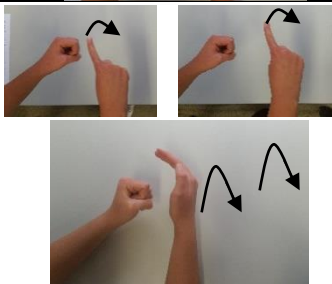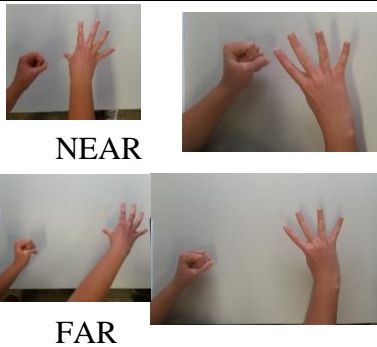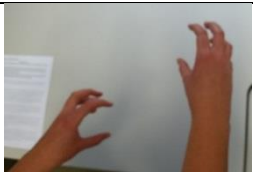
Summary

Ok, so today we have talked about what makes something a planet, the reason why everything stays in the solar system and the different surfaces of the planets. Thank you for listening so well and participating with me. We are all done now, I hope you had fun.

Now in order to help me with my study and make the results really good, it is very important that you don't talk to the other kids about what you learnt with me today until the end of the study. Once the study is finished you can talk about today all you like, but until then please don't tell the others what you learnt, as we might get weird findings and it will also make this less exciting for the other children. So please don't tell the other children what you learnt until the end of the study. Thank you!
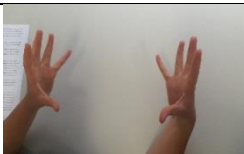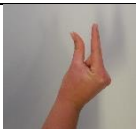
Appendix D
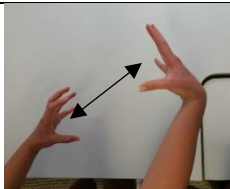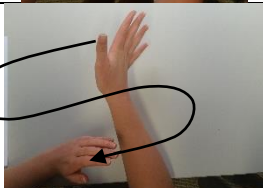
Scripted gestures
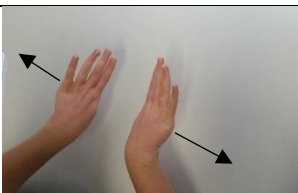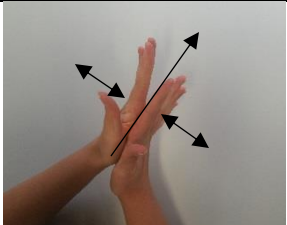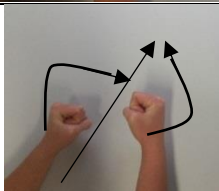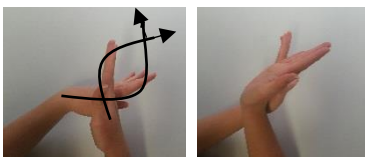
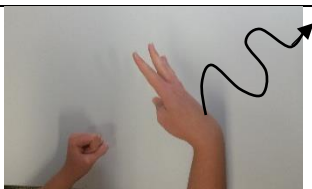| Commonly used gestures		
Picture	**Gesture description**	**Verbal content**
	Single finger moves around left fist OR Fist moves around other fist	Orbit refers to planets around sun, sun around black hole or ball around hand
	Beginning of an orbit, but no motion in the hands	Orientation of the planets.
	Finger points to successive locations in a line further from the left fist. Pointing starts further from fist when talking about the far planets	Refers to the locations of the planets and each orbit being further from the sun
NEAR FAR	The hand or body act as a placeholder for the location of the sun. Another hand (either held open or with four fingers up) indicates the near and far placement relative to the sun.	There are planets close to the sun and planets that are far from the sun.
	Hands are held in a circular shape to indicate a group. This motion is produced either side of the body (not directly centred in front of the gesturer.	Planets are in two groups: those close to the sun and those far from the sun.
	Right hands anchors a central object. Both hands moves to various locations around the indicated centre.	Planets around sun and the suns/solar systems are positioned around the black hole.
	Hands are held	Things stay in the right place

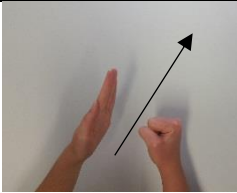|	Hands are held in a circular shape to indicate a group or circular objects and one hand points to the centre of the shape.	Middle of solar system or galaxy Middle of gas planets is liquid
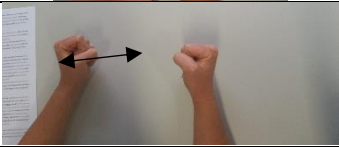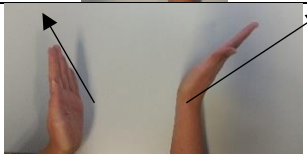
Counting gestures considered as informative		
Picture	**Gestural description**	**Verbal content**
	Hands held	Three rules to make a planet
	Hands held	Refers to the number of planets in each group (rocky and gas)
	Hands held	Eight planets in total, four in each group
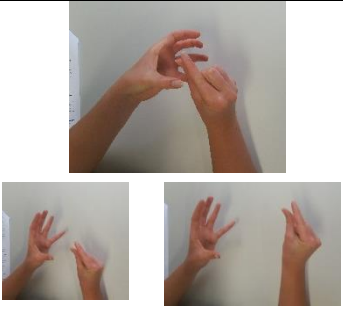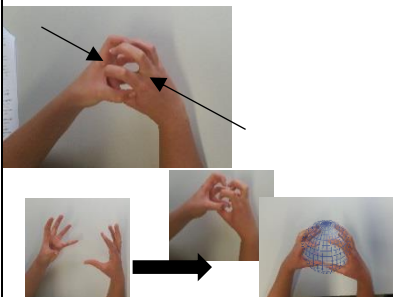
Introduction to the solar system		
Picture	**Gestural description**	**Verbal content**
	Hands held in a circular shape.	Representing the sun, planets, black hole or solar system/ galaxy as circular objects. Also used to refer to the planets as being round and solids having a defined shape.
	Hands are held far apart with fingers up.	Refers to the planets being big, or far apart. The child may draw a line between the two hands to indicate the long distance planets have to travel to get around the sun.
	Thumb and forefinger held close together	Planets are small
	Hands are held close together with fingers up.	Refers to the planets being close together. The child may draw a line between the hands to indicate the short distance the planets travel.

