CHAPTER 1

Introduction

1.1 Working Memory

WM occupies a central role in contemporary views of human cognition. WM can be broadly conceptualised as the set of processes which temporarily store, selectively attend to, and operate on the products of perceptual processing and activated long-term memory (LTM), generating conscious awareness, and intentional, goal directed mental activity such as task focus, problem solving, and planning (Baddeley, 1986; Miyake & Shah, 1999a; Shah & Miyake, 1999).

The construct of WM evolved from earlier conceptions of short-term memory (STM; e.g., Broadbent, 1957, 1958) when it was recognised that control processes beyond those subserving *maintenance* were necessary to explain the manipulation (processing) of information in the service of higher cognition. For example, Schiffrin and colleagues (R. C. Atkinson & Shiffrin, 1971; Schneider & Shiffrin, 1977) developed the modal model of human information processing, based on the flow of information through short-term memory (STM; which they termed *short-term store*). In this model environmental input passes through modality specific sensory registers ‘into’ a limited capacity STM, which also contains the activated portions of unlimited capacity LTM. Control processes such as rehearsal, coding, retrieval strategies, and decision rules operate on the contents of STM to achieve encoding into and retrieval from LTM, and to generate response output. LTM is considered to be functionally separate from WM. Interaction between the two memory systems occurs when portions of LTM become activated and are thus stored in STM, and when representations curently in STM are operated on by control processes and stored in

LTM. LTM can be activated by control processes, thus utilising WM capacity, or it can be activated automatically, independent of active control or attention from WM. The control processes are volitional. Atkinson and Schiffrin (1971, p. 83) conceived this amalgamation of activated representations in short-term storage and control processes as "… a working memory: a system in which decisions are made, problems are solved and information flow is directed."

This early model illustrates some key issues that have driven subsequent research into WM. First and foremost, WM is conceived as a dynamic *system* which is capacity limited, and fundamentally determines the characteristics of human cognition. Secondly, the modal model includes candidate control processes and posits a complex role for these in organising the flow of information through STM. Control processes (such as rehearsal) are implicated in keeping information active, and in facilitating activation (retrieval) of information from LTM, and encoding of new information into LTM. Thus, by inference control processes are implicated in WM capacity limitations. Thirdly, while the model assumes a functional distinction between WM and LTM, it also assumes a fundamental integration between these two components of what Atkinson and Schiffrin describe as the 'human memory system'. This integration can be mediated by volition, or by implicit—contextual—automatic activation. In either case, WM access to information from LTM is domain general, including related memory traces from all modalities. Fourth, the model explicitly recognises the possibility of domain specific components, and thus the *fractionation* of WM. It includes modality (domain) specific modules that deliver sensory input to WM buffers, and the authors note that interference effects are reduced when different *types* of information are competing for WM capacity. Finally, the modal model does not focus on attention, or capacity limits. However, in terms of WM capacity it makes the important association between the *focus* of attention, volitional control processes, and limitations in capacity of central cognition. Together these issues have shaped the progression of WM research, and the development of contemporary models of WM.

Because of the scope of WM in relation to higher cognition, any encompassing theory of WM would have to account for the broad range of empirical data that constitute the profile of cognition, and explain how executive functions (control) operate, how storage is accomplished, how WM and LTM integrate, and ultimately how these component processes might be constrained such that the capacity limitations which define human cognition emerge. Given such a large problem space, contemporary models of WM approach the construct from a variety of perspectives, drawing on different data streams from the breadth of cognitive psychology and the neurosciences to build both functional and computational information processing models of the WM system. This multi-faceted approach means the WM literature is wide ranging, with different models emphasising different aspects of cognition and different levels of analysis and explanation (see Miyake & Shah, 1999b). As such, no single model offers a comprehensive description of higher cognition, with some aspects better explained by one model, or another. The sections that follow identify areas of consensus across a range of WM models in terms of how the system stores and integrates different types of information (e.g., phonological/verbal; visuospatial); how it interacts with LTM; and how capacity constraints may be determined in the system. When the term *WM capacity* is used in this thesis it will refer broadly to the ability of the system to perform concurrent storage and processing activities without significant time cost.

1.1.1 Structure of the WM System

No contemporary model of WM subscribes to a purely unitary structural view, with general recognition of some distinction between those WM *processes* that engage in storage, and those that control and organise processing. However, there is some consensus that WM capacity limitations are determined by a limited pool of resources that is spread across all of these processes. Additionally, there is also recognition that there is some degree of domain separation, or specificity, in the way that the WM system deals with verbal and spatial information although conceptions as to how this is implemented vary (Miyake & Shah, 1999b).

1.1.1.1 Modalities: Verbal and Spatial

That WM has somewhat separable ways to deal with modality specific information, particularly verbal, and spatial information, is supported from a number of different lines of research. For example, factor analysis suggests some separability between the two domains. In a study where participants were asked to complete a range of tests assessing different types of information processing (verbal, visuo-spatial, and reasoning tasks) factor analysis revealed three clear factors. One was related to age and speed of processing, while verbal tasks clearly loaded on another factor, and visuo-spatial tasks on the third. When subjects performed a series of progressively more complex Tower of London tasks, a task demanding on spatial manipulation as well as executive function (though note that there are questions about its construct validity regarding executive function: Kafer & Hunter, 1997), their performance correlated with the visuo-spatial factor scores, not the verbal factor (Gilhooly, Wynn, Philips, Llogie, & Della Sala, 2002).

Dual task approaches have also shown a domain separability. Luck and Vogel (1997) were examining WM spatial capacity for features when they also required participants to keep numbers in memory. Holding a memory load which was maintained by articulatory rehearsal did not effect the capacity of the spatial store, suggesting a clear separation between the WM mechanisms for spatial and verbal information. Shah and Miyake found similar dissociation using dual tasks designed to measure WM span by invoking both processing and storage loads. They devised a spatial span task that made demands on processing and storage components of spatial working memory. Their results showed that this measure correlated with other spatial manipulation measures, but not with verbal measures. They also administered a classic verbal based span task, the reading span task (Daneman & Carpenter, 1980), and found that it correlated with verbal ability measures, but not with spatial ability measures. Additionally, a factor analysis clearly separated factors for spatial, and for verbal performance. In a second experiment they developed four different dual (span measure) tasks by crossing the processing task in the span task (mental rotation or sentence verification) with the type of information to be maintained (spatial orientations indicated by arrows or two-syllable words). They found patterns of interference when the processing and storage requirements were from the same domain, and improved performance when they were from different domains. Together, the findings from this elegant study strongly support the idea that there are separable domain specific components in WM, a conclusion supported by other studies looking at interference effects in dual task paradigms where visual or verbal primary tasks have been crossed with visual or verbal secondary tasks (Deyzac, Logie, & Denis, 2006; Hale, Bronik, & Fry, 1997).

Proactive interference (PI) has been another area of examination regarding domain separability in WM. PI is a reduction in the ability to perform a cognitive task because of prior performance of the same or a related task. It has been suggested that

PI occurs because information stored during previous trials interferes with the storage and processing of current information (Carretti, Mammarella, & Borella, 2011). PI has mostly been examined in relation to information presented in the auditory or the visual domain, with the idea that auditory information is received in the verbal domain, and visual in the spatial (Ronald H. Hopkins, Edwards, & Cook, 1973a). PI occurs most strongly when previous trials have required manipulation or storage of information in the same domain as the information that needs to be currently processed (Bowles & Salthouse, 2003; Emery, Hale, & Myerson, 2008; Lustig, May, & Hasher, 2001). If it is the case that there are separable components of WM that deal with verbal, and with spatial information, and prior engagement with one sort interferes with current processing of that same type of information, then after a series of trials presenting information in one modality switching modality for the final trial (on some occassions) should result in a release from PI, and therefore a performance improvement relative to no shift trials. This is precisely what has been found in studies of release from PI (e.g., Dean, Garabedian, & Yekovich, 1983; Ronald H Hopkins, 1973; Ronald H. Hopkins, Edwards, & Cook, 1973b; Kroll, Bee, & Gurski, 1973).

Finally, PET examination of brain regions involved in either verbal or spatial memory tasks has shown a dissociation in regions involved for the different tasks (Smith, Jonides, & Koeppe, 1996). The verbal task activated mostly left-hemisphere regions while the spatial task activated only right-hemisphere regions. In a follow up experiment subjects were shown an identical sequence of letters in all conditions, with a varied requirement to remember the names of the letters (which would use verbal memory) or the positions the letters occupied in the display (which would use spatial memory). In the verbal task, activation was concentrated more in the left than the right hemisphere; in the spatial task, there was substantial activation in both hemispheres, though in key regions there was more activation in the right than the left hemisphere. The researchers ran a third experiment where subjects engaged in a continuous memory task using only verbal material. This task activated the same regions as the verbal task used in the first experiment. In sum, these results indicate clearly that verbal and spatial working memory are implemented by different neural structures. Taken together the findings presented in this section support the idea that spatial and verbal information are treated in some separable fashion in WM.

1.1.1.2 Storage and Processing

The seminal evidence for a dissociation between storage and processing in WM came from tasks requiring two different activities to be performed at the same time, what are known as *dual*-tasks. For example, Hitch and Baddeley (1976) conducted an experiment in which participants were asked to perform a task which required them to repeat a list of numbers while performing a verbal reasoning task where they had to answer true or false to various questions. As the number of digits in the digit span task increased to six, participants took longer to answer the reasoning questions, and this effect was strongest for the more difficult questions, although they didn't make any more errors in the verbal reasoning tasks as the number of digits increased. The authors suggested that the process of articulating was performed by a different WM component than was used for reasoning, and only plays a minor role in reasoning activities. Once the load on both components increased sufficiently (to six digits, and the most difficult reasoning questions) a performance decrement in reasoning emerged, but prior to this none was evident.

The distinction between processing and storage in WM has also been illustrated by comparing the predictive capacity of simple versus compound memory tasks on performance in real world higher order cognitive tasks such as problem solving, and reading comprehension (see Cowan, 2001). While no single task can be seen as construct pure, simple tasks predominantly tap storage (i.e., STM) resources (e.g., Digit Span task; Wechsler, 1981), whereas compound tasks involve both storage requirements and a processing load (e.g., Reading Span task; Daneman & Carpenter, 1980). Thus, compound tasks can be seen as *dual tasks* which engage complex WM resources involving both short-term storage, and processing of information. Measures of WM capacity from simple tasks are not good predictors of performance at higher order cognitive tasks (Engle, Tuholski, Laughlin, & Conway, 1999), whereas WM measures from dual tasks do show a relationship to performance on higher order cognitive tasks (Daneman & Carpenter, 1980; Engle, Tuholski et al., 1999; King & Just, 1991; Kyllonen & Christal, 1990; Kyllonen & Stephens, 1990). These findings support the idea that the construct of WM encompasses the types of cognitive behaviour underpinning daily interaction with the world, and that to account for this human behaviour we need a construct that differs from a solely STM based temporary store by including mechanisms that actively manipulate stored information.

Evidence refining the relationship between STM and WM was found by Engle and colleagues (1999), who gave 133 participants two measures of general fluid intelligence (Ravens Progressive Matrices; Cattells Culture Free Test) and had them complete a battery of 11 memory tasks, some simple, thought to reflect predominantly STM demands (e.g., backward span; forward span, dissimilar; forward span, similar), and some compound, thought to reflect overall WM demands (e.g., reading span; operation span; counting span). The authors generated a series of path models which showed two distinct factors, with the dual tasks loading on one factor (WM), and the simple tasks loading on the second (STM). Performance on the WM tasks was significantly related to the measures of fluid intelligence, whereas performance on the STM tasks was not. However, WM and STM were highly related, and it was through their common variance that STM had any relationship to fluid intelligence. This finding indicates the importance of the processing component of WM in accomplishing complex cognitive tasks, as well as its reliance on the storage of information required to complete the tasks. Thus, this evidence suggests that storage and processing are seperable but fundamentally related components of the overall WM system.[1]

These findings suggest the obvious assumption that WM capacity limitations are related primarily to the mechanisms responsible for processing in WM, not those responsible for storage. This position is, however, contentious. The WM system is complex. Functionally, it must subserve the range of behaviours evident in higher cognition, and we know from the neuroscience literature that complex interrelationships unfold between many neural structures during the progression of a processing episode (Davidson, Ekman, Saron, Senulis, & et al., 1990; Klimesch, 1997; Krause, 2003; Potts & Tucker, 2001; Ruchkin, Grafman, Cameron, & Berndt, 2003). Given the scope of such a system it is entirely possible that fundamental constraints on the capacity to perform complex tasks could be determined by multiple factors (see Baddeley, 2000). Alternatively, analogous to Spearman's *g* (Spearman, 1923), there could be a unitary resource that ultimately determines cognitive constraints. The following sections present three models of WM and draw from them some consensus that capacity limitations result from a limited resource of *activation* needing to be shared across both processing and storage requirements.

---

[1] In light of this evidence, for the remainder of this thesis the term STM will be used to refer to a component of WM involved solely with *storage*. When the term WM is used it will refer to the overall processing / storage system, of which STM is a subset.

1.1.1.3 Engle’s Model

Engle and colleagues (Engle, 2001, 2002; Engle, Kane, & Tuholski, 1999) have presented a parsimonious model of WM which illustrates a relatively unitary perspective on capacity limitations. They suggest that WM is composed of a STM which consists of portions of LTM activated above threshold, a set of processes for generating and maintaining the activation, and *controlled attention* (CA), which they conceptualise as the ability to activate task relevant representations and maintain activation in the face of competition or interference. Thus, the activation fundamental to CA also encompasses the construct of inhibition. The CA model accounts for domain seperability by proposing that information is represented in various domain specific representational formats, all of which are accessable by CA, which is thus a unitary, domain general resource. Citing evidence from an array of studies examining the relative performance on higher cognitive tasks of high- versus low-WM capacity groups, assigned through performance on dual tasks like the Reading Span (Daneman & Carpenter, 1980) and Operation Span (Turner & Engle, 1989) tasks, they claim that it is CA which determines WM capacity, and so mediates performance on higher cognitive tasks. Engle’s conception of an *attention or activation based* constraint on WM capacity reflects a position that can be seen as common across a number of WM models which nevertheless differ in terms of specified structural properties.