	Hands form a circular shape to anchor the position of the sun and then each hand moves outwards from this position.	Sun sends out heat and light.
	Hand makes a swiping motions back and forward in front of the body.	Referring to all of the solar system.

What makes something a planet		
Picture	**Gestural description**	**Verbal content**
	Both hands move outwards with palms facing outwards	For something to be counted as a planet it needs to have a clear pathway
	Palms hit into each other while moving forwards.	Dwarf planets bump into things
	Fists move consecutively in front of each other and then the orbit line is drawn through the middle.	Dwarf planets have stuff in their orbits
	Hands criss-cross over each other.	None of the orbits overlap

Gravity and the solar system		
Picture	**Gestural description**	**Verbal content**
	Fist make a pulling motion towards the body. May be either each hand consecutively or both hands at once.	Gravity is a pull. Gravity pulls on the planets and pulls them in towards the sun. The black hole pulls on the suns.
	One hand anchors the location of the sun and the other makes a flicking motion moving away.	Planets don't float into space

	Either the hand or body are used to anchor the location of the sun. One hand moves quickly away from the central point.	Planets are moving. Without the sun, plants would shoot off into space. If you let go of the ball after swinging it around it will shoot off.
	Hands move in towards the body and then fingers spread out once contact has been made with the body.	Planets don't crash into the sun.
	One fist is held. The other hand pulls away from the stationary fist in a straight line.	Tennis ball tied to string
	Fist creates a small circular rotation.	Spin the ball around
	One fist remains stationary while the other moves back and forth slightly towards and backwards from stationary fist.	Sun pulls on the moving planet
	Hand is held in fist and fingers open outwards.	Let go of the string when you spin the ball around. This is like the sun having no gravity.
	Hands move outwards in straight lines going in different directions	Without gravity the planets and suns would shoot off

Types of planets in the solar system		
Picture	**Gestural description**	**Verbal content**
	A circular shape is created to represent the planet and inside it the thumb and forefinger move to various locations, either close together or far apart.	Referring to the particles that make up the planet. The particles are either close together (rocky planets) or far apart (gas planets).
	Hands create a circular shape and hands move inwards toward each other. Fingers create squeezing motion.	Representing the squeezing together of the particles in the solid rocky planets. Also represents how the gas particles come together in the gas planets to create liquid.

	Hands stay in a circular shape, but move so that the fingers and wrists move closer then further, or the fingers rotate from left to right. When referring to the gas planets this motion is created with the hands moving outwards from each other	Refers to the surface of the planets as round. May refer to the lack of surface on the gas planets as well.
	One had presses on top of the other hand, fingers may also walk on top of hand.	Rock planets can support your weight so you can stand on them.
	Finger points down towards the ground	Rock planets have a ground like here on earth.
	Fingers attempt to press on the back of the hand, but fall through the fingers. Two fingers may also be walking on hand and fall through fingers.	Gas planets can't support your weight so you will sink/fall through the surface.
BLOW ONTO HAND	Breath is blown onto hand	The air we breathe is a gas

211

References

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal, 31*, 313-337. Retrieved from http://www.jstor.org/stable/1163312

Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation, 5*, 307-331. doi:10.1207/s15427633scc0504_2

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hand reveal about a child's state of mind. *Cognitive Psychology, 25*, 468-523. doi:10.1006/cogp.1993.1012

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language, 44*, 169-188. doi:10.1006/jmla.2000.2752

Alibali, M. W., & Kita, S. (2010). Gesture highlights perceptually present information for speakers. *Gesture, 10*, 3-28. doi:10.1075/gest.10.1.02ali

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes, 15*, 593-613. doi:10.1080/016909600750040571

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*, 247-286. doi:10.1080/10508406.2011.611446

Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Jacobs, S. A., Johnson Martinez, C., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction, 32*, 65-100. doi:10.1080/07370008.2013.858161

Allen, R., & Shatz, M. (1983). 'What says meow?'. The role of context and linguistic experience in very young children's responses to what-questions. *Journal of Child Language, 10*, 321-335. doi:10.1017/S0305000900007790

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1063-1087. doi:10.1037/0278-7393.20.5.1063

Andre, T., Whigham, M., Hendrickson, A., & Chambers, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. *Journal of Research in Science Teaching, 36*, 719-747. doi:10.1002/(SICI)1098-2736(199908)36:6<719::AID-TEA8>3.0.CO;2-R

Austin, E. E., & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of Experimental Child Psychology, 122*, 92-103. doi:10.1016/j.jecp.2013.12.008

Baddeley, A. (2003a). Working memory and language: An overview. *Journal of communication disorders, 36*(3), 189-208. doi:10.1016/S0021-9924(03)00019-4

Baddeley, A. (2003b). Working memory: Looking back and looking forward. *Nature reviews neuroscience, 4*, 829-839. doi:10.1038/nrn1201

Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (1990). The effects of involvement on children's memory for events. *Cognitive Development, 5*, 55-69. doi:10.1016/0885-2014(90)90012-I

Baker-Ward, L., Ornstein, P. A., & Holden, D. J. (1984). The expression of memorization in early childhood. *Journal of Experimental Child Psychology, 37*, 555-575. doi:10.1016/0022-0965(84)90076-6

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental psychobiology, 40*, 293-310. doi:10.1002/dev.10034

Beattie, G., & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica, 123*(1-2), 1-30. doi:10.1515/semi.1999.123.1-2.1

Beaudoin-Ryan, L., & Goldin-Meadow, S. (2014). Teaching moral reasoning through gesture. *Developmental Science, 17*, 984-990. doi:10.1111/desc.12180

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science, 21*, 1605-1610. doi:10.1177/0956797610385353

Benjamin, N., Haden, C. A., & Wilkerson, E. (2010). Enhancing building, conversation, and learning through caregiver-child interactions in a children's museum. *Developmental Psychology, 46*, 502-515. doi:10.1037/a0017822

Berk, L. E. (2009). *Child development*. Boston, Mass. London: Pearson/Allyn and Bacon.