1.1.1.4 Baddeley’s Model

A more fractionated conception of WM can be introduced by referring to Baddeley’s influential multiple component model (Baddeley, 1986; Baddeley & Logie, 1999; Ruchkin et al., 2003). This model consists of two domain specific storage systems, and a domain general storage and modelling space, all coordinated by an executive component, the central executive (CE). The two proposed domain specific storage components are the phonological loop for auditory information, and the visuospatial sketchpad for visuospatial information, each comprised of a ‘passive’ component—the phonological store, and the visual cache, respectively—and an ‘active’ component—the rehearsal system, and the inner scribe, respectively. A third component, the episodic buffer, has recently been proposed as a limited capacity short-term store where integrated representations—chunks—are generated through the binding of representations from the two storage buffers, and from LTM (Baddeley, 2000). The activity of these three sub-systems is coordinated by the central executive (CE) which executes control processes, manipulating information in the slave systems by switching attention, engaging in encoding and retrieval strategies, and organising the activation and binding (manipulation) in the episodic buffer of appropriate representations from the phonological loop, the visuospatial sketchpad, and LTM. The CE is fundamentally an attentional control system with no independent storage capability. Information is stored in the three sub-systems. The phonological loop and the visuospatial sketchpad are storage only components, able to engage in passive storage, or rehearsal processes, however the episodic buffer has the added properties of a domain general *modelling space*, where representations can be combined and manipulated to generate integrated representations in conscious awareness, and to solve novel problems. Importantly, even though combination and manipulation occurs on representations stored in the episodic buffer this processing is under the control of the CE and does not utilise resources from the buffer.

Thus, Baddeley’s model fractionates WM in terms of short-term storage, and processing, including the formal idea of domain specific storage components in WM that are further fractionated into functionally discrete subcomponents, and a domain general store with workspace capacities. While the model posits a mechanism for the activation and control of representations, the CE is a relatively undeveloped concept, essentially embodying the same type of intuitively derived control processes as in the earlier modal model. Thus, the broad terms in which its role is specified in the model (control of attention; manipulation of representations) allow it to ‘explain’ an essential aspect of cognition while not making claims yet to be supported empirically.

Baddeley does not specify a locus for the capacity limits of WM, claiming that each component of the model has capacity limits in terms of its specialised function and that limits can thus arise from multiple causes, such as capacity for activation, capacity for rehearsal, capacity for complexity of material, or efficiency at using acquired strategies and prior knowledge (Baddeley & Logie, 1999). However, in view of the organisation of the model an argument can be made that attention determines capacity in ways similar to Engle’s model. The CE has overarching control in the WM system, and has an intimate relationship with attention control, and the implementation of processing, with the subsystems only possessing storage properties. Because capacity measures using dual tasks show that efficiency in processing while carrying storage load predicts higher cognition performance (Cowan, 2001; Engle, Tuholski et al., 1999; King & Just, 1991; Kyllonen & Christal, 1990), it is reasonable to infer that WM capacity may be directly related to limitations in the CE’s capability to direct attention, and thus to activate representations and engage in binding operations in the episodic buffer. Therefore, as in Engle’s CA model, the multiple component model can be seen as supporting to some degree an attentional, or activation based account of capacity limitations.

However, inclusion of the episodic buffer presents some problems for this interpretation. The component is necessary in order to explain processing acitivities within the model—it is where the marriage of storage and processing occurs. For representations to become involved in processing by the system they need to be activated by the CE and moved into the ambit of the episodic buffer. Because the episodic buffer has a storage limit, it is possible that it could be the bottleneck in Baddeley's model. If information needs to be in the episodic buffer prior to the occurrence of any integration, or modelling of representations, then its limits could interact with the capacity for attentional deployment of the CE in two ways. If the capacity of the episodic buffer is *larger* than the attention available to be deployed by the CE, then CE capacity would constrain the capacity to perform cognitive tasks. If the capacity of the episodic buffer is *smaller* than the attention available to be deployed by the CE, then episodic buffer capacity would constrain the capacity to perform cognitive tasks. In either case, WM capacity could ultimately be defined as *attention available for utilisation*. The key difference would be the nature of the constraint—either attention itself, or storage capacity within which processing could occur.

1.1.1.5 Cowan's Model

Cowan (1999) also views WM capacity in terms of activation, citing it as a unitary constraint. As such, he contends that WM is basically a unitary system consisting of all currently activated LTM. A subset of this LTM is activated over threshold *a* and occupies STM. A subset of this activated information is further activated, over threshold *b*, and this occupies the focus of attention. Thus, activation and attention are integrally related. As in Baddeley's model, Cowan includes a CE component for directing attention, and controlling voluntary processing. He specifies some structure in terms of domain specificity by proposing, as does Engle, that information is represented in various domain specific representational formats, all of which are available for activation, which is a unitary, domain general resource. He contends that capacity limitations result from two constraints on activation. Firstly, a time limit on memory activation of approximately 10 to 20 seconds, and secondly, a size limit on the number of unrelated items that can be held active in the focus of attention.

In general these three models agree that WM involves abilities to store and process a spectrum of domain specific information types, with views ranging from Baddeley's conception of separate specialised components, to the existence of different representational codes proposed by Engle, and by Cowan. Each view recognises explicitly that WM shares a close integration with LTM. There can also be seen to be a consistent view that capacity to activate representations may be a critical determinant of WM capacity, and that the capacity for activation can be linked to a faculty of attention related to the control and organisation of storage and processing activities of WM. This faculty of attention may represent a general resource of available activation, which includes the capacity to integrate (process) representations in the service of task requirements.

1.1.2 The Activation Model

The key defining feature of WM in terms of the way it determines *how* humans can perform cognitive activities is its limited capacity. The three influential models of WM just presented can be seen to each view WM capacity as constrained by the availability of attention for activation. On the basis of this agreement the position taken on WM in this thesis will be summarised around a model which will be referred to throughout the thesis as the *Activation Model* of WM—a term coined by the author to capture the nature of this consensus. This model views the availability of attention for activation of representations as the determining factor in WM capacity limits. A crucial component of this position is the concept of capacity limitations in terms of a *processing episode*, which occurs over not more than perhaps ¾ of a second (and under less load may occur over much less time), a position which differs considerably from the general conception of how WM capacity is defined, and which it is hoped is introduced sufficiently in the following two paragraphs before being expanded on in Chapter 2.

Throughout the WM literature capacity has been operationally defined in ways that reflect an ability to engage in some demanding processing task while holding information active in storage. Classic examples of these *dual tasks* are the *reading span task* (Daneman & Carpenter, 1980) and the *operation span task* (Turner & Engle, 1989). In the reading span task subjects must read aloud sets of from 1 to *n* sentences, and after each set recall the terminal word of each sentence in the order they were presented. Recall is followed by the requirement to answer a veracity or comprehension question. The operation span is similar, with subjects having to read aloud and solve sets of from 1 to *n* arithmetic problems followed by a word, and then recall the words in order of presentation. The reading span task measures WM capacity, or ‘span’, in terms of the highest set size at which recall is accurate. The operation span measures capacity as the total words correctly recalled summed over all sets. The key point here is that these measures involve WM operations that progress over many seconds and fundamentally involve *ongoing* maintenance of information while processing activities proceed. In contrast, the type of capacity which is of interest to the research program presented in this thesis involves the *momentary* activation of representations and the swift binding or integration of these representations during a processing episode—a concept that will be referred to as *chunking capacity*, and that is fundamentally related to the use of event-related potentials as the primary dependent variables in this research program, and what they tell us about the time course of context updating in WM. These details will be made clear in Chapter 2. For the moment the reader is asked to become familiar with the *concept* of this type of capacity in the context of the activation model of WM developed so far.

To summarise this position in terms of performing a specific task of increasing levels of complexity, where at some level the task requirements can no longer be completed within a single processing episode: in *the activation model of WM*, attention is seen as a mechanism with three operational properties: (a) Attention activates verbal and / or spatial (and perhaps tactile, and auditory non-verbal) sensory representations, and domain general representations from LTM related to current sensory input, to the task, and to the task procedure (e.g., rules); (b) It *maintains* representations active during the sequence of operations required to complete a trial episode; and (c) It *binds* these various representations to accomplish these operations, if necessary producing *intermediate representations* over which further binding operations may occur to produce task relevant output—*integrated representations* that involve evaluative outcome and response selection. These properties are constrained by a fourth, fundamental factor which determines a processing episode—only a limited amount of activation is available to the system *at any given time*. While there are sufficient resources of activation available, binding can occur across multiple representations, including intermediate representations, without *substantial* time cost—that is, within a processing episode. However, once storage and binding operations exceed the available activation, continued binding is achieved through sequential implementation of operations in subsequent, ongoing processing episodes, as serial processing occurs. It is important to note here that this conception does not preclude the operation of serial processing *within* a processing episode. Put another way, the concept of a processing episode does not imply that all processing done during an episode is parallel, nor that there is an all or none transition from parallel processing, which can be accomplished within the duration of an episode, to serial processing which exceeds the capacity of an episode. Rather, as shall be shown in Chapter 2, the capacity of a processing episode is sufficient to absorb some serial processing prior to these costs becoming too great, at which time the processing can no longer be accomplished within the short duration of a single episode, and fully blown serial processing ensues.

1.1.2.1 The Range of Attention

Another concept important to the inferential scope of this thesis is the *range of attention*, a term that is used to define the boundaries of central cognition with respect to the perceptual processes that deliver its sensory input (see Block, 2007; Fodor, 1986). In this sense attention refers to top-down influence on the operations required to process a stimulus through to conscious evaluation, and often in the case of laboratory studies, to generate task specific output. This top-down behaviour may not be entirely reliant on volition. Task requirements, once learnt, would constitute activated LTM structures, utilising little resources while determining the activation of appropriate semantic traces, and the organisation of processing resources and the progression of the operations required for task completion, an idea proposed by Schneider and Shiffrin in 1977 (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977), that is consistent with contemporary theory (Baddeley & Logie, 1999; Ericsson & Delaney, 1999). In terms of the functional sequence of operations required to transduce a stimulus and process it through to evaluation and then response selection, the earliest stage of processing at which these top-down determinants have influence constitutes the range of attention, and thus the functional scope of the WM system.

The range of attention defines the beginning of integrative processes in WM. *Thus, a process that changes behaviour as a function of task dependent activation demands can be said to be within the range of attention, and a part of WM.*

1.1.3 Summary

This section has developed an operational definition of WM, the *Activation Model*, introduced the concept of a *processing episode*, and suggested in terms of the Activation Model that there is a relationship between the amount of information which can be processed across a single processing episode and WM capacity to perform binding and maintenance operations—a concept introduced as *chunking capacity*. A further important concept, the *range of attention*, has also been introduced. Together these concepts are important tools referred to throughout this thesis to make inferences about the involvement of WM in subitizing during item enumeration.

1.2 Item Enumeration and Subitizing

1.2.1 Subitizing and Counting

Subitizing has been generally defined as the capacity limited process involved in swiftly, accurately, and effortlessly apprehending the numerosity of small collections of perhaps three or four items (Kaufman, Lord, Reese, & Volkmann, 1949). While subitizing has been examined in the haptic (e.g., Plaisier, Bergmann Tiest, & Kappers, 2009; Plaisier & Smeets, 2011; Plaisier, Tiest, & Kappers, 2010), and auditory domains (e.g., Camos & Tillmann, 2008; Repp, 2007) the construct was first recognised (Jevons, 1871) and has been examined predominantly in the visual domain. In this approach subitizing occurs in the context of visually presented arrays of varying number of items presented simultaneously, with the task requirement being to enumerate the array and produce a response indicating that enumeration has occurred; that is, in the context of what is referred to throughout this thesis as *item enumeration*. Subitizing is defined in contrast to the time consuming, or error prone, construct of *counting*, which is required to enumerate collections of items exceeding the subitizing capacity.

Two different methods have been used to examine item enumeration—tachistoscopic displays with short display duration, often around 200ms (e.g., Dehaene & Cohen, 1994; Kaufman et al., 1949; Mandler & Shebo, 1982 Experiments 1, 2, 3, & 4; Nan, Knösche, & Luo, 2006), and subject determined display durations where display terminates upon subject response, typically a voice key or button press, and accuracy and speed are stressed as task requirements (e.g., Balakrishnan & Ashby, 1991; Jensen, Reese, & Reese, 1950; Logie & Baddeley, 1987; Mandler & Shebo, 1982 exp. 5; Trick, 1990; Van Oeffelen & Vos, 1982b; Wolters, Van Kempen, & Wijlhuizen, 1987). Tachistoscopic presentations have typically yielded a contrast between three processes defined by an interaction between RT and accuracy: Fast, accurate responses for one to three item arrays, reflecting *subitizing*; more prolonged RT and small but increasing proportion of errors for four, five and six items, reflecting post-perceptual *counting*; and a levelling of the RT slope from six to seven items and beyond, with sharply increasing error rates as numerosity increases, post-perceptual processing resources are taxed, and a process of *estimation* comes into play (for review see Gallistel & Gelman, 1991; Mandler & Shebo, 1982). Because the subject determined display time which defines the response contingent display method precludes the need for estimation to occur as array numerosity increases, and because accuracy is stressed, this method has typically shown a contrast between two processes: Arrays of perhaps three or four items are enumerated swiftly and accurately with only a small time cost per item—they are subitized—while in arrays with greater number of items accurate enumeration of each extra item incurs a much larger time cost, as the time consuming process of counting occurs. Thus, in response contingent displays subitizing and counting are defined by differing characteristics of the RT slope function, while in tachistoscopic displays interaction between the RT slope, and error rate functions defines the transitions between subitizing, counting, and estimation (see Figure 1.01).

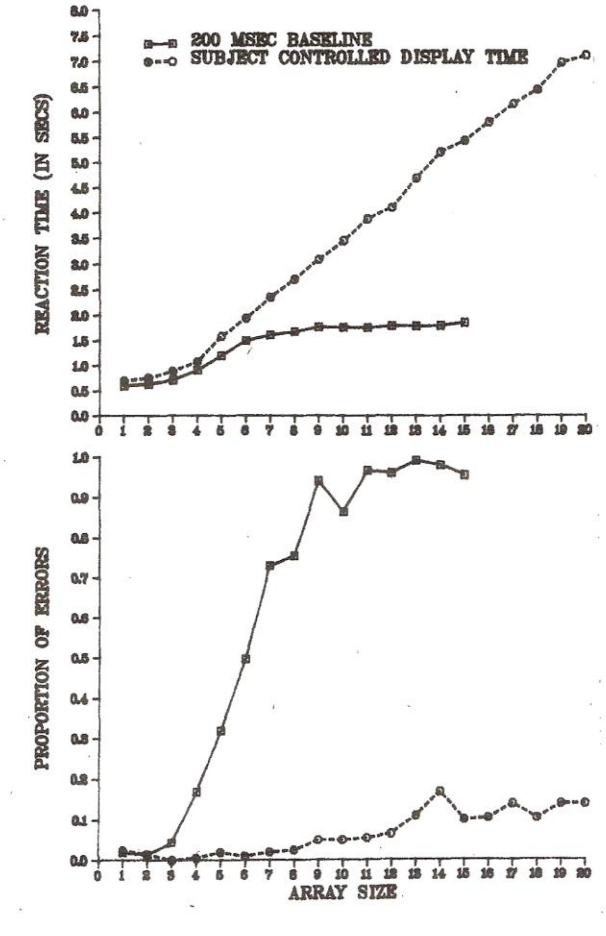


Figure 1.01. Reaction time and proportion of errors for response contingent v. tachistoscopic (200 ms) display durations as a function of array size. Taken from Mandler and Shebo, 1982 (Exp. 5 and Exp. 1, respectively).

The s*ubitizing span* (or *range*) varies between people. Some have suggested an upper limit of perhaps six or seven items (e.g., Mandler & Shebo, 1982) based on tachistoscopic displays, though a more generally accepted range is typically 4+/-1 items (Basak & Verhaeghen, 2003; Folk, Egeth, & Kwak, 1988; Gallistel & Gelman, 1991; Jevons, 1871; Trick, Enns, & Brodeur, 1996; Trick & Pylyshyn, 1993; Warren, 1897). Given that most adults are still subitizing at 3 items, the typical *subitizing slope* taken over the range of 1 to 3 items *averages* between 30ms and 60ms per item, in both tachistoscopic and response contingent displays (see Figure 1.01). However, in response contingent displays, for 5 items and above the *counting slope* is typically 250ms to 350ms / item (see Gallistel & Gelman, 1991; Mandler & Shebo, 1982, Experiment 5; Svenson & Sjöberg, 1983; Trick & Pylyshyn, 1993, 1994b). In response contingent displays this discontinuity in RT slopes has been the classical indicator of transition from subitizing to counting, with the point of significant change in function taken to indicate the extent of the subitizing range (though for dissenting view see Balakrishnan & Ashby, 1991).

This pattern of results, which is exemplified in Figure 1.01, indicating a contrast between swift, accurate enumeration of small numerosities, and time consuming, or error prone enumeration of larger ones, has been robust across a plethora of studies(e.g., Gallistel & Gelman, 1991; Kaufman et al., 1949; Mandler & Shebo, 1982; Saltzman & Garner, 1948; Svenson & Sjöberg, 1983; Trick & Pylyshyn, 1993, 1994b; Warren, 1897) since first documented by Jevons in 1871 . These data indicate incontrovertably that there is some sort of difference in the way that small versus larger collections of items are enumerated; that is, between subitizing, and counting. There have been two enduring schools of thought regarding the nature of this difference. The first claims that the two processes are *qualitatively different*; that is, that subitizing is some specialised process (of one sort or another) dedicated to the apprehension of small numerosities, while counting involves serial processes common to central cognition; that is, WM processes (e.g., Butterworth, 1999; Pylyshyn, 1989; Trick, 1990; Tuholski, Engle, & Baylis, 2001). Throughout the remainder of this thesis views of this nature will be referred to as *dual-process* explanations. The second school of thought claims that both subitizing and counting are subserved by processes common to central cognition, but in *quantitatively different* ways; that is, that WM constraints determine the defining characteristics of both subitizing, and counting (e.g., Boles, Phillips, & Givens, 2007; Burr, Turi, & Anobile, 2010; Chi & Klahr, 1975; Cowan, 2001; Egeth, Leonard, & Palomares, 2008; Jevons, 1871; Klahr, 1973a, 1973b; Klahr & Wallace, 1973; Olivers & Watson, 2008; Poiese, Spalek, & Di Lollo, 2008; Vetter, Butterworth, & Bahrami, 2008; Xu & Liu, 2008). Throughout the remainder of this thesis views of this nature will be referred to as *single-process* explanations.

For both single-and dual-process views a common explanation of the swift enumeration and limited capacity typical of subitizing has been that it reflects parallel processing of the array elements by some capacity limited system, or process. In common dual-process explanations this might be achieved variously by: (a) a capacity limited perceptual mechanism dedicated to pre-attentive item individuation, such as Pylyshyn’s Fingers of INSTantiation (FINST) mechanism (Pylyshyn, 1989, 2001); (b) a modular processing mechanism such as Butterworth’s proposed number module (Butterworth), an innate, capacity limited neural facility dedicated to number processing; (c) a specialised, innate non-verbal counting mechanism which processes representations of ‘number’ as anologue magnitudes along a number line, similar to the way temporal durations may be represented (the approximate number system; ANS Gallistel & Gelman, 1991); (d) an associative mechanism linking learned canonical patterns to numerosity concepts, such as “oneness” or “twoness” (Mandler

& Shebo, 1982); or (e) a more active pattern matching mechanism that determines degree of similarity to past exemplars prior to association of a numerosity concept (Logan & Zbrodoff, 2003). In typical single-process views explanations have historically appealed to the limited capacity of *storage* in WM, with the idea that there are a limited number of storage ‘slots’ in visual WM. Prior to this capacity being exceeded the quantification operations required for enumeration can proceed in parallel, however, once the slots are filled, quantification operations must ensue serially, and thus the time costs associated with counting emerge (e.g., Chi & Klahr, 1975; Cowan, 2001; Klahr, 1973a, 1973b; Klahr & Wallace, 1973).

Between these typical views there is no argument that WM resources are fundamentally engaged in counting, given what we know about its component processes (Butterworth, 1999; Logie & Baddeley, 1987). When the number of elements in an array exceeds the subitizing capacity and counting must occur, items or item clusters (see Van Oeffelen & Vos, 1982a for a discussion on group and quantify strategies) must be selected and enumerated in turn until all items have been processed, the correct number word is retrieved, and a response made. Counting requires spatial storage and processing—to avoid directing attention towards previously counted items, the locations of these items need to be stored (marked; see Atchley, Jones, & Hoffman, 2003; Kunar, Humphreys, Smith, & Hulleman, 2003; Olivers & Watson, 2003), and this information must be accessible to inform subsequent deployment of attention, and ultimately to signal search completion when no unmarked items remain, and thus cue final number word retrieval, and response. Secondly, verbal storage and processing is also involved. A running tally of items counted must be maintained, and updated, a process that many believe involves sub-vocalisation, and thus verbal processing and storage resources (see Logie, 1987;

Pylyshyn, 1999, 2001; Trick, 1994; Trick, 1990). Finally, in order to tally across successive episodes of grouping and quantifying, access is required to an ordered list of number names, and to rules for generating these number names in terms of addition. Presumably this semantic information is accessed from long-term memory, and must be held active to be available for use. Thus, counting is characterised by spatial attention shifts, and storage and processing demands in both the spatial and verbal domains, and the requirement to activate and utilise task relevant information from LTM. Together, these taxing demands on WM resources require serial operations, which determine the time costs of the counting slope.

There is a key element which challenges the perhaps implicit, and certainly enduring belief that, in contrast to the time consuming, serial processing necessary for counting, the characteristics of subitizing are driven by parallel processing of some sort or another, *and that as numerosity increases within the subitizing range these processes do not demand additional attentional resources,* as each of the typical explanations—both single, and dual—has suggested. If the proposed mechanisms are truly underpinned by parallel processing then the addition of extra items *within the subitizing range* should not incur a time cost. However, this is not the case. With very few exceptions studies of item enumeration *do* produce a subitizing slope (of between 30 and 60 ms average increment per item, as previously stated), and the few exceptions (e.g., Atkinson, Campbell, & Francis, 1976; Mandler & Shebo, 1982, Experiment 4; Sagi & Julesz, 1985a, 1985b) all include atypical methodological features (e.g., manual response that may have occurred prior to enumeration being completed; specific, non-variant shapes for each numerosity with learnt associations to number words; blocked numerosities with only binary response such as 1 vs. 2, 2 vs. 3, or 3 vs. 4, within blocks) which can be seen to influence the processing which typically might generate the subitizing slope (for review and analysis see Folk et al., 1988). Therefore, any compelling explanation of subitizing needs to provide an explanation of the subitizing slope.

1.2.2 The Non-Linear Subitizing Slope

A common explanation for the subitizing slope has been an appeal to a serial, self-terminating scan through a list of ordered number names stored in LTM (Gallistel & Gelman, 1991; Klahr & Wallace, 1973; Mandler & Shebo, 1982; Trick & Pylyshyn, 1994b), with the notion that scanning time per list item should be around 30 to 45 ms as Sternberg's (1966) memory scanning data indicated. However, while this might be appealing when examining slope gradients computed as *average* increment per item (see Trick & Pylyshyn, 1994a p.89 for a summary of slope average gradients), a finding noted across a range of studies has been that the subitizing slope is described by a slowly accelerating function—that is, the increments between one and two items (perhaps 20 ms), and between two and three items (perhaps 50 ms) are not of equal magnitude (Balakrishnan & Ashby, 1991; Dehaene & Cohen, 1994; Gallistel & Gelman, 1991; Mandler & Shebo, 1982; Railo, Koivisto, Revonsuo, & Hannula, 2008). For the memory scanning explanation to be compelling there should be no differences in access time across integer words in this range because scanning time per list item should be equal, thus the non-linear steps suggest that the subitizing slope is determined by something other than WM scanning an ordered list of number names. The obvious suggestion here is that some serial, attention demanding processing during subitizing is causing the subitizing slope, and the slope's non-linear function indicates that the addition of each extra item increases the processing demands in an exponential rather than a linear fashion.

Such an explanation would be very difficult to reconcile with dual-process views of subitizing, which by definition dismiss any role for WM in the quantification process(s) underlying subitizing, and which already appeal to the additional and supposedly serial, *linear* WM process of memory scanning to explain the subitizing slope. It would also be difficult to reconcile with typical STM capacity based single-process views, which to explain the slope also appeal to the same serial, supposedly linear WM process occurring *after* parallel subitizing activity. An obvious alternative explanation is that rather than an all or nothing transition from parallel processing (subitizing) to serial processing (counting), item enumeration involves a *gradual transition* from very fast counting in WM when only one or two items need enumerating, to slow and steady counting in WM for five items and beyond, with a transitional phase of serial processing occurring across two, three, and four items. This proposed single-process *serial counting explanation* of subitizing would be consistent with the gradually accelerating function of the RT slope. It is supported by a recent body of evidence suggesting that serial processes under the control of attention are involved in the enumeration of even two items (Burr et al., 2010; Egeth et al., 2008; Olivers & Watson, 2008; Poiese et al., 2008; Vetter et al., 2008; Xu & Liu, 2008), and it can explain the non-linear subitizing slope in terms of the activation model of WM function outlined in section 1.1.2, through a proliferation of binding operations and intermediate storage requirements as processing load (item number) increases.

1.2.3 The Serial Counting Explanation

The serial counting model of subitizing is put forward here as an explanation of subitizing in terms of serial WM activity against which to test the suppositions of other models which claim that alternative types of activity determine subitizing.

Based on the activation model of WM function the basic proposition of the serial counting explanation is that the characteristics of the subitizing slope, and the limits of the subitizing span, are fundamentally determined by a gradual loading of WM resources as the demands of both storage (number of items, and intermediate products) and processing (number of binding operations to perform quantification operations on the collection of items) increase, until they can no longer be completed within the capacity of a single processing episode. These demands on *activation resources* (i.e., attention) might involve two processes, even within the supposed subitizing range; individuating items as separate quantification units, which may be *a spatial process,* along with the operation on these units of further quantification operations which proceed serially, involve binding operations which produce both intermediate and final products, and which may or may not involve subvocalisation, and thus recruit *verbal processing resources* (Wender & Rothkegel, 2000). Note that individuation in this sense does not mean figure ground segregation or the like, but the assigning of constituent identity to the item. For example, one item would hypothetically require two binding operations interceded by a storage requirement, prior to a response being initiated; individuation in terms of a quantification unit with the numerosity value ‘1’, maintaining the integrated representation in an active mode, then binding it with the number word ‘one’. However, two items would require four operations, as well as maintenance of two intermediate products; individuation of each item in terms of a quantification unit with the numerosity value ‘1’ (i.e., two operations), storage of these intermediate representations, binding the individuated items into a quantification unit with the value ‘2’, storing that integrated representation, and finally binding it with the number word ‘two’. In turn, three items would require six operations, and the associated increase in intermediate storage demands. Thus, each increment requires some increasing time to resolve the operations, until the capacity of activation in WM to absorb the processing costs within a single processing episode is fully exceeded and lengthy delays ensue as counting must occur.

This description illustrates how serial activity in WM could determine the characteristics of both the subitizing and counting slopes, with each determined as a function of activation available over time. For subitizing, when two, and then three units of information are presented there is sufficient activation available to accomplish the required binding and maintenance operations *with minimal but increasing processing time cost*, but still within a single processing episode. Once activation is fully utilised—that is, when the subitizing span is reached—the binding and maintenance operations required over additional units must be accomplished over more time as available activation performs work and must then be redeployed to perform the remaining operations necessary to complete the task. This process represents what we term ‘subitizing’ as very fast counting, without the major shifts of spatial attention required for counting larger numbers of items, and importantly, characterises WM activity during subitizing as being serial from the very start of processing load, a position which can be supported by evidence from some recent studies addressing the role of attentional resources in subitizing.

1.2.4 Attentional Involvement in Subitizing

A number of recent, ingenious studies have found convincing evidence suggesting that subitizing recruits attentional resources, by showing that disrupting attentional deployment disrupts subitizing for even two items. Moreover, there are indications that this disruption can be greater for three items than for two, findings which support the explanation for the non-linear subitizing slope provided by the serial counting explanation.