Bjorklund, D. F., Miller, P. H., Coyle, T. R., & Slawinski, J. L. (1997). Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training studies. *Developmental Review, 17*, 411-441. doi:10.1006/drev.1997.0440

214

Bjorklund, D. F., Ornstein, P. A., & Haig, J. R. (1977). Developmental differences in

    organization and recall: Training in the use of organizational techniques.

    Developmental Psychology, 13, 175-183. doi:10.1037/0012-1649.13.3.175

Boland, A. M., Haden, C. A., & Ornstein, P. A. (2003). Boosting children's memory by

    training mothers in the use of an elaborative conversational style as an event

    unfolds. *Journal of Cognition and Development, 4*, 39-65.

    doi:10.1207/S15327647JCD4,1-02

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children

    gesture brings out implicit knowledge and leads to learning. *Journal of

    Experimental Psychology: General, 136*, 539-550. doi:10.1037/0096-

    3445.136.4.539

Brooks, N., & Goldin-Meadow, S. (2016). Moving to learn: How guiding the hands can set

    the stage for learning. *Cognitive Science, 40*, 1831–1849. doi:10.1111/cogs.12292

Brown, A. L., & Campione, J. C. (1972). Recognition memory for perceptually similar

    pictures in preschool children. *Journal of Experimental Psychology, 95*, 55-62.

    doi:10.1037/h0033276

Butler, L. P., & Markman, E. M. (2012). Finding the cause: Verbal framing helps children

    extract causal evidence embedded in a complex scene. *Journal of Cognition and

    Development, 13*, 38-66. doi:10.1080/15248372.2011.567201

Cameron, H., & Xu, X. (2011). Representational gesture, pointing gesture, and memory

    recall of preschool children. *Journal of Nonverbal Behavior, 35*, 155-171.

    doi:10.1007/s10919-010-0101-2

Capirci, O., Montanari, S., & Volterra, V. (1998). Gestures, signs, and words in early

    language development. *New Directions for Child and Adolescent Development,

    1998*(79), 45-60. doi:10.1002/cd.23219987904

Chi, M. T. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.),

    Children's thinking: What develops (pp. 73-96). Hillsdale, New Jersey, USA:

    Lawrence Erlbaum Associates, Inc.

Chin, C. (2006). Classroom interaction in science: Teacher questioning and feedback to

    students' responses. *International Journal of Science Education, 28*, 1315-1346.

    doi:10.1080/09500690600621100

Chu, M., & Hagoort, P. (2014). Synchronization of speech and gesture: Evidence for

    interaction in action. *Journal of Experimental Psychology: General, 143*, 1726-

    1741. doi:10.1037/a0036281

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General, 140*, 102-116. doi:10.1037/a0021790

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology: General, 143*, 694-709. doi:10.1037/a0033861

Chui, K. (2014). Mimicked gestures and the joint construction of meaning in conversation. *Journal of Pragmatics, 70*, 68-85. doi:http://dx.doi.org/10.1016/j.pragma.2014.06.005

Church, R. B. (1999). Using gesture and speech to capture transitions in learning. *Cognitive Development, 14*, 313-342. doi:10.1016/S0885-2014(99)00007-6

Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism, 7*, 303-319. doi:10.1080/13670050408667815

Church, R. B., Garber, P., & Rogalski, K. (2007). The role of gesture in memory and social communication. *Gesture, 7*, 137-158. doi:10.1075/gest.7.2.02bre

Church, R. B., Kelly, S. D., & Lynch, K. (2000). Immediate memory for mismatched speech and representational gesture across development. *Journal of Nonverbal Behavior, 24*, 151-174. doi:10.1023/A:1006610013873

Cleveland, E. S., & Reese, E. (2005). Maternal structure and autonomy support in conversations about the past: Contributions to children's autobiographical memory. *Developmental Psychology, 41*, 376-388. doi:10.1037/0012-1649.41.2.376

Coffman, J. L., Ornstein, P. A., McCall, L. E., & Curran, P. J. (2008). Linking teachers' memory-relevant language and the development of children's memory skills. *Developmental Psychology, 44*, 1640-1654. doi:10.1037/a0013859

Colletta, J.-M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication, 52*, 565-576. doi:10.1016/j.specom.2010.02.009

Cook, S. W. (2011). Abstract thinking in space and time: Using gesture to learn math. *Cognitie, Creier, Comportament / Cognition, Brain, Behavior, 15*, 553-570. Retrieved from https://search.proquest.com/docview/916330919?accountid=14782

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development, 84*, 1863–1871. doi:10.1111/cdev.12097

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7*, 211-232. doi:10.1207/s15327647jcd0702_4

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106*, 1047-1058. doi:10.1016/j.cognition.2007.04.010

Cook, S. W., & Tanenhaus, M. K. (2009). Embodied communication: Speakers' gestures affect listeners' actions. *Cognition, 113*, 98-104. doi:10.1016/j.cognition.2009.06.006

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal of Memory and Language, 63*, 465-475. doi:10.1016/j.jml.2010.07.002

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes, 27*, 594-610. doi:10.1080/01690965.2011.567074

Cooperrider, K., Wakefield, E., & Goldin-Meadow, S. (2015). More than meets the eye: Gesture changes thought, even without visual feedback. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 441-446. Retrieved from https://mindmodeling.org/cogsci2015/papers/0085/paper0085.pdf

Corballis, M. C. (2002). *From hand to mouth: The origins of language*. New Jersey, USA: Princeton University Press.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior, 11*, 671-684. doi:10.1016/S0022-5371(72)80001-X

Cumming, G., & Calin-Jageman, R. (2016). *Introduction to the new statistics: Estimation, open science, and beyond*. New York, USA: Routledge.

Cutica, I., & Bucciarelli, M. (2008). The deep versus the shallow: Effects of co-speech gestures in learning from discourse. *Cognitive Science, 32*, 921–935. doi:10.1080/03640210802222039

Cutica, I., & Bucciarelli, M. (2011). "The more you gesture, the less I gesture": Co-speech gestures as a measure of mental model quality. *Journal of Nonverbal Behavior, 35*, 173-187. doi:10.1007/s10919-011-0112-7

Cutica, I., & Bucciarelli, M. (2013). Cognitive change in learning from text: Gesturing enhances the construction of the text mental model. *Journal of Cognitive Psychology, 25*, 201-209. doi:10.1080/20445911.2012.743987

de Nooijer, J. A., van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica, 144*, 173-179. doi:10.1016/j.actpsy.2013.05.013

De Ruiter, J. P. (2007). Postcards from the mind: The relationship between speech, imagistic gesture, and thought. *Gesture, 7*, 21-38. doi:10.1075/gest.7.1.03rui

De Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science, 4*, 232-248. doi:10.1111/j.1756-8765.2012.01183.x

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology, 36*, 511-523. doi:10.1037/0012-1649.36.4.511

Deloache, J. S., & Marzolf, D. P. (1995). The use of dolls to interview young children: Issues of symbolic representation. *Journal of Experimental Child Psychology, 60*, 155-173. doi:10.1006/jecp.1995.1036

Dick, A. S., Goldin-Meadow, S., Solodkin, A., & Small, S. L. (2012). Gesture in the developing brain. *Developmental Science, 15*, 165-180. doi:10.1111/j.1467-7687.2011.01100.x

Dreyfus, A., Jungwirth, E., & Eliovitch, R. (1990). Applying the "cognitive conflict" strategy for conceptual change—some implications, difficulties, and problems. *Science Education, 74*, 555-569. doi:10.1002/sce.3730740506

Driskell, J. E., & Radtke, P. H. (2003). The effect of gesture on speech production and comprehension. *Human Factors, 45*, 445-454. doi:10.1518%2Fhfes.45.3.445.27258

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. United States of America: National Academies Press.

Emmorey, K., & Casey, S. (2002). Gesture, thought, and spatial language. In K. R. Coventry & P. Olivier (Eds.), *Spatial Language: Cognitive and Computational Perspectives* (pp. 87-101). Dordrecht: Springer Netherlands.

Eshach, H., Dor-Ziderman, Y., & Yefroimsky, Y. (2014). Question asking in the science classroom: Teacher attitudes and practices. *Journal of Science Education and Technology, 23*, 67-81. doi:10.1007/s10956-013-9451-y

Farrant, K., & Reese, E. (2000). Maternal style and children's participation in reminiscing: Stepping stones in children's autobiographical memory development. *Journal of Cognition and Development, 1*, 193-225. doi:10.1207/S15327647JCD010203

Feyereisen, P. (2006). Further investigation on the mnemonic effect of gestures: Their meaning matters. *European Journal of Cognitive Psychology, 18*, 185-205. doi:10.1080/09541440540000158

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.

Fivush, R., Gray, J. T., & Fromhoff, F. A. (1987). Two-year-old talk about the past. *Cognitive Development, 2*, 393-409. doi:10.1016/S0885-2014(87)80015-1

Fivush, R., Haden, C. A., & Reese, E. (2006). Elaborating on elaborations: Role of maternal reminiscing style in cognitive and socioemotional development. *Child Development, 77*, 1568-1588. doi:10.1111/j.1467-8624.2006.00960.x

Frick-Horbury, D. (2002). The use of hand gestures as self-generated cues for recall of verbally associated targets. *The American Journal of Psychology, 115*, 1-20. doi:10.2307/1423671

Fyfe, E. R., DeCaro, M. S., & Rittle-Johnson, B. (2015). When feedback is cognitively-demanding: The importance of working memory capacity. *Instructional Science, 43*, 73-91. doi:10.1007/s11251-014-9323-8

Fyfe, E. R., & Rittle-Johnson, B. (2015). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology, 108*, 82-97. doi:10.1037/edu0000053

Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology, 104*, 1094-1108. doi:10.1037/a0028389

Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science, 26*, 817-831. doi:10.1016/S0364-0213(02)00087-3

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse processes, 28*, 1-26. doi:10.1080/01638539909545067

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. United States of America: Harvard University Press.

Goldin-Meadow, S. (2005). The two faces of gesture: Language and thought. *Gesture, 5*, 241-257. doi:10.1075/gest.5.1.16gol

Goldin-Meadow, S. (2015). Gesture and cognitive development. *Handbook of Child Psychology and Developmental Science, 2*(9), 1-42. doi:10.1002/9781118963418.childpsy209

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100*, 279-297. doi:10.1037/0033-295X.100.2.279

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20*, 267-272. doi:10.1111/j.1467-9280.2009.02297.x

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology, 91*, 720-730. doi:10.1037/0022-0663.91.4.720

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12*, 516-522. doi:10.1111%2F1467-9280.00395

Goldin-Meadow, S., Shield, A., Lenzen, D., Herzig, M., & Padden, C. (2012). The gestures ASL signers use tell us when they are ready to learn math. *Cognition, 123*, 448-453. doi:10.1016/j.cognition.2012.02.006

Goldin-Meadow, S., & Singer, M. A. (2003). From children's hands to adults' ears: Gesture's role in the learning process. *Developmental Psychology, 39*, 509-520. doi:10.1037/0012-1649.39.3.509

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive Sciences, 9*, 234-241. doi:10.1016/j.tics.2005.03.006

Goldin-Meadow, S., Levine, S. C., Zinchenko, E., Yip, T. K., Hemani, N., & Factor, L. (2012). Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science, 15*, 876-884. doi:10.1111/j.1467-7687.2012.01185.x

Göncü, A., & Gauvain, M. (2012). Sociocultural approaches to educational psychology: Theory, research, and application. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues.* (pp. 123-152).