In 2008 Olivers and Watson used a rapid serial visual presentation (RSVP) paradigm to assess enumeration performance under conditions of attentional blink—that is, where attending to a target stimulus in the stream induces a failure to detect (attend to) a stimulus following closely (perhaps 100 to 400 ms) in the stream. Olivers and Watson reported three experiments. The first simply presented dot arrays for 133 ms, with pre- and post-masks, with no RSVP, to confirm that their stimuli could be subitized at this duration. Subitizing performance (lack of errors for 1 to 3 dots; increasing errors thereafter) was consistent with typical tachistoscopic presentation. The second experiment consisted of RSVP of between 15 and 20 letters displayed for 100 ms each with intervening 33 ms blank screen. One letter, the target stimulus, differed in colour from the others and was presented 1, 2, 3, 5, or 8 temporal positions from the end of the letter sequence. At the end of the series a dot array (from 0 to 5 dots) was displayed for 133 ms, followed by a mask. Thus, the dot array followed the target letter by lags of either 133, 267, 400, 667, or 1067 ms. Subjects had to report, accurately, the identity of the target letter first, followed by the number of dots they had seen. This response order was designed to maintain processing on the primary task (letter identification) during and after presentation of the dot array. The logic of this experiment was simple, given the established finding of an attentional blink under brief lag conditions—if subitizing requires attentional resources then subitizing performance would suffer at the shorter lags, and be preserved at the lengthier ones. This is precisely what Olivers and Watson found. For two, and three items, relative to the 1067 ms lag condition, there were over 20%, and 25% more errors, respectively, in the 137 ms lag condition. There were no differences for a single item. The parameters of the third experiment were identical to experiment 2 except primary and secondary tasks were swapped—the RSVP stream was composed of dot arrays (0 to 5 dots), one of which (the target array) differed in colour from the rest and required enumeration, and each series was terminated by a target letter. Letter identification was compromised at short lags, but not at the long lag, a pattern identical to that found in experiment 2 for dot enumeration. Critically, at the shortest lag letter identification was impaired even when only a single dot had to be enumerated, and this impairment became worse when two dots were presented. Thus, enumeration of even two dots was compromised by the attentional blink, an effect that was worse for three dots, and subitizing of even a single dot caused an attentional blink that compromised subsequent letter identification.

Using similar lags, stimulus presentation, and response parameters Burr (2010, Experiment 1) had subjects either perform the classic attentional blink paradigm—that is, attend to a target letter, and attempt to enumerate a following dot array—or ignore the letters, and attend only to the dot array, constituting a *dual*- or *single*-task condition, respectively. They then computed *Weber fractions* (standard deviation of accuracy / number of items) as measures of performance precision for each numerosity, and lag. Relative to the single-task condition, performance in the dual-task condition for 2, and for 3 items was significantly impaired at the shorter lags, but not at the longer one, and this impairment was greater for three items than for two, exactly the pattern of results found by Olivers and Watson.

Egeth (2008) also used the attentional blink to examine the recruitment of attentional resources in subitizing. Even though there were differences in methodology, with much shorter display durations, their findings were consistent with those of Olivers and Watson (2008), and with Burr et al. (2010). Egeth et al. used

RSVP of 16 letters displayed for 50 ms each, followed with a 50 ms blank screen, with one letter being red designated as the target letter. Once in each 16 letter stream a letter was ringed by a circular array of from 1 to 9 dots, with this array appearing either simultaneously with the target letter, or 1, 2, 3, or 4 letters later. Thus, the dot array lags were either 0, 100, 200, 300, or 400 ms. The letter displayed after the dot array was ringed by a mask. At the end of each stream subjects were required to report accurately the target letter, followed by the numerosity of the dot array. In the subitizing range dot enumeration accuracy was unaffected at the 400 ms lag, however, at the shorter lags performance was impaired for two items, and more so for three items.

The final pertinent attentional blink study was conducted by Xu and Liu (2008). Similar to Burr et al. (2010), Xu and Liu included dual- and single-task conditions, although there were some other considerable methodological differences to the attentional blink studies detailed above, notably: (a) dot arrays were displayed for three times as long as the preceding RSVP stimuli; (b) the shortest lag corresponded to the median lags from the other three experiments; (c) the response was a forced binary choice with a cue (that is, a same / different response); and (d) the response for T2, the dot array numerosity, was asked for prior to response for T1. In their experiment the RSVP stream consisted of 14 uppercase letters each displayed in black, except for the target letter (T1) which was green, for 50 ms followed by a 40 ms blank screen. The target letter was always either “A” or “U”, and appeared as either the eighth or twelfth letter. Following the letter stream a display of from 1 to 6 dots (T2) randomly placed around a circle was presented for 150 ms. Thus, T1 to T2 lag was either 270 ms or 630 ms. T2 was followed by two successive 100 ms masks, after which a prompt appeared along with a digit (1-6), and the subject was required to respond as to whether the perceived numerosity of the dot array was the same as or different to the displayed digit. A same / different response was then asked for in relation to the T1 identity. Relative to the single-task condition, in the dual-task condition Xu and Liu found a reduction in enumeration accuracy for the dot arrays of three items, but not also for arrays of two items as Olivers and Watson (2008), Burr et al. (2010), and Egeth et al. (2008) all found. However, the methodological differences noted above could have had an influence on this result. In the other studies the strongest influence of the attentional blink was at the shortest lags, and Xu and Liu only included a median length lag of 270 ms. Additionally, having subjects respond first with the forced choice, cued numerosity judgement might have reduced the impact of the dual-task manipulation (having to process T1), as well as reducing the amount of errors reported by only having to make a same / different judgement. Therefore, together these issues may well have reduced the effectiveness of the paradigm at reducing attention available to process the dot array, as well as making it easier to provide the correct answer regarding numerosity (Note: Xu and Liu also collected event-related potential data which also indicated recruitment of attentional resources during subitizing, but this will be discussed in a later section).

These studies using the attentional blink to examine recruitment of attention during subitizing have been presented in some detail to illustrate that even across divergent methodology the findings have been consistent, showing that (in three of the four studies detailed) when attentional resources are compromised there is a performance decrement for the enumeration of even two items, and that this decrement is greater with the addition of a third item. In the Xu and Liu study which didn’t show this pattern at two items, there was an effect for three items, and the lack of effect at two items can be explained by the methodological differences to the other three studies. Overall, the atentional blink manipulation of attentional resource deployment during subitizing robustly generates results which are consistent with the idea that incremental loading of attentional resources with additional item load during subitizing generates the non-linear subitizing slope, a body of evidence which supports the speculations of the serial counting explanation.

Another similar approach used to examine attentional involvement in subitizing has been to introduce various *dual-task* manipulations into tachistoscopic presentations and measure accuracy, either raw error rates or by Weber fractions, relative to single-task performance. For example, Vetter et al. (2008) presented 200 ms displays where a central diamond pattern comprised of four coloured triangles was surrounded by a circular array of 9-13 gabor patches, with from 1-8 being targets (vertical oriented, high contrast), and the remainder being horizontally oriented low contrast distractor items. In the dual-task condition subjects were given instructions to detect in the central diamond either a simple feature (low load condition, primary task), or a conjunction of features (high-load condition, primary task), as well as detect the numerosity of targets in the surrounding array (secondary task). Subjects had to report whether the target feature was present or not, and then the numerosity of the targets. In the single-task condition subjects performed the primary, and secondary tasks separately. For the enumeration task, for each of 1, 2, and 3 items Weber fractions were significantly greater for the dual-task relative to the single task, with the high-load condition being substantially greater than the low-load condition. Thus, under conditions where attention was utilised most by the primary task, subitizing efficiency was most impaired, and this impairment occurred even for a single item. Burr et al. (2010, Experiment 2) used an almost identical spatial dual load procedure and also found exactly the same profile of Weber fractions for 1, 2, and 3 items.

Finally, in a clever design Poiese et al. (2008) utilised information about the time course of primary visual processing found from intracranial recordings in monkeys, and from transcranial magnetic stimulation in humans, to manipulate SOAs between a dot array and a mask (which would disrupt any ongoing, incompleted processing) to assess whether attentive processing is necessary for subitizing. Their logic suggested that if subitizing is determined by preattentive processing this processing would either be accomplished in primary visual cortex (V1), where it would be completed within less than 50 ms post-stimulus, or in extrastriate cortex during the subsequent feedforward sweep where it should be completed prior to 100 ms post-stimulus. In two experiments they showed target arrays of either 0, 1, or 2 items followed by a mask, and varied the array-mask SOAs from 50 to 83 ms in Experiment 1, to test area V1, and from 100 to 150 ms in Experiment 2, to test the feedforward sweep, and measured enumeration accuracy. Because both SOAs in Experiment 1 were less than 100 ms their hypothesis was that if subitizing is determined by preattentive processes in V1 then because the information will have left V1 on the feedforward sweep prior to the backward mask effect occurring, accuracy should be identical for both SOAs, and because two items is within the subitizing range accuracy should be identical for 1, and for 2 items. However, if subsequent attentive processes are crucial in subitizing, they reasoned that the longer SOA should allow more information to be transmitted to these processes leading to an improvement in target detection and so subitizing accuracy relative to the shorter SOA, and also, because in the 2 item condition attention would need to be shared (split) between the items, that accuracy should be better for 1 item than for 2 at both SOAs. The same predictions were made for Experiment 2, which tested the feedforward sweep as the locus for preattentive influence on subitizing. In each experiment accuracy was better at the longer SOA, and for 1 item over 2 at both SOAs. The authors concluded that rather than subitizing being determined by preattentive processes, attentional processing occurring after 150 ms post-stimulus was crucial for subitizing to proceed.

Together these studies provide convincing evidence supporting the idea that attention is involved in subitizing, and thus supporting single-process views of item enumeration rather than dual-process views. Additionally, these findings question the typical single-process view regarding WM *storage capacity* (i.e., STM) as the determinant of the limited subitizing capacity. This basic assumption that WM storage limits determine subitizing contraints has endured, perhaps because of an implicit belief drawn from classic dual-process models, that subitizing is indeed driven by parallel processing of some sort or another. However, the studies reviewed above have conceptualised and examined WM involvement in subitizing in terms of utilisation of attentional resources, a position more consistent with contemporary views of WM, with the activation model introduced in section 1.1, and with the serial counting explanation of subitizing presented in section 1.2.3. Importantly, the robust findings that disrupting attentional deployment impairs enumeration performance for just two items, and then generates additional impairment for three items, do not just speak against dual-process explanations of subitizing, but against the existence of *any* parallel processing mechanism underpinning the swift enumeration of small collections of items, and particularly against any preattentive mechanism.

The idea that preattentive processing underpins subitizing has been put forward in two main dual-process explanations of item enumeration, which will be addressed in the following two sections. The first claims that subitizing comes about through the operation of preattentive visual processes— the FINST hypothesis (Pylyshyn, 1989;

2001; Trick, 1990; Trick & Pylyshyn, 1994)—while the second claims that subitizing is an innate capacity determined by modular facilitation through dedicated neural structure(s) specialised for number processing (Butterworth, 1999), or for representing approximate numerosity (Cantlon, Brannon, Carter, & Pelphrey, ; Gallistel & Gelman, 1991, 2000).

1.2.5 The FINST Hypothesis

An influential preattentive account of subitizing has been put forward by Pylyshyn and Trick (Pylyshyn, 1989; Trick, 1990, 1992; Trick & Pylyshyn, 1993, 1994b), which claims that, because item enumeration is fundamentally a visual task, subitizing is determined by preattentive visual processes. Trick and Pylyshyn have proposed that a limited capacity spatially parallel mechanism operates preattentively to individuate items by binding feature clusters to ‘mental reference tokens’, FINSTs, once basic visual processing (e.g., figure ground segregation; feature registration) has occurred. These tokens then become available to WM for processing. Capacity of this system is determined by a basic limitation in the number of tokens, with estimates varying from four to six (Pylyshyn, 1989, 2001). The FINST hypothesis was developed in relation to a variety of visual phenomena such as multiple object tracking, with a primary focus on the property of *indexing* objects or items in the visual field. Indexing involves *individuation* of feature clusters, which are then bound to a token along with information *locating* the item in the visual field. The binding of a feature cluster to a token is unique, and persistent through changes in the item’s location (Scholl, Pylyshyn, & Feldman, 2001; Sears & Pylyshyn, 2000). As movement occurs new location information is dynamically bound to the index. Thus, indexing provides a mechanism whereby constituent identity is possible. Such a process is obviously necessary prior to *relative* locations being computed, and readily explains our capacity to keep track of dynamic objects in the visual field which change their relative spatial relationships, a critical component of everyday visual cognition.

Trick and Pylyshyn suggest that FINSTs allow subitizing to occur during static item enumeration, primarily by performing the critical function of individuating the target items. To support the role of automatic individuation in subitizing they present evidence showing that the subitizing slope is increased substantially, from around 60ms / item to 200ms / item, when, for example, item individuation is disrupted by embedding target items amongst conjunctive distractors (e.g., ‘O’ amongst ‘Qs’ Trick & Pylyshyn, 1993). Because these stimuli share a major conjunction, target items do not *pop-out*, and individuation requires an effortful comparison process by WM. All available tokens are utilised in the comparisons, so automatic target individuation is impossible, subitizing cannot occur, and serial processing, signaled by attention shifts, and marking, updating, and storage demands, becomes necessary even when there are only one or two target items. Thus, the authors claim that the capacity to index items in parallel determines the capacity to subitize, and that this capacity is disrupted by manipulations that force focused spatial attention. They go on to suggest that enumeration involves the interaction of two separate processes. Counting is reliant on WM through goal directed (attentional) processing of the FINSTs’ content, and reallocation of their focus once available tokens are full. However, the attentional ‘work’ that underlies this grouping process (see Van Oeffelen & Vos, 1982a for a discussion on group and quantify strategies), and thus subitizing, occurs preattentively, in the visual processing system at the stage of item individuation.

While it is entirely plausible that aspects of the visual system constrain the capacity for information to be *delivered* to WM, and so determine the rate of information processing, it is not entirely clear that this equates to subitizing being an automatic, pre-attentive phenomena determined by the number of available FINSTs. FINSTs don’t bind with just *any* feature clusters in the visual scene. Under conditions of goal orientation their focus can be directed towards specific features and feature conjunctions so long as these features have already been indexed (Pylyshyn, 1989, 2001). Thus, rather than being exclusively pre-attentive, FINSTs might well reflect early integrative processes in WM as *task relevant* features of the visual information become preferentially activated and are thus selected for subsequent processing. This supposition is supported by a study comparing visual WM (vWM) capacity, numerosity estimation precision, and subitizing capacity (Piazza, Fumarola, Chinello, & Melcher, 2011), which found that neither vWM capacity, or subitizing was associated with numerosity estimation precision, but that vWM capacity and subitizing capacity were strongly positively correlated. Additionally, subitizing performance was impaired by a concurrent vWM load, with subitizing capacity decreasing in proportion to the vWM load. The authors concluded that loading vWM impaired efficiency at object individuation, thus suggesting that individuation processes necessary for efficient subitizing cannot be preattentive, the same conclusion arrived at from the attentional studies summarised in section 1.2.4.