Washington, D.C., US: American Psychological Association. doi:10.1037/13273-006

Goossens, N. A., Camp, G., Verkoeijen, P. P., Tabbers, H. K., & Zwaan, R. A. (2014). The benefit of retrieval practice over elaborative restudy in primary school vocabulary learning. *Journal of Applied Research in Memory and Cognition, 3*, 177-182. doi:10.1016/j.jarmac.2014.05.003

Graham, T. A. (1999). The role of gesture in children's learning to count. *Journal of Experimental Child Psychology, 74*, 333-355. doi:10.1006/jecp.1999.2520

Grimminger, A., Rohlfing, K. J., & Stenneken, P. (2010). Children's lexical skills and task demands affect gestural behavior in mothers of late-talking children and children with typical language development. *Gesture, 10*, 251-278. doi:10.1075/gest.10.2-3.07gri

Güler, O. E., Larkina, M., Kleinknecht, E., & Bauer, P. J. (2010). Memory strategies and retrieval success in preschool children: Relations to maternal behavior over time. *Journal of Cognition and Development, 11*, 159-184. doi:10.1080/15248371003699910

Haden, C. A., Haine, R. A., & Fivush, R. (1997). Developing narrative structure in parent–child reminiscing across the preschool years. *Developmental Psychology, 33*, 295-307. doi:10.1037/0012-1649.33.2.295

Haden, C. A., Ornstein, P. A., Eckerman, C. O., & Didow, S. M. (2001). Mother–child conversational interactions as events unfold: Linkages to subsequent remembering. *Child Development, 72*, 1016-1031. doi:10.1111/1467-8624.00332

Haden, C. A., Ornstein, P. A., Rudek, D. J., & Cameron, D. (2009). Reminiscing in the early years: Patterns of maternal elaborativeness and children's remembering. *International Journal of Behavioral Development, 33*, 118-130. doi:10.1177/0165025408098038

Hatano, G., Miyake, Y., & Binks, M. G. (1977). Performance of expert abacus operators. *Cognition, 5*, 47-55. doi:10.1016/0010-0277(77)90016-6

Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition, 15*, 95-110. doi:10.1016/0010-0277(83)90035-5

Hayne, H. (2004). Infant memory development: Implications for childhood amnesia. *Developmental Review, 24*, 33-73. doi:10.1016/j.dr.2003.09.007

Hayne, H., Boniface, J., & Barr, R. (2000). The development of declarative memory in human infants: Age-related changes in deffered imitation. *Behavioral neuroscience, 114*, 77-83. doi:10.1037/0735-7044.114.1.77

Hedrick, A. M., Haden, C. A., & Ornstein, P. A. (2009). Elaborative talk during and after an event: Conversational style influences children's memory reports. *Journal of Cognition and Development, 10*, 188-209. doi:10.1080/15248370903155841

Hedrick, A. M., San Souci, P., Haden, C. A., & Ornstein, P. A. (2009). Mother-child joint conversational exchanges during events: Linkages to children's memory reports over time. *Journal of Cognition and Development, 10*, 143-161. doi:10.1016/j.dr.2004.08.004

Hodapp, R. M., Goldfield, E. C., & Boyatzis, C. J. (1984). The use and effectiveness of maternal scaffolding in mother-infant games. *Child Development, 55*, 772-781. doi:10.2307/1130128

Hoetjes, M., Krahmer, E., & Swerts, M. (2014). Does our speech change when we cannot gesture? *Speech Communication, 57*, 257-267. doi:10.1016/j.specom.2013.06.007

Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology, 26*, 4-27. doi:10.1177/0261927X06296428

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological bulletin, 137*, 297-315. doi:10.1037/a0022128

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15*, 495-514. doi:10.3758/PBR.15.3.495

Hostetter, A. B., Alibali, M. W., & Kita, S. (2007a). Does sitting on your hands make you bite your tongue? The effects of gesture prohibition on speech during motor descriptions. *Proceedings of the 29th annual meeting of the Cognitive Science Society*, 1097-1102. Retrieved from http://escholarship.org/uc/item/6g02f7bv

Hostetter, A. B., Alibali, M. W., & Kita, S. (2007b). I see it in my hands' eye: Representational gestures reflect conceptual demands. *Language and Cognitive Processes, 22*, 313-336. doi:10.1080/01690960600632812

Hudson, J. A., & Shapiro, L. R. (1991). From knowing to telling: The development of children's scripts, stories, and personal narratives. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure* (pp. 89-136). Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers.

Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development, 9*, 23-43. doi:10.1016/0885-2014(94)90018-3

Iverson, J. M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication in Italian vs. American children. *First Language, 28*, 164-181. doi:10.1177/0142723707087736

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature, 396*, 228. doi:10.1038/24300

Iverson, J. M., & Goldin-Meadow, S. (2001). The resilience of gesture in talk: Gesture in blind speakers and listeners. *Developmental Science, 4*, 416-422. doi:10.1111/1467-7687.00183

Iverson, J. M., Longobardi, E., Spampinato, K., & Caselli, M. C. (2006). Gesture and speech in maternal input to children with Down's syndrome. *International Journal of Language & Communication Disorders, 41*, 235-251. doi:10.1080/13682820500312151

Jant, E. A., Haden, C. A., Uttal, D. H., & Babcock, E. (2014). Conversation and object manipulation influence children's learning in a museum. *Child Development, 85*, 2029-2045. doi:10.1111/cdev.12252

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin, 114*, 3-28. doi:10.1037/0033-2909.114.1.3

Jones, C. H., & Pipe, M.-E. (2002). How quickly do children forget events? A systematic study of children's event reports as a function of delay. *Applied Cognitive Psychology, 16*, 755-768. doi:10.1002/acp.826

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23-31. doi:10.1207/S15326985EP3801_4

Kang, S., Hallman, G., Son, L., & Black, J. (2013). The different benefits from different gestures in understanding a concept. *Journal of Science Education and Technology, 22*, 825-837. doi:10.1007/s10956-012-9433-5

Kang, S., Tversky, B., & Black, J. B. (2015). Coordinating gesture, word, and diagram: Explanations for experts and novices. *Spatial Cognition & Computation, 15*, 1-26. doi:10.1080/13875868.2014.958837

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331*, 772-775. doi:10.1126/science.1199327

Karpicke, J. D., Blunt, J. R., Smith, M. A., & Karpicke, S. S. (2014). Retrieval-based learning: The need for guided retrieval in elementary school children. *Journal of Applied Research in Memory and Cognition, 3*, 198-206. doi:10.1016/j.jarmac.2014.07.008

Kelly, S. D. (2001). Broadening the units of analysis in communication: Speech and nonverbal behaviours in pragmatic comprehension. *Journal of Child Language, 28*, 325-349. doi:10.1017/S0305000901004664

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language, 40*, 577-592. doi:10.1006/jmla.1999.2634

Kelly, S. D., & Church, R. B. (1997). Can children detect conceptual information conveyed through other children's nonverbal behavior's. *Cognition and Instruction, 15*, 107-134. doi:10.1207/s1532690xci1501_4

Kelly, S. D., & Church, R. B. (1998). A comparison between children's and adults' ability to detect conceptual information conveyed through representational gestures. *Child Development, 69*, 85-93. doi:10.2307/1132072

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and language, 89*, 253-260. doi:10.1016/S0093-934X(03)00335-3

Kendon, A. (2004). *Gesture: Visible action as utterance*. United Kingdom: Cambridge University Press.

King, A., & Rosenshine, B. (1993). Effects of guided cooperative questioning on children's knowledge construction. *Journal of Experimental Education, 61*, 127-148. doi:10.1080/00220973.1993.9943857

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162-185). United Kingdom: Cambridge University Press.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*, 16-32. doi:10.1016/S0749-596X(02)00505-3

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of

speech and gesture production. *Language and Cognitive Processes, 22*, 1212-1236. doi:10.1080/01690960701461426

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin, 119*, 254-284. doi:10.1037/0033-2909.119.2.254

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual review of psychology, 51*, 481-537. doi:10.1146/annurev.psych.51.1.481

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? *Advances in experimental social psychology, 28*, 389-450. doi:10.1016/S0065-2601(08)60241-5

Krauss, R. M., Chen, Y., & Gotfexnum, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-283). United Kingdom: Cambridge University Press.

Krauss, R. M., Dushay, R. A., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gesture. *Journal of Experimental Social Psychology, 31*, 533-552. doi:10.1006/jesp.1995.1024

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic books.

Leung, E. H., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology, 17*, 215-220. doi:10.1037/0012-1649.17.2.215

Lipko-Speed, A., Dunlosky, J., & Rawson, K. A. (2014). Does testing with feedback help grade-school children learn key concepts in science? *Journal of Applied Research in Memory and Cognition, 3*, 171-176. doi:10.1016/j.jarmac.2014.04.002

MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review, 18*, 1197-1202. doi:10.3758/s13423-011-0168-8

Macoun, A., & Sweller, N. (2016). Listening and watching: The effects of observing gesture on preschoolers' narrative comprehension. *Cognitive Development, 40*, 68-81. doi:10.1016/j.cogdev.2016.08.005

Marstaller, L., & Burianová, H. (2013). Individual differences in the gesture effect on working memory. *Psychonomic Bulletin & Review, 20*, 496-500. doi:10.3758/s13423-012-0365-0

Mayberry, R. I., Jaques, J., & DeDe, G. (1998). What stuttering reveals about the development of the gesture-speech relationship. *New Directions for Child Development, 1998*, 77-87. doi:10.1002/cd.23219987906

Mayberry, R. I., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Current Directions in Psychological Science, 9*, 192-196. doi:10.1111/1467-8721.00092

McGuigan, F., & Salmon, K. (2004). The time to talk: The influence of the timing of adult-child talk on children's event memory. *Child Development, 75*, 669-686. doi:10.1111/j.1467-8624.2004.00700.x

McGuigan, F., & Salmon, K. (2006). The influence of talking on showing and telling: Adult-child talk and children's verbal and nonverbal event recall. *Applied Cognitive Psychology, 20*, 365-381. doi:10.1002/acp.1183

McNeil, N. M., Alibali, M. W., & Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior, 24*, 131-150. doi:10.1023/A:1006657929803

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

McShane, J. (1991). *Cognitive development: An information processing approach*. Massachusetts, USA: Basil Blackwell.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes, 22*, 473-500. doi:10.1080/01690960600696916

Miller, P. H. (1990). The development of strategies of selective attention. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 157-184). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 615-622. doi:10.1037/0278-7393.18.3.615

Morris, G., & Baker-Ward, L. (2007). Fragile but real: Children's capacity to use newly acquired words to convey preverbal memories. *Child Development, 78*, 448-458. doi:10.1111/j.1467-8624.2007.01008.x

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *The American Journal of Psychology, 117*, 411-424. Retrieved from http://www.jstor.org/stable/4149008