In terms of the range of attention, FINSTs could thus be conceived, at least in part, as an early selective mechanism of visual WM, a position supported by Poiese et al. (2008) findings regarding the time course of visual processing in subitizing, which suggested that if any visual processing constraints are involved in determining subitizing then this processing occurs sometime after 150 ms post-stimulus, when we know that top-down processes are already influencing selective attention (Steven J. Luck, 1995; Steven J. Luck & Hillyard, 1995; Martinez et al., 1999).

Given this explanation, and Pylyshyn’s (2001) indication that only 4 to 6 FINSTs are available at any one time, it is plausible that the quantity of FINSTs may relate to vWM capacity, and be causal in the subitizing *limit*. It is also plausible that the integration of the items referenced by each FINST token into a quantification unit occurs in WM, during which the FINSTs are ‘emptied’, by being *counted*, a position taken by Pylyshyn himself (2001, p. 135):

> “According to our explanation of the subitizing phenomenon, small sets are enumerated faster than large sets when items are preattentively individuated because in that case each item attracts an index, so observers only need to count the number of active indexes without having to first search for the items.”

Thus, processing, or ‘emptying’ of the FINSTS must be accomplished by WM, as the items—each referenced by the token bound to them—are counted. In this scenario the subitizing slope must be determined by quantification processes over each FINST. The existence of the subitizing slope means that even if the passage of the FINST’s content into WM (unloading) might occur in parallel, the quantification operations can’t, or the slope would be zero. In essence then, the FINST account can be seen as a counting explanation of subitizing with a specified individuation mechanism, which can be construed as a capacity limited input stage for visual WM. This mechanism achieves goal directed individuation based on perceptual features, however there is no indication in the theory that there is any automatic binding of two or more FINSTs in terms of quantification units. As such, quantification identity must be bound to the individuated items in WM as the counting process proceeds.

Interpreted in this way, the FINST account is compatible with the serial counting explanation of subitizing.

1.2.6 Modular Views and Neural Activity

Another dual-process view which implies that preattentive processing underpins subitizing has been that subitizing is an innate, specialised faculty. For example, Butterworth (1999) has presented the idea of a number module, which he states (p.226) "… human infants are born with, and which we share with non-human species, [and it] is known to function for numerosities up to about 4, 5 at the very most. That is, our brain is genetically programmed to represent numerosities up to 4." However, because modular processing is held to be very fast, and automatic (Fodor, 1986), the number module view cannot easily account for the non-linear subitizing slope over three items, which is within the supposed 4, or 5 item capacity of the module, within which processing should be parallel. Indeed, Butterworth (p.274) states that "… it takes no longer to name the number of dots when there are four dots than when there is just one dot", a statement which clearly disregards the numerous findings of a subitizing slope, non-linear or otherwise. Additionally, the studies reviewed earlier which indicated a key role for attentional resources in subitizing are also inconsistent with the modular conception of preattentive processing.

A second conception of some innate mechanism dedicated to numerosity judgements which can be swift and accurate over small collections of items has been put forward by Gallistel and Gelman (1991; 2000). Their preverbal counting explanation of subitizing appeals to the existence of an ANS which nonverbally represents number as a magnitude representation along a mental number line, in a similar sense as temporal durations are represented (Gallistel & Gelman, 1991, 2000), and lays down this representation as a memory trace which can then be accessed to make judgements and subsequently, by accessing the correct number word, to make statements about numerosity. The ANS mechanism is shared across species, and is present in infants, developing throughout the first years (for review see Dehaene, 1997). However, the ANS represents number only approximately (Dehaene, 1997; Gallistel & Gelman, 2000), and due to the property of scalar variance—the signals which encode magnitudes are imprecise, and the signal variance increases as numerosity increases—precision deteriorates markedly with increasing numerosity (Gallistel & Gelman, 2000). For this reason preverbal counting by the ANS has been proposed as a viable ‘very fast counting’ mechanism that can be accurate over only a few items, and thus can subserve subitizing, but after only a few items the growing uncertainty means that counting proper is the process of choice for accurate enumeration. The non-linear subitizing slope could be explained by reference to the increasing variability at 2, and then at 3 items requiring additional ‘checking for accuracy’ processes, however it is not clear how this process would then actually differ in practice from the serial counting explanation, particularly for 3 items. Additionally, as for the number module, the evidence suggesting a key role for attentional resources in subitizing is also inconsistent with the preverbal counting suggestion of preattentive processing.

A key feature of both the number module and the ANS views is that subitizing is determined by innate endowment that is realised through the operation of a dedicated neural structure or structures. Both views propose similar brain regions as being responsible for the neural activity subserving subitizing, notably the intraparietal sulcus, with the middle and inferior frontal gyrus also being implicated, and there are suggestions of a left hemisphere bias (Butterworth, 1999; Cantlon et al., 2006; Kucian et al., 2006). Thus, if these views are correct there should be clear evidence of differences in neural structures being active during subitizing, and during counting.

There is some evidence suggesting that different neural structures may be involved in subitizing, and counting, however this does not necessarily imply that the processes are qualitatively different. In one study Dehaene and Cohen (1994) used a sample of patients with a deficit in visual perception of complex scenes, while still being able to recognise individual objects (simultanagnosia), to examine naming time while performing an enumeration task, and a visual search task. In the enumeration task subjects counted arrays of 1-6 rectangles, varying in stimulus duration, either 200ms, or until 1500ms after response, and configuration, either canonical patterns, or random dispersement. In the visual search task subjects had to identify the presence or absence of a single target (a red horizontal rectangle) amongst an array of 1-6 vertical or horizontal red or green rectangles which were displayed until 900ms after a response. There were three conditions: 1) a ‘pop-out’ condition where all distractors were green rectangles; 2) an orientation condition, where all distractors were red vertical rectangles; and 3) a conjunction condition, where half the distractors were green horizontal rectangles, and the other were red and vertical. Subjects were variously able to subitize 1, 2, or 3 items, and performance in the pop-out condition was mostly preserved. Together, these findings suggested that low-level visual processes allowing individuation were relatively intact in these patients, as was the capacity to bind quantification identity to items. However, beyond 2 or 3 items serial counting was markedly slowed, with many errors, and performance in the conjunction search condition was also very slow, though few errors were made. Of the five patients, four of them had lesions around the right parietal region, an area where damage has been associated with disruption of spatial attention. Because the patients could individuate items, but could only quantify small, not large sets, the authors reasoned that counting must recruit spatial processing resources that are beyond the item individuation required for subitizing.

On face value these findings seem to support a dual-process view of item enumeration, with the suggestion that subitizing and counting utilise different functional mechanisms, which, in line with modular views, may be implemented by separable neural ensembles, with counting but not subitizing specifically reliant on capacity for spatial attention and memory demands (*which* items have been examined) which might be subserved by activity in the right parietal lobe (Corbetta, Shulman, Miezin, & Petersen, 1995). However, an alternative interpretation is more consistent with a single-process explanation—that is, that there are some WM related processes specifically related to counting which are not required for subitizing, notably coding of numerous spatial relations, and memory for inspected locations, marking. Thus, as spatial demands increase at three items and beyond, deficits in these visuo-spatial WM processes due to lesions lead to the observed impaired performance. In fact, links between specific visuo-spatial WM impairments and behavioural performance dissociation between subitizing (preserved) and counting (impaired) have also been found by Demeyere et al.(2010) in a detailed neuropsychological assessment of a patient with bilateral posterior parietal lesions, who failed in marking efficiency (memory for inspected locations) and who exhibited seriously impaired enumeration of 4 or more items, but retained subitizing ability.

Conversely, while certain parietal damage is associated with deficits in coding of numerous spatial relations, and memory for inspected locations, other parietal structures have been implicated in both visual short-term memory, as well as efficiency of enumeration estimation. The role of the intraparietal sulcus, notably left hemisphere, in visual short-term memory has been clearly illustrated by fMRI data collected in a series of studies reported by Todd and Marois (2004), and the functional role of the intraparietal sulcus in relation to enumeration estimation has been established in both adults, and young children (Cantlon et al., 2006; Kucian et al., 2006). Thus, a clear link can be drawn between neural structures involved in attention and visual short-term memory, and those which are involved in subitizing.

Other examinations of brain regions that may exhibit differential involvement in subitizing, and counting, have used methods such as positron-emission tomography (PET), and electroencephalography (EEG). For example, Sathian et al. (1999) used PET to examine item enumeration and reported that subitizing engaged an area in the occipital extrastriate cortex, most likely linked to early visual processing, while counting activated a wider network of neural structures, particularly the superior parietal cortex, and the right inferior frontal cortex, consistent with activity involved with shifting visual attention. Also using PET, Piazza and colleagues (2002) found extrastriate, middle occipital, and intraparietal regions were activated for both subitizing and for counting. This network became more active during the counting range (6-9 dots in this case). In another study, Nan et al. (2006) used EEG to assess enumeration with arrays presented for 300ms. Using source analysis they found indications of activation in almost identical areas in the inferior parietal cortex associated with both subitizing and counting.

While the activation of additional, or wider neural areas might at first glance seem to support the modular, dual-process view of item enumeration, together these findings can be interpreted in terms of the activation model as providing some neural evidence reflecting a transition from storage to storage and processing activities in WM, as more brain regions are recruited to join the functional network performing the processing as load of operations increases, or already active regions engage in more intense processing. The data are broadly consistent with this view, suggesting increased involvement of components of a wide ranging neural network as processing load increases. This pattern could reflect the operation of functional connectivity—the moment by moment binding of diverse brain regions into a dynamic functional system by the synchronous oscillations of the ensembles of neurons which are required to perform the various operations during unfolding of a processing episode (John, Easton, & Isenhart, 1997). In this sense, rather than suggesting support for dual-process accounts of item enumeration such as the number module, or ANS approaches, the apparent increases in sites of neural activity between subitizing and counting could reflect the emergence of indicators of the gradually increasing demands on activation of the task as maintenance and processing requirements increase. Thus, while some differences in neural activity between subitizing and counting have been noted this neural activity can be interpreted in terms of WM activity both within and beyond the subitizing range and so does not necessarily preclude a single-process explanation of item enumeration, and could be consistent with the speculations of the serial counting explanation.

1.2.7 Canonical Patterns and Pattern Matching

There are two other closely related dual-process views that need to be outlined. The first, made popular by Mandler and Shebo (1982), appeals to associative spatial matching mechanisms, to explain subitizing. They suggested that small numbers of items necessarily appear in configurations that reflect simple geometric patterns—canonical patterns—such as pairs which are perceived as lines, triples as triangles (and quadruples as quadrilaterals, although there would be more variance in this production, and so the pattern’s occurrence would be less reliable), and that associations between these patterns and numerosities are acquired through multiple exposures as we develop from child to adult. Mandler and Shebo suggested that this explanation explains the change from the steep subitizing slope displayed by young children (perhaps 180 ms per item) to the much shallower one adults typically produce (< 60 ms item). Thus, in this view subitizing is subserved by pattern matching mechanisms, and associative learning.

The learnability of canonical patterns, and their influence on reducing enumeration times for multi-item collections was shown by Allen (2008, Experiment 2) where air traffic controllers (whose work involved long-term exposure to engaging with spatial arrangements) were significantly better than matched controls at enumerating arrays of 5 items and beyond when these were arranged canonically, but not for linear arrays. Wolters et al. (Wolters et al., 1987) also found evidence for learnability of canonical arrays when, after 5 days of training in associating invariantly configured arrays of from 4 to 18 dots with the appropriate numerosities, subjects were capable of all but immediate, accurate recognition of numerosity for all arrays—that is, they learnt to ‘subitize’ the arrays. In contrast, while there were some practice effects, no firm learning took place on randomly configured arrays which continued to generate counting slopes throughout the 5 days of exposure. With arrays within the subitizing range, Wender and Rothkegel (2000) also found that randomly configured arrays resisted learning, with canonical arrays being enumerated quicker even after equal pre-exposure. In fact, after pre-exposure random arrays were enumerated quicker than at initial exposure, but so were the canonical arrays. This latter finding was interesting because if canonical patterns are enumerated quickly because of pre-existing overlearning, a further advantage due to only a few extra exposures during an experimental procedure would be unexpected. The authors took this finding to suggest that it may not solely be overlearning that drives recognition of canonical patterns, but a more active pattern recognition mechanism.

This suggestion is interesting because a purely associative mechanism would imply that there should be no additional time cost for enumeration of two items over one item, or three items over two items because in each case the only operation required would be to retrieve the number word associated with the prelearned pattern—that is, the association should imply a direct relationship with no requirement for any serial memory scanning through a list of number names—so there should be no subitizing slope, and certainly no non-linear slope. Thus, existence of the slope further supports the idea that if recognition of canonical patterns is implicated in determining the swift enumeration of subitizing then it is more likely to be the result of some more active pattern recognition process than a purely associative process resulting from overlearning.

Logan and Zbrodoff (2003) approached this active pattern matching theory in a way that is consistent with the activation model. They suggested that retrieval of a numerosity related to a pattern relies on the degree of activation in a pattern matching mechanism, with this being determined by degree of similarity to previous displays of the same or similar numerosity. Activation is high to highly similar displays, and low to highly dissimilar displays. Because the addition of each extra item increases the ways in which configuration of arrays can vary from other exemplars, smaller numerosities should reliably produce more activation, and so the pattern matching mechanism should be more efficient. When they assessed similarity ratings to simultaneously presented arrays of randomly configured dots that were either the same, or different numerosities, subjects were far more accurate in rating correctly arrays in the subitizing range (1-3 items) than those in the counting range. The authors reasoned that this finding supported the idea of an active pattern matching mechanism being implicated in subitizing, but being dismissed in favour of serial counting when numerosity made the process too inaccurate.

While pattern matching accounts of subitizing are appealing because they draw on general cognitive mechanisms—the capacities to generate associations, and to perceive configural similarities—the existence of which are well established, and which we use in a range of circumstances interacting with the world at large, it may be that for subitizing, pattern matching and association are not used exclusively. Under some circumstances configuration of the items in an array undoubtedly facilitates swift enumeration, but existence of the subitizing slope, and more so the non-linear subitizing slope, calls into question the assumption that pattern matching suffices to explain entirely the characteristics of subitizing. If a pattern of two items, which can only be a ‘line’ is matched to previous exemplars of ‘line’ and there is a learned association between ‘line’ and numerosity ‘2’, then why should the matching operation and then retrieving this numerosity take any longer than the same operations for a single item (dot = ‘dot’, association = ‘1’)? The same question can be asked for three items versus two. While some studies presenting canonical patterns do not find a subitizing slope (e.g., Mandler & Shebo, 1982, Experiment 4), the majority of those presenting randomly configured arrays do find a slope (see section 1.2.1). The obvious suggestion here is that when specific patterns are not evident in the array (and this may even be specific orientation of ‘line’ for 2 items) further, time consuming processing of the items is required to bind them into a ‘configuration’ to which a quantification identity can be bound, which, if such configurations are associated with a particular number word, would then allow the appropriate numerosity to be retrieved. The non-linear subitizing slope in particular suggests that, certainly for three items, some triplets may not be readily matched to previously seen ‘triangle’ configurations, and thus would require additional, time consuming processing for enumeration to occur. This account provides one possible explanation (over and above concentration on the task) for the trial by trial variability seen within the subitizing range, as well as existence of the subitizing slope.

Additionally, it makes explicit the idea that while learned configurations and associations might provide one path towards swift, accurate enumeration of a few items, real world configurational variability might require that other, more effortful processes be recruited to perform enumeration, at least sometimes.

Indeed, if pattern recognition were the process by which subitizing *always* occurs, but other, effortful processing by WM is how counting proceeds, then it would be expected that different brain regions would be involved in subitizing, and in counting, by the same logic that this would have been expected under the modular accounts. As already explained in section 1.2.6, there is no compelling evidence that this is the case, with differences in neural activity between subitizing and counting equally well explained by appeal to an explanation implicating increasing storage and processing demands in WM, such as the serial counting explanation. This conclusion is supported by PET data from Piazza et al. (2002) who presented subjects with both random and canonical configurations to enumerate and found that in the subitizing range RTs were the same for canonical and random configurations, while areas in the bilateral occipitotemporal cortex that have been linked with object recognition and spatial relations were activated for both types of configuration, as was the intraparietal sulcus, which, as has been stated, is known for involvement in numerical processing. These areas became more active as a function of number of items, but there were no effects reported specific to the canonical configurations. However, in the counting range (6–9 items) for random but not canonical configurations there was additional activity noted in the right superior parietal lobe, an area known for involvement in shifting of spatial attention (Corbetta et al., 1995).The authors concluded that during subitizing pattern recognition processes were recruited by *both* canonical and random configurations, as were processes related to numerical processing, while during counting random configurations recruited additional, spatial resources than did canonical configurations.

1.3 Summary and Progression

The preceding sections have made clear that longstanding single- and dual-process explanations of item enumeration which suggest that subitizing is determined by parallel processing mechanisms of one sort or another, which may or may not be preattentive, are not clearly supported by a range of findings. The common existence of a subitizing slope, often non-linear, suggests some sort of time consuming processing occurs over each additional item in the subitizing range, notably when array configuration is random and enumeration may require greater involvement of spatial processes. There have also been a number of findings indicating that disrupting attentional deployment disrupts subitizing of even two items, a disruption that becomes greater as a function of additional item load, and this disruption has been linked to processes operating in vWM. Additionally, examination of neural activity during item enumeration does not reveal the clear dissociation in active regions that would be expected if subitizing and counting were entirely separable processes. Rather, brain regions subserving object recognition, spatial attention, and numerical processing are active in both subitizing and counting, with activity increasing as a function of item load, and additional regions linked to attentional processing ‘coming online’ in the counting range. Finally, regular configurations of items can facilitate enumeration in both the subitizing and counting ranges, suggesting that patterns can be learnt, and that pattern matching processes can effectively identify patterns that are associated with particular numerosities. However, irregular configurations still result in a subitizing slope, which is often non-linear, and neural activity associated with object recognition and spatial relations is evident for both regular and irregular configurations, suggesting that pattern matching cannot entirely account for subitizing, and that these visuospatial processes may be engaged at different intensities, depending on configuration, to accomplish the visuospatial manipulations necessary for enumeration to occur. Thus, rather than supporting dual-process views of item enumeration based on preattentive, parallel accounts of subitizing, or single-process views based on parallel processing, together these findings instead point to the idea that attentional resources, and serial processing in WM, might well be involved in enumerating even two or three items, and thus to a single-process explanation of item enumeration such as the serial counting explanation put forward in section 1.2.3, that does not assume that preattentive, or parallel processing determines subitizing.

As was noted in sections 1.2.2 and 1.2.3, the proposed serial counting explanation of subitizing is based on the activation model of WM function developed in section 1.1 and outlined in section 1.1.2, where the limited capacity of WM to perform maintenance and binding operations without significant time cost during a *processing episode* is determined by available activation for attention, a central coordinating resource which is shared across both storage and processing demands. Within this operational definition of WM, the serial counting explanation made speculations as to the types of operations involved in subitizing, suggesting that for even two or three items it may not be accomplished by parallel or preattentive processes, but require active resources of attention and may involve time consuming, serial processing.

Of course, the way to examine whether or not attention, and perhaps serial operations, are involved with subitizing would be to follow from stimulus presentation the flow of information processing involved with the operations which lead to the behavioural manifestation of subitizing and assess the impact of attentional processing at stages in this flow. However behavioral measures cannot follow information flow at such a fine grained level, and the predominant dependent measures used throughout studies of item enumeration and subitizing have been the behavioural measures of error rates, and RT. Error rates are a summary measure and while informative, provide no direct insight into what processes might determine subitizing. In the same vein, RT is an ‘offline’ measure signaling culmination of stimulus evaluation, and response selection and execution, and therefore can provide no real information in terms of the flow of perceptual and cognitive processes that may be recruited for subitizing during item enumeration. As such, these measures cannot answer important questions about WM involvement in subitizing, notably regarding the location of the boundary between preattentive and attentive processes, or between parallel and serial processing, if indeed one exists. Thus, if subitizing is defined by the operation of parallel or preattentive processes, and counting by serial processes, it might not be possible to determine accurately through behavioural measures the item level at which a transition between these processes occurs.

To achieve insight into the interplay of these processes a dependent variable is needed which can provide fine grained information about the flow of functional information processing during enumeration. Event-related potentials are ideally suited to this task. They have millisecond resolution, provide continuous information throughout the unfolding of a processing episode right from stimulus presentation through to response execution, and a body of knowledge exists regarding the possible functional significance of a number of event-related potential components related to both sensory and cognitive processing. While there have been a handful of event-related potential studies that have briefly addressed item enumeration using tachistoscopic displays (reviewed in section 2.1.4), to the author's knowledge there have been no published studies using event-related potentials to examine item enumeration and subitizing using response contingent displays. The research program presented in this thesis uses response contingent displays to examine the role that WM might play during subitizing, with event-related potentials as the main dependent measure. The following chapter introduces event-related potentials, outlines the components of interest, and makes predictions as to how these might be expected to behave when indicating WM activity during item enumeration.

CHAPTER 2

Event-related Potentials

2.1 Introduction

From electrodes affixed to the scalp it is possible to record bioelectric potentials generated by brain activity; the electroencephalogram (EEG). Precise segments, *epochs*, of the ongoing EEG (which may have durations of from a few hundred milliseconds, to 3 seconds or more) can be time-locked around the presentation of stimuli requiring performance of tasks reflecting distinct types of psychological events; sensory processing, selective attention, memory updating, or semantic analysis, for example. Numerous epochs relative to presentation of the same class of stimuli can be extracted from the ongoing EEG and averaged together to extract from the general electrical ‘noise’ of ongoing brain activity signals representing specific processing activity reflecting these sensory or cognitive operations. These averaged, time-locked signals are known as event-related potentials (ERPs). Because of their low signal to noise ratio ERPs (signal, perhaps +/- 5 μV) need to be extracted from the ongoing EEG (noise, perhaps +/- 50 μV) by this averaging procedure (Rugg & Coles, 1995a) Note that there are methods to estimate ERPs from single trials but these are not relevant to this thesis (e.g., Effern et al., 2000; Laskaris & Ioannides, 2002; Shin, Talnov, Matsumoto, & Brankack, 2001; Wagner, Röschke, Grözinger, & Mann, 2000).

In the derived waveforms generated by this averaging process, ERP *components* are typically measured relative to a prestimulus baseline (perhaps 100-200 ms) and identified by reference firstly to their *relative* polarity, positive or negative—that is, whether the peak results from positive *going* or negative *going* activity independent of its absolute amplitude value, though some components such as the P300 and the N400 do tend to have absolute amplitude values which are positive, and negative, respectively—and their latency from stimulus onset, either in ms, or as rank. Thus, if an averaged waveform’s 3rd major positive going peak had a latency of around 300 ms, it could be referred to as a P300, or a P3. Scalp topography, and type of eliciting stimulus or task are also important in determining the nature of an ERP component. For instance a P3 generated in response to an attentionally demanding task, and which has a centro-parietal *maximum* (often called a P3b) is likely to be associated with memory or context updating (Donchin & Coles, 1988). In contrast, a P3 in response to a novel stimulus, and which has a more frontal maximum (often termed a P3a) appears to reflect an orienting response to novel stimuli (Fabiani, Gratton, Karis, & Donchin, 1987). Similarly, tasks requiring resolution of semantic anomalies generate pronounced negative going deflections with a peak latency of approximately 400+ ms post-stimulus onset, and a predominantly anterior distribution, the N400 component (Marta Kutas & Iragui, 1998), and an anterior negative going component peaking around 200-300 ms post-stimulus, the N2, has been associated with cognitive control and attentional deployment in Go/NoGo, and sequential matching tasks, while a posterior negative going component with similar latency has been associated with visual attention tasks (Folstein & Van Petten, 2008). It is important to note two issues here. Firstly, some components, such as the P3b and the N2 examples mentioned here, have a wide scalp distribution, and for proper identification and interpretation both the task demands and sites of maximum expression need to be taken into account (Coles & Rugg, 1995; Fabiani, Gratton, & Coles, 2000; Rugg & Coles, 1995b). Secondly, nomenclature throughout the ERP literature is quite inconsistent, mostly because of the issues related to defining a component. Thus, components identified by different names by different researchers in differing paradigms may have similar functional significance.

ERPs represent the net dipolar electrical field produced through summation of the synchronous activation of one or more neural populations (generators), where within each population the individual neurons are oriented roughly in parallel, as in the cortex (Coles & Rugg, 1995), and so their postsynaptic potentials summate (Davidson, Jackson, & Larson, 2000). Where the internal orientation of neuron populations is inappropriate (e.g., in many mid-brain structures such as the thalamus) their processing activity does not summate in a dipolar field, and so does not contribute *directly* to the scalp recorded EEG. Consequently, while the thalamus, for instance, is implicated in regulating the oscillation of cortical EEG (Bastiaansen & Brunia, 2001; Danos, Guich, Abel, & Buchsbaum, 2001; Davidson et al., 2000; Larson et al., 1998), neural activity of the thalamus itself cannot be recorded as electrical activity at the scalp (intracranial recordings are required for this). Therefore only a subset of the brain’s electrical activity is available *directly* for the EEG. Furthermore, the brain is a volume conductor, and the skull is a low conductance electrically distortive medium. These features together make it difficult to infer accurately the actual neural generator sites of ERPs (Fabiani et al., 2000), although substantial gains have been made in the area of source analysis with the advent of high density recordings (128 sites and above), and improved signal processing algorithms (e.g., Dümpelmann, Ball, & Schulze-Bonhage, 2011). However, the primary focus of the ERP research program reported in this thesis is not on identifying brain regions involved in subitizing, but rather on indexing the temporal dynamics of functional aspects of mental activity during subitizing. So, while the EEG has poor spatial resolution, it has excellent temporal resolution, in the order of milliseconds, and thus, because ERP components can provide information about the time course of possible operations as a processing episode unfolds they are an ideal tool for making inferences regarding the types of cognitive operations which might be involved in subitizing.

2.1.1 Making Inferences From ERPs

ERPs have been studied since the early 1960s (e.g., Donchin & Lindsley, 1965; Donchin, Wicke, & Lindsley, 1963) and during this time relationships between a number of ERP components and some relatively specific cognitive (or sensory) processes have been established (Fabiani et al., 2000; Rugg & Coles, 1995a, 1995b). However, defining a ‘component’ can be problematic because the underlying neural processes of some cognitive activity or processing can overlap in both space and time with other processes, meaning that the critical features of the ERP waveform such as peaks or troughs can be caused by the summation of several contributing sources, and so may not reflect functionally homogeneous neural or cognitive processes. This makes clear interpretation a matter of careful consideration of the functional components of the task that is eliciting the processing activity. In fact, the underlying logic of this interpretive approach can be applied even if the features of the ERP waveform do not overtly conform to ‘established’ components—indeed, this is a common way that new information about some components is derived, and new component definitions developed (Davidson et al., 2000; Fabiani et al., 2000; Rugg & Coles, 1995a). This inferential approach relies on presenting conditions that differ in some way that can be interpreted as implying a functional difference, or not, in cognitive or sensory demands, and then observing three aspects of ERP differences which might manifest between conditions—differences in time course, amplitude, and distribution across the scalp. Differences, or not, in these three properties can then be used to infer the degree of functional equivalence of the underlying cognitive or sensory processes.

2.1.2 Inferential Approach

In the research reported in this thesis inferences regarding the role of WM activity during subitizing will be made using as a starting point the well recognised ERP component, the P3b, and its established properties of reflecting aspects of cognitive activity associated with the time course of stimulus evaluation in WM, and of the processing load associated with this (Dien, Spencer, & Donchin, 2004; Donchin & Coles, 1988; Fabiani et al., 1987; Johnson, 1988; Kok, 2001; Magliero, Bashore, Coles, & Donchin, 1984; Verleger, 1997). Using this information, the speculated operations underlying the serial counting explanation of subitizing will be used to make inferences about the functional significance of an earlier component referred to throughout this thesis as the N2, an anterior negative going component which may be a useful indicator of processing load in WM. Finally, the behaviour of this component will be used to make inferences regarding the operation of WM during subitizing.