Namy, L. L., Acredolo, L., & Goodwyn, S. (2000). Verbal labels and gestural routines in parental communication with young children. *Journal of Nonverbal Behavior, 24*, 63-79. doi:10.1023/A:1006601812056

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development, 69*, 295-308. doi:10.2307/1132165

Nelson, K. (1974). Variations in children's concepts by age and category. *Child Development, 45*, 577-584. doi:10.2307/1127823

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review, 111*, 486-511. doi:10.1037/0033-295X.111.2.486

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences, 9*, 467-473. doi:10.1016/j.tics.2005.08.004

Novack, M. A., Congdon, E., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction using the hands to learn math. *Psychological Science, 25*, 903-910. doi:10.1177/0956797613518351

Novack, M. A., & Goldin-Meadow, S. (2015). Learning from gesture: How our hands change our minds. *Educational Psychology Review, 27*, 405-412. doi:10.1007/s10648-015-9325-3

Novack, M. A., Goldin-Meadow, S., & Woodward, A. L. (2015). Learning from gesture: How early does it happen? *Cognition, 142*, 138-147. doi:10.1016/j.cognition.2015.05.018

Novack, M. A., Wakefield, E. M., & Goldin-Meadow, S. (2016). What makes a movement a gesture? *Cognition, 146*, 339-348. doi:10.1016/j.cognition.2015.10.014

Ornstein, P. A., Haden, C. A., & Hedrick, A. M. (2004). Learning to remember: Social-communicative exchanges and the development of children's memory skills. *Developmental Review, 24*, 374-395. doi:10.1016/j.dr.2004.08.004

Ornstein, P. A., & Trabasso, T. (1974). To organize is to remember: The effects of instructions to organize and to recall. *Journal of Experimental Psychology, 103*, 1014-1018. doi:10.1037/h0037407

Osman, M. E., & Hannafin, M. J. (1994). Effects of advance questioning and prior knowledge on science learning. *The Journal of Educational Research, 88*, 5-13. doi:10.1080/00220671.1994.9944829

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016). Is seeing gesture necessary to gesture like a native speaker? *Psychological Science, 27*, 737-747. doi:10.1177/0956797616629931

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie, 45*, 255-287. doi:10.1037/h0084295

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science, 11*, 137-138. doi:10.3758/bf03331011

Pennebaker, J., Chung, C., Ireland, M., Gonzales, A., & Booth, R. (2014). The development and psychological properties of LIWC2007. In *LIWC2007 Manual*. Austin, Texas: LIWC.net. Retrieved from https://pdfs.semanticscholar.org/5842/736189064114be6cbe04e6e8c239a9312c4e.pdf.

Pine, K. J., Lufkin, N., & Messer, D. (2004). More gestures than answers: Children learning about balance. *Developmental Psychology, 40*, 1059-1067. doi:10.1037/0012-1649.40.6.1059

Ping, R. M., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology, 44*, 1277-1287. doi:10.1037/0012-1649.44.5.1277

Ping, R. M., & Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science, 34*, 602-619. doi:10.1111/j.1551-6709.2010.01102.x

Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology, 5*, 1-14. doi:10.3389/fpsyg.2014.00359

Ramsey, J. L., & Fowler, M. L. (2004). 'What do you notice?' Using posters containing questions and general instructions to guide preschoolers' science and mathematics learning. *Early Child Development and Care, 174*, 31-45. doi:10.1080/0300443032000153471

Rapp, D. N., & Kurby, C. A. (2008). The 'ins' and 'outs' of learning: Internal representations and external visualizations. In J. K. Gilbert, M. Reiner, & M.

Nakhleh (Eds.), *Visualization: Theory and practice in science education.* (Vol. 3, pp. 29-52). Netherlands: Springer. doi:10.1007/978-1-4020-5267-5

Ratner, H. H. (1984). Memory demands and the development of young children's memory. *Child Development, 55*, 2173-2191. doi:10.2307/1129790

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*, 226-231. doi:10.1111/j.1467-9280.1996.tb00364.x

Reese, E., Haden, C. A., & Fivush, R. (1993). Mother-child conversations about the past: Relationships of style and memory over time. *Cognitive Development, 8*, 403-430. doi:10.1016/S0885-2014(05)80002-4

Reese, E., & Newcombe, R. (2007). Training mothers in elaborative reminiscing enhances children's autobiographical memory and narrative. *Child Development, 78*, 1153-1170. doi:10.1111/j.1467-8624.2007.01058.x

Rimé, B., Schiaratura, L., Hupet, M., & Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion, 8*, 311-325. doi:10.1007/BF00991870

Ritter, K. (1978). The development of knowledge of an external retrieval cue strategy. *Child Development, 49*, 1227-1230. doi:10.2307/1128766

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research, 3*, 131-141. doi:10.1016/0926-6410(95)00038-0

Roth, W.-M. (2003). Gesture-speech phenomena, learning, and development. *Educational Psychologist, 38*, 249-263. doi:10.1207/S15326985EP3804_4

Roth, W.-M., & Lawless, D. (2002). Scientific investigations, metaphorical gestures, and the emergence of abstract scientific concepts. *Learning and Instruction, 12*, 285-304. doi:10.1016/S0959-4752(01)00023-8

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science, 323*, 951-953. doi:10.1126/science.1167025

Sauter, M., Uttal, D. H., Alman, A. S., Goldin-Meadow, S., & Levine, S. C. (2012). Learning what children know about space from looking at their hands: The added value of gesture in spatial communication. *Journal of Experimental Child Psychology, 111*, 587-606. doi:10.1016/j.jecp.2011.11.009

Schacter, D. L. (2001). *How the mind forgets and remembers: The seven sins of memory*. USA: Houghton Mifflin Harcourt.