2.1.3 Type of Item Enumeration Task

The item enumeration task used throughout this thesis will be a *response contingent task* using simple dot stimuli. Dot stimuli have been chosen because these very simple stimuli have no semantic associates and thus little likelihood of initiating extraneous activation operations being in WM through spreading activation affecting LTM representations (as detailed in section 1.1.4). As such each dot will represent one single unit of information and allow information load between item levels to be precisely incremented one unit at a time. Note that this task can be construed as a variant of a simple visual search task, with the goal being to note each item. The response contingent display method has been chosen over tachistoscopic presentation for four reasons related to using ERPs as the prime dependent variable: (a) it precludes the use of *estimation processes* as item load increases into the early counting range, and therefore only the processes of subitizing, and counting will be reflected by ERP behaviour, thus providing the greatest possibility of observing clearly interpretable activity reflecting any transitional phase between the two; (b) the P3b ERP component is known to be sensitive to stimulus evaluation time (Donchin & Coles, 1988; Magliero et al., 1984) and response contingent displays provide the greatest opportunity for evaluation of the actual stimuli to occur, rather than adding the confound of possible individual differences in vWM encoding efficiency (Piazza, Fumarola, Chinello, & Melcher, 2011), or associated iconic storage differences which tachistoscopic displays could introduce; (c) peak latency of the P3b is known to show *sensitivity* to the same task demands that delay RT (Verleger, 1997), so requiring a speeded response will allow this relationship to be examined; and (d) the time course of RT performance in the subitizing range through transition to counting (perhaps 400-700 ms) fits nicely with the time course of the ERP components of primary interest, with the P3b peak, signaling end of stimulus evaluation, tending to fall between 300-600 ms post-stimulus.

2.1.4 ERPs and Item Enumeration

There have been only a handful of ERP studies that have included, or focused on, item enumeration, and none of these have used response contingent displays. For the moment these studies will be introduced and the ERP components that were identified in them as being associated with item enumeration in their paradigms will be noted. Further details regarding the possible functional significance of these components will be expanded on in following sections.

One of these studies, by Xu and Liu (2008; the behavioural data from this study was introduced in section 1.2.4), examined attentional involvement in subitizing in an RSVP paradigm, incorporating single- and dual-task conditions, with 50 ms presentations with 40 ms ISI, and arrays of from 1-6 dots appearing at the end of each series. The arrays were followed by 2 successive 100 ms masks, and then a 200 ms blank screen prior to a prompt requesting response. This design has a problem in that the ERPs generated by the requirement to enumerate the dot array were overlapped by signals generated by the preceding and following screens. Thus, some end of series presentations were of a blank screen, and the ERPs generated by this display were subtracted from those generated to the dot arrays to provide what the authors (p.141) termed the "… pure dots-evoked ERP components". The ERPs were then collapsed across the 'subitizing range' of 1-3 items, and the 'counting range' of 4-6 items. A P3 component was identified and analyses were conducted on its amplitude and latency data at a single site, Pz. In the dual-task but not the single-task condition P3 amplitude was reduced, and peak latency increased for the counting range relative to the subitizing range. The authors reasoned that when WM was taxed with the dual task requirements stimulus evaluation was prolonged (indicated by the increased P3 latency), and processing load was increased when item load increased (as will be detailed shortly, P3 amplitude decreases under conditions of heightened processing load). Because the P3 data was collapsed across item levels within the subitizing, and counting ranges, no data is available from this study about possible differences in P3 behaviour as item level increased one item at a time. However, examination of the waveforms for the subitizing range shows that when the attentional manipulations of the RSVP paradigm would have been taxing attentional resources the most, during the dual-task condition, there was a relative amplitude reduction in the 200-400 ms window preceding the P3 peak. This negative deflection, while not discussed explicitly by the authors, is an important observation in terms of the information presented in the following sections regarding the possible functional importance of negative deflections such as the N2 in this window as indicators of attentional load on WM. Unfortunately, no data from fronto-central sites was presented, so distribution of the component cannot be established.

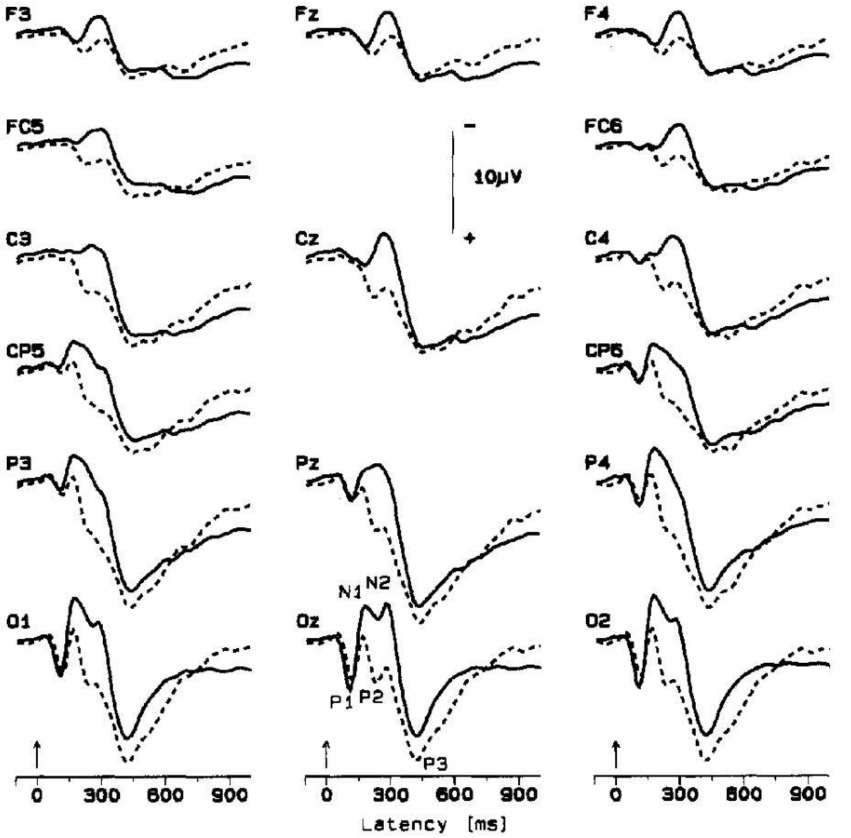


Figure 2.01. Grand average ERP waveforms for people with schizophrenia (dotted lines) and controls (solid lines) for right visual field, averaged across conditions (Negative is plotted up; From Bruder et al., 1998).

In another study people with schizophrenia and matched controls were given a dot enumeration task with 180 ms tachistoscopic displays (Bruder et al., 1998), comprising 2-6 dots presented to either right or left visual field along with a centrally presented digit which either matched or didn't match the numerosity of the array Waveforms from this study are presented in Figure 2.01. Subjects had to make a judgement as to whether the number of dots matched the digit. The authors were interested in the time course of visuospatial processing deficits noted in schizophrenic patients, in terms of claims of WM deficits (e.g., Cameron et al., 2003; Fleming et al., 1997; McGrath, Chapple, & Wright, 2001; Spitzer, 1993). They reasoned that if this deficit was in early stimulus driven attentional allocation, rather than later WM operations, then differences between patients and controls would manifest in early sensory driven components such as the visual N100. However, if it was driven by allocation of conceptual resources to stimulus classification it should manifest in a later N2 component, and if it was in later, final evaluation processes it should manifest in the P3 component.

Significantly poorer enumeration performance by the patients was reflected by notably reduced N2 amplitude in the window 200-400 ms. In fact, as is evident from Figure 2.01, patients barely produced an N2, while the component was large for the controls. On the basis of previous work they cite by Knight (1984) and Rabinowicz (1996) the authors reasoned that the controls applied better recognition and comparison processes than the patients, a supportable claim, and that this was reflected by the large N2 differences. As in the Xu and Liu study, enumeration produced a P3 component preceded by this negative deflection in the 200-400 ms window, which was maximum at fronto-central sites. However, again data was collapsed across conditions so it is not possible to assess differences in N2 or P3 behaviour as item level increased one item at a time.

A third study presented tachistoscopic arrays of 1-6 rectangles (targets) and either zero, the same, or double the number of circles, which were distractors, for 300 ms, after which subjects were asked to make a parity judgement (Nan, Knösche, & Luo, 2006). Again, these authors collapsed their data across items 1-3, as the subitizing range, and items 4-6 as the counting range so unfortunately it is not possible to assess clearly any item by item changes as item load increased. However, they did plot N2 amplitude, and perform pairwise comparisons for the zero, and same number distractor conditions. They found significant amplitude reductions in the N2 window (200-400 ms), that is, greater N2, for 4 items relative to each item in the subitizing range of 1-3, an effect that was maximum at fronto-central sites. They also identified a P3 component and found that as target number increased P3 amplitude decreased, as noted above for the Xu and Liu study.

Finally, Kamijo and Takeda (2009) presented younger and older adults 200 ms tachistoscopic displays of from one to four items and required them to respond by pressing one of four buttons corresponding to the appropriate numerosity. Both younger and older participants generated waveforms with a P2-N2-P3 morphology. For both groups the P2 became more pronounced at three and four items. However, the subsequent N2 showed no amplitude changes across item levels for the younger participants, while there was a significant increase at two, and at three items for the older participants. Kamijo and Takeda suggested that this reflected an age related failure in automatic processing during subitizing that meant the older participants had to utilise WM related processes to achieve enumeration. The authors also assessed P3 amplitude, separately at frontal, and at parietal sites. Unlike the other enumeration studies, there was no reduction in P3 amplitude across the four item range. Instead they found increased amplitudes for four items relative to one, two, or three at frontal sites, and steadily increasing amplitude across item levels at parietal sites.

Three of these four ERP studies that examined item enumeration collapsed data across items 1-3, in the subitizing range, and so any changes in the ERP components across item levels during subitizing were obscured. This approach was most probably based on the common assumption that subitizing is a qualitatively different process than counting—just the assumption that the research presented in this thesis is addressing. However, even though these studies did obscure any differences between conditions within the subitizing range both the Xu and the Nan studies did find differences in component behaviour between the subitizing range and the counting range—that is, between 3 or less items, and 4 or more—and these differences involved increased negative going amplitude of a component in a window 200-400 ms post-stimulus, referred to so far as an N2, and decreased amplitude of a following P3 component. The Bruder study was not looking at enumeration per se, but at between group differences in visuospatial tasks related to differences in allocation of conceptual resources that may have been related to differences in WM functionality between patients and controls, and they also identified similar N2 and P3 ERP components being generated in response to a dot enumeration task. Finally, the Kamijo and Takeda study did provide data from each item level of their four level study, which also showed the presence of an N2-P3 complex where N2 amplitudes became more negative going with increasing item level. However, there was no P3 amplitude reduction in this study. Instead, P3 amplitude increased for four items, the opposite to its behaviour in the other three studies. Additionally, this study clearly identified a P2 component which preceded the N2, and which increased in amplitude within the subitizing range. Even though the item enumeration tasks used in these studies differed considerably—one being embedded in an RSVP presentation, another requiring same / different numerosity judgement, while the third included disjunctive distractors and required a parity judgement—they consistently generated a fronto-central N2-P3 ERP component structure where the N2 became more pronounced as item level increased within the subitizing range. Additionally, the Kamijo and Takeda study also examined a P2 component preceding the N2, and found that it also responded to increasing item load in the subitizing range. Together with the association of changes in the presentation of the N2 component related to differences in WM functionality in the Bruder et al. study these findings provide a useful point of departure for developing a rationale about ERP behaviour during item enumeration which might indicate whether serial WM activity is implicated in determining the characteristics of subitizing—the focus of this research program.

2.1.5 The Late Positive Complex

The N2-P3 complex of ERPs forms part of what for the remainder of this thesis will be termed the *Late Positive Complex* (LPC). In terms of the trial based presentation typical of experiments which examine cognition the ERP waveform generated in response to a trial presentation can be seen as representing the unfolding of a *processing episode* where very early modulations in the waveform represent sensory processing (e.g., the P50; Jerger, Biggins, & Fein, 1992) and may be elicited independent of any processing activities required of the subject for task performance. These are known as exogenous components. The following modulations in the waveform reflect processing activities which become more related to task specific processing requirements as the episode unfolds—endogenous components (Coles & Rugg, 1995). These two types of components can be seen as reflecting bottom-up, and top-down processing respectively. While there is no definitive transition point between exogenous and endogenous components, from perhaps 100 ms on the influences of selective attention begin to emerge (Coles & Rugg, 1995; Luck, 1995a;

Rugg & Coles, 1995a), and by 200 ms post-stimulus task relevant information has been shown clearly to be influencing the characteristics of the ERP waveform (Bokura, Yamaguchi, & Kobayashi, 2001; Bruin, Wijers, & van Staveren, 2001; Corbetta, Shulman, Miezin, & Petersen, 1995). It is the collection of endogenous ERP components which unfold in temporal sequence beginning from perhaps 100-150 ms that constitutes the LPC. The dominant feature of the LPC is the large positive deflection peaking somewhere between 300-700 ms post-stimulus that has already been introduced as the P3. In its simplest presentation the P3 is generated from any stimulus that deviates from some template or series, and in this simplest of eliciting paradigm is a unitary component (Chapman & Bragdon, 1964). Once extra complexity is folded into the processing requirements of the eliciting task we start to see the emergence of recognisable endogenous components preceding the P3 (Azizian, Freitas, Parvaz, & Squires, 2006; Christensen, Ivkovich, & Drake, 2001; Dien et al., 2004), an example being the classic N2, which forms the N2-P3 complex already mentioned as being evident during item enumeration. The change in *morphology* reflected by appearance of the N2 component is often accompanied by a positive going peak which precedes the N2 (Makeig et al., 1999; Potts, 2004), and which will be referred to from hereon as the P2.

The primary inferential tool used in this thesis will be the characteristics of the morphology of the LPC to each additional item within the subitizing range, and beyond as the transition from subitizing to counting occurs. The following sections outline the key features of the ERP components comprising the LPC—the P3b, the N2, and the P2—and present predictions as to what their behaviour during subitizing might be if the serial counting explanation of subitizing is to be supported.

2.1.5.1 The P3b

The P3 is the most widely studied ERP component. As introduced in section 1.3 it is generally accepted that there are at least two types of P3, the P3a and the P3b, differentiated by their eliciting conditions and predominant scalp distribution. The 'novelty P3', or P3a refers to the P3 component which is elicited by events about which the subject has not been instructed prior to the experiment, and often occurs in response to stimuli which *deviate* from a common pattern—a high-pitch tone in a stream of low-pitched tones, a capital letter in a stream of lower-case letters, or a red symbol in a stream of blue symbols, for example. Thus, it is speculated that it reflects an orienting response as attention is captured by possibly meaningful changes in the environment (Friedman, Cycowicz, & Gaeta, 2001). While the component is evident over a wide area, the P3a tends to have maximum amplitude at fronto-central sites. In contrast, the P3b, while also being wide spread, tends to have maximum amplitude over centro-parietal sites, and is elicited under conditions where the subject has been instructed to attend to certain features of stimulus presentation and to perform some sort of task or tasks in relation to the stimuli, which lead to a response requirement which may be overt, such as a speeded button press or verbal response, or covert such as keeping a running count (Fabiani et al., 1987; Kok, 2001; Rugg & Coles, 1995a; Verleger, 1997). Thus, the component is generated in response to intentional processing demands and variations in its presenting properties have been used to examine intentional processing dynamics across a huge range of circumstances (e.g., depression, schizophrenia, anxiety) and phenomena (e.g., prepulse inhibition, face processing, somatosensory response, numerical processing).