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual review of psychology, 49*, 289-318. doi:10.1146/annurev.psych.49.1.289

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology, 6*, 147-151. doi:10.1027/1614-2241/a000016

Schneider, W., & Sodian, B. (1997). Memory strategy development: Lessons from longitudinal research. *Developmental Review, 17*, 442-461. doi:10.1006/drev.1997.0441

Sekine, K., Sowden, H., & Kita, S. (2015). The development of the ability to semantically integrate information in speech and iconic gesture in comprehension. *Cognitive Science, 39*, 1855-1880. doi:10.1111/cogs.12221

Shore, C., Bates, E., Bretherton, I., Beeghly, M., & O'Connell, B. (1990). Vocal and gestural symbols: Similarities and differences from 13 to 28 months. In V. Volterra & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (Vol. 27, pp. 79-91). Berlin: Springer. doi:10.1007/978-3-642-74859-2_8

Simcock, G., & Hayne, H. (2003). Age-related changes in verbal and nonverbal memory during early childhood. *Developmental Psychology, 39*, 805-814. doi:10.1037/0012-1649.39.5.805

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16*, 85-89. doi:10.1111/j.0956-7976.2005.00786.x

So, W.-C., Chen-Hui, C. S., & Wei-Shan, J. L. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes, 27*, 665-681. doi:10.1080/01690965.2011.573220

So, W.-C., Ching, T. H.-W., Lim, P. E., Cheng, X., & Ip, K. Y. (2014). Producing gestures facilitates route learning. *PloS one, 9*(11), 1-21. doi:10.1371/journal.pone.0112543

So, W.-C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science, 33*, 115-125. doi:10.1111/j.1551-6709.2008.01006.x

So, W.-C., Shum, P. L.-C., & Wong, M. K.-Y. (2015). Gesture is more effective than spatial language in encoding spatial information. *The Quarterly Journal of Experimental Psychology, 68*, 2384-2401. doi:10.1080/17470218.2015.1015431

Stevanoni, E., & Salmon, K. (2005). Giving memory a hand: Instructing children to gesture enhances their event recall. *Journal of Nonverbal Behavior, 29*, 217-233. doi:10.1007/s10919-005-7721-y

Stieff, M., Lira, M. E., & Scopelitis, S. A. (2016). Gesture supports spatial thinking in STEM. *Cognition and Instruction, 34*, 80-99. doi:10.1080/07370008.2016.1145122

Suppes, A., Tzeng, C., & Galguera, L. (2015). Using and seeing co-speech gesture in a spatial task. *Journal of Nonverbal Behavior, 39*, 241-257. doi:10.1007/s10919-015-0207-7

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. New York, USA: Springer Science Business Media.

Tare, M., French, J., Frazier, B. N., Diamond, J., & Evans, E. M. (2011). Explanatory parent–child conversation predominates at an evolution exhibit. *Science Education, 95*, 720-744. doi:10.1002/sce.20433

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*, 24-54. doi:10.1177/0261927X09351676

Tenenbaum, H. R., Snow, C. E., Roach, K. A., & Kurland, B. (2005). Talking and reading science: Longitudinal data on sex differences in mother–child conversations in low-income families. *Journal of Applied Developmental Psychology, 26*, 1-19. doi:10.1016/j.appdev.2004.10.004

Thompson, L. A., Driscoll, D., & Markson, L. (1998). Memory for visual-spoken language in children and adults. *Journal of Nonverbal Behavior, 22*, 167-187. doi:10.1023/A:1022914521401

Thompson, L. A., & Massaro, D. W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology, 42*, 144-168. doi:10.1016/0022-0965(86)90020-2

Trafton, J. G., Trickett, S. B., Stitzlein, C. A., Saner, L., Schunn, C. D., & Kirschenbaum, S. S. (2006). The relationship between spatial transformations and iconic gestures. *Spatial Cognition & Computation, 6*, 1-29. doi:10.1207/s15427633scc0601_1

Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of verbal learning and verbal behavior, 6*, 175-184. doi:10.1016/S0022-5371(67)80092-6

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*, 187-204. doi:10.1016/S0361-476X(02)00007-3

Vasc, D., & Ionescu, T. (2013). Embodying cognition: Gestures and their role in the development of thinking. *Cognitie, Creier, Comportament, 17*, 149-169. Retrieved from https://search.proquest.com/docview/1399278125?accountid=14782

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner & E. Souberman., Eds.). Cambridge, Massachusetts: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin., Ed. and Trans.). Cambridge, Massachusetts: The Massachusetts Institute of Technology.

Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language, 50*, 395-407. doi:10.1016/j.jml.2004.01.002

Wang, X.-L., Bernas, R., & Eberhard, P. (2001). Effects of teachers' verbal and non-verbal scaffolding on everyday classroom performance of students with Down syndrome. *International Journal of Early Years Education, 9*, 71-80. doi:10.1080/09669760125651

Wermeskerken, M., Fijan, N., Eielts, C., & Pouw, W. T. J. L. (2016). Observation of depictive versus tracing gestures selectively aids verbal versus visual–spatial learning in primary school children. *Applied Cognitive Psychology, 30*, 806-814. doi:10.1002/acp.3256

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *The American Journal of Psychology, 114*, 591-600. doi:10.2307/1423612

Wheeler, M., Ewers, M., & Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory, 11*, 571-580. doi:10.1080/09658210244000414

Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concept. *Cognitive Science, 29*, 719-736. Retrieved from

https://pdfs.semanticscholar.org/7368/defd9b0ac6e48ccbb963c9290641eee8a563.p
df

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*, 625-636. doi:10.3758/bf03196322

Yoon, J., & Onchwari, J. A. (2006). Teaching young children science: Three key points. *Early Childhood Education Journal, 33*, 419-423. doi:10.1007/s10643-006-0064-4

Zohar, A., & Aharon-Kravetsky, S. (2005). Exploring the effects of cognitive conflict and direct teaching for students of different academic levels. *Journal of Research in Science Teaching, 42*, 829-855. doi:10.1002/tea.20075