Whether display is tachistoscopic or response contingent the item enumeration task demands intentional engagement with the stimuli in terms of engaging in quantification operations of some sort, followed by a response. Therefore it is likely that the P3 component evident in each of the tachistoscopic item enumeration studies reviewed in section 2.1.4 can be identified as the P3b, and that any P3 component generated from the response contingent displays used throughout the research program reported in this thesis will be a P3b. Thus, it is the properties of this component which will be outlined in the following sections.

2.1.5.1.1 Latency and Sensitivity to RT

One view on the functional significance of the P3b is that it reflects processes related to *stimulus evaluation* and *context updating in WM* (Donchin & Coles, 1988; Fabiani et al., 1987), with the idea that the *peak latency* of the P3b can be used to infer information regarding the time course of evaluation, or of the processing operations necessary to perform the immediate task at hand. A number of task properties have been shown to prolong P3b latency including increased difficulty of visual search (Hoffman, Simons, & Houck, 1983), increased memory load (Brumaghim, Elorman, Strauss, Lewine, & Goldstein, 1987; A. F. Kramer, Strayer, & Buckley, 1991), increased demands on semantic memory access (Bentin & McCarthy, 1994), difficulty of categorisation (McCarthy & Donchin, 1981), increased task complexity (Ragot, 1984), and instructions where accuracy is stressed over speed of response (Falkenstein, Hohnsbein, & Hoormann, 1994). Particularly when accuracy is stressed P3b latency increases are highly correlated with associated RT increases (M. Kutas, McCarthy, & Donchin, 1977). Importantly, all of these effects on P3b latency are evident when the task allows swift responses to be made, that is where RT is similar in magnitude to the P3b peak latency. However, as RTs stretch substantially beyond the timing of the peak latency, task demands that have been associated with increases in P3b latency cease to show such a relationship, most probably because the processing required to fully evaluate the stimuli in terms of task requirements can no longer occur within one episode of *perceptual closure* (Verleger, 1997), or *cognitive epoch* (Donchin & Coles, 1988), or what has been referred to so far in this thesis as a *processing episode*. Conversely, it should be noted here that even when accuracy is stressed, when task demands are simple and so errors are unlikely it is quite common for RT to be shorter than P3B peak latency, perhaps because of ease of information transmission (Johnson, 1988), or because for simple tasks evaluation and response selection may occur in parallel, or response may be selected on the basis of partial processing of the stimulus information (Pfefferbaum, Christensen, Ford, & Kopell, 1986).

At this point it is important to make explicit the relationship between the P3b and WM proposed in the context updating model. A key claim underpinning this model is that " … the process underlying the P300 is involved in the creation and maintenance of representations" (Donchin & Coles, 1988, p.373). The authors distinctly situate this process in an information processing framework akin to Baddeley's WM model that was introduced in section 1.1.1.4 (Baddeley, 1986), and make clear reference that the evidence supporting their model suggests that the P3 is related to this *capacity limited process*. In terms of the concept of a processing episode the limitations on the complexity of context updating that can take place in a short amount of time (what equates to the duration of the LPC, with the P3b peak as the endpoint) reflect the limitations of the WM system, notably in terms of the idea of *chunking capacity* that was introduced in section 1.1.2. When the binding and maintenance operations required to update context—such as arrive at the appropriate output to achieve task completion—exceed the immediate capacity of activation available to the system the remainder of the processing must wait until resources are available to be reallocated to it, thus processing becomes ongoing (often reflected by slow wave activity; Ruchkin, Canoune, Johnson, & Ritter, 1995).

This proposed relationship between WM and the limitation in capacity to update context in the course of a single processing episode—chunking capacity—is neatly illuminated by looking, in terms of the activation model of WM, at the types of task demands that prolong P3b latency. Increasing the difficulty of each of the task properties mentioned earlier can be seen to equate to some increase in the number of representations needed to be held active, or in the number or complexity of binding operations required to perform the task. Difficult visual search requires effortful comparison to search parameters for accurate discrimination (Atchley, Jones, & Hoffman, 2003; Luck, 1995a; Treisman & Sato, 1990); categorisation requires memory search and comparison to exemplar characteristics (Axel Mecklinger, Kramer, & Strayer, 1992; Walter Ritter, Simson, & Vaughan, 1983; Salisbury, 2004); extra memory load directly recruits additional activation resources, as does increased demand for semantic memory access (Conway & Engle, 1994; Kane & Engle, 2000); task complexity means keeping active more elaborate rules required to perform the task, and assessing when to implement them (A. Kramer, Schneider, Fisk, & Donchin, 1986; Van der Lubbe & Verleger, 2002); and focus on accuracy recruits resources to monitor performance and ensure that guessing, or heuristics, are avoided (Christensen et al., 2001). Thus, operations that directly tax WM activation resources are the same sorts of operations that have been shown to prolong P3b latency, and many of these manipulations also result in increased time to complete the tasks.

If complexity of stimulus evaluation and maintaining and creating representations in WM influences both P3b latency and RT the obvious question to address is whether there is a relationship between changes in P3b latency and in RT in the sorts of tasks where response can be generated in roughly the same time as it takes the P3b to peak—that is, within the chunking capacity of the system. While early conceptions of the stimulus evaluation model made some distinction between evaluation processes and those involved in response selection, the subsequent body of evidence has suggested that evaluation time, as indexed by P3b latency, can indeed be highly related to variations in RT, presumably because sufficient stimulus evaluation is required for subsequent appropriate response selection (M. Kutas et al., 1977; Verleger, 1997). Thus, P3b latency can show considerable *sensitivity* to the same types of task demands which increase RT—that is, increases in task difficulty, or complexity. In a review of P3b sensitivity to the task manipulations which prolong RT (Verleger, 1997) it was noted that there is some degree of 'slop' in the comparisons, which are computed as percentage scores using the formula: *sensitivity = delay in P3b / delay in RT * 100*. Because of the assumption that this value reflects an indicative rather than absolute relationship, Verleger proposed three categories of sensitivity: 67% to ≥100% = High; 33-66%= Medium; 0-32% = Low. Given (a) that the item enumeration task can be construed as a simple variant of a visual search task (as noted in section 1.3.3); and (b) that memory scanning has been proposed as one explanation for the subitizing slope (as noted in section 1.2.2), P3b sensitivity to RT in visual search, and in Sternberg tasks is of interest. Latency in visual search tasks exhibited a very high sensitivity to changes in RT across differing difficulty levels (79% median values across the assessed data sets), while a large body of work assessing the Sternberg task (60 studies) showed a median sensitivity of 43%. In contrast, tasks manipulating spatially incompatible stimulus-response options (e.g., respond *left* to the word *right*) showed a low median sensitivity of 11%, and those evoking mental rotation showed no sensitivity.

To summarise briefly: P3b latency tends to become longer as task properties increase demands on activation and binding operations in WM, so long as a speeded response is required as part of the task structure, and that response can be generated roughly within a single processing episode, the time frame of which is defined by the endpoint of the LPC which is the P3b peak. In some paradigms, but not in others, P3b peak latency can be sensitive to the same task manipulations which prolong RT, and the two most relevant paradigms to the item enumeration task as presented in this research program—visual search, and the Sternberg task—show high and medium sensitivity respectively.

2.1.5.1.2 Amplitude

While P3b peak latency behaviour has been used as an effective index of duration of processing, task elicited changes in P3b amplitude have been used to make inferences about resource allocation, and processing demands (Fabiani et al., 1987; Johnson, 1986, 1988; Rugg & Coles, 1995b). A range of different task manipulations influence P3b amplitude. For example, in the classic ‘oddball’ task where a stream of two stimulus categories are presented with one being less common than the other (the ‘oddball) and the subject is instructed to attend to the rare stimuli and either respond to them (perhaps by button press) or to silently maintain a running tally, P3b amplitude to the oddball category is much greater than to the common stimulus. Thus, stimulus probability, and task relevance are two properties which influence P3b amplitude (Donchin & Coles, 1988), with improbable, relevant stimuli increasing amplitude. In this sense, the degree of attention allocated to a stimulus (irrespective of its probability) can be seen as a powerful determinant of increased amplitude, and in simple tasks greater P3b amplitude has been used effectively to infer greater attentional engagement.

An important point here is that the attentional deployment indicated by increases in P3b amplitude (in fact, by the very generation of a P3b) reflects processing the *meaning* of the evoking stimulus. This meaning is conveyed by the stimulus itself, but only in the context of task relevance. The exact same stimulus can have multiple meanings in terms of its relationship to changes in the current context (e.g., task instructions; surrounding, or preceding stimuli), and the relative difficulty in resolving this relationship can obviously influence the extent to, or ease with which meaning can be extracted (ascertained) from the stimulus (Johnson, 1988). As the relationship of the stimulus to the context becomes more complex, that is, the task becomes more difficult, even though more attentional resources must likely be deployed to resolve the meaning of the stimulus (i.e., meet task requirements) P3b amplitude tends to *decrease*. So, task difficulty is another important influence on P3b amplitude, and at first glance the inverse relationship between task difficulty and P3b amplitude seems at odds with the opposite relationship seen in simpler paradigms of positive variation between amplitude and allocation of attentional resources.

Decreased P3b amplitude in response to increased task difficulty has been illustrated over a range of paradigms. During visual search simply increasing perceptual load sufficiently has been shown to reduce amplitude (Lorist, Snel, Kok, & Mulder, 1996), as has increasing memory load (Hoffman et al., 1983), or requiring varied (versus consistent) mapping (A. Kramer et al., 1986). Increasing the complexity of conceptual operations (Wijers, Otten, Feenstra, Mulder, & Mulder, 1989), and some dual task manipulations such as target tracking with covert counting of one class of two classes of sporadic auditory or visual stimuli (Isreal, Chesney, Wickens, & Donchin, 1980; Wickens, Kramer, Vanasse, & Donchin, 1983) decrease P3b amplitude. Additionally, dual-tasks with a secondary task focus on memory search and memory load (Strayer & Kramer, 1990; Wijers, Otten et al., 1989) also show this same strong pattern of decreased P3b amplitude. The increasing difficulty of these tasks is confirmed by increasing RT across difficulty conditions. Invariably the P3b amplitude reductions seen in these difficult tasks are accompanied by increases in P3b latency, as the stimuli become more difficult to categorise in terms of the task context, and so stimulus evaluation is prolonged. Of key importance for this thesis is the observation that paradigms requiring complex visual search, heavy memory load, memory search, and generally more complex conceptual operations obviously place greater demands on WM resources, and these paradigms generate the combination of prolonged P3b latency, and reduced amplitude.

Two possible explanations of this pattern of amplitude reduction and latency increase when processing load increases and WM resources might become taxed are, firstly that the large amplitude, unitary P3b peak generated under less complex processing conditions is actually comprised of more than one positive ERP component, and these components reflect at least two cognitive processes. These processes either co-occur, or operate in close temporal sequence when processing demands on WM resources are low. When demands tax WM resources these processes devolve, with the first one staying static while the second process is prolonged (or begins later) and so ‘moves forward’ in time. Thus, under low processing load amplitude of the ERPs reflecting the two processes summates, while under high processing load the two processes do not co-occur and so the ERPs do not summate, leading to reduced amplitude (Brookhuis et al., 1981). Indeed, multiple P3 peaks have been identified when task difficulty has been increased (Christensen et al., 2001), and the temporal distance between the peaks has increased when task demands have slowed responding, a situation also accompanied by reduced amplitude over the latter P3 window. A difficulty with this explanation is found in those studies where amplitude decreases, latency increases, but there also appears a fundamental change in the morphology of the *earlier* part of the LPC, with the appearance of the N2 component already highlighted as part of the common N2-P3 complex in sections 2.1.4 and 2.1.5 (Christensen et al., 2001; Isreal et al., 1980; Strayer & Kramer, 1990; Wickens et al., 1983; Wijers, Mulder, Okita, Mulder, & Scheffers, 1989; Wijers, Otten et al., 1989).

This circumstance is relevant to a second explanation, which suggests that high WM demands are reflected by negative going slow potential ERPs which co-occur with the positive going wave comprising the P3b, and so reduce its amplitude (Falkenstein, Hoormann, Hohnsbein, & Kleinsorge, 2003; Rösler, Heil, Bajric, Pauls, & Hennighausen, 1995; Rösler, Heil, & Glowalla, 1993; Wijers, Mulder, Okita, Mulder et al., 1989; Wijers, Otten et al., 1989). Slow, negative waves have been commonly identified in a range of studies addressing WM *retention*. Indeed, many ERP studies overtly claiming to look at WM have focused on examining slow wave activity that is generated robustly during *maintenance* tasks (Cameron et al., 2003; Deldin, Deveney, Kim, Casas, & Best, 2001; Deveney & Deldin, 2004; García-Larrea & Cézanne-Bert, 1998; Geffen et al., 1997; Axel Mecklinger & Müller, 1996; A. Mecklinger & Pfeifer, 1996; Rösler & Heil, 1991; Ruchkin et al., 1995; Schubotz & Friederici, 1997; Smolnik, Perras, Mölle, Fehm, & Born, 2000; Vos, Gunter, Kolk, & Mulder, 2001). However, while these studies have demonstrated that effortful maintenance of information, often (but not always) concurrent with some primary processing task, is reflected in the EEG by these slow negative potentials, this approach is not relevant to the focus on WM in this thesis. Indeed, this approach reflects an oft seen conception of WM as being predominantly a maintenance operation, which is possibly a hangover of older days when STM was a focus of information processing. In these studies some sort of maintenance over time is required and differences in slow wave activity (perhaps extending over some seconds) between conditions, or between groups, are examined. This conception of WM is not consistent with the time course over which subitizing occurs, nor with the dynamic activation and binding activities posited by the serial counting explanation as occurring over perhaps 500-600 ms—the course of a single processing episode—and as possibly underlying the characteristics of the subitizing slope. However, in some studies using tasks that can be seen to tax WM and that *do* conform to this time course, emergence of the N2-P3 complex has been identified, consistent with those components found in the enumeration studies.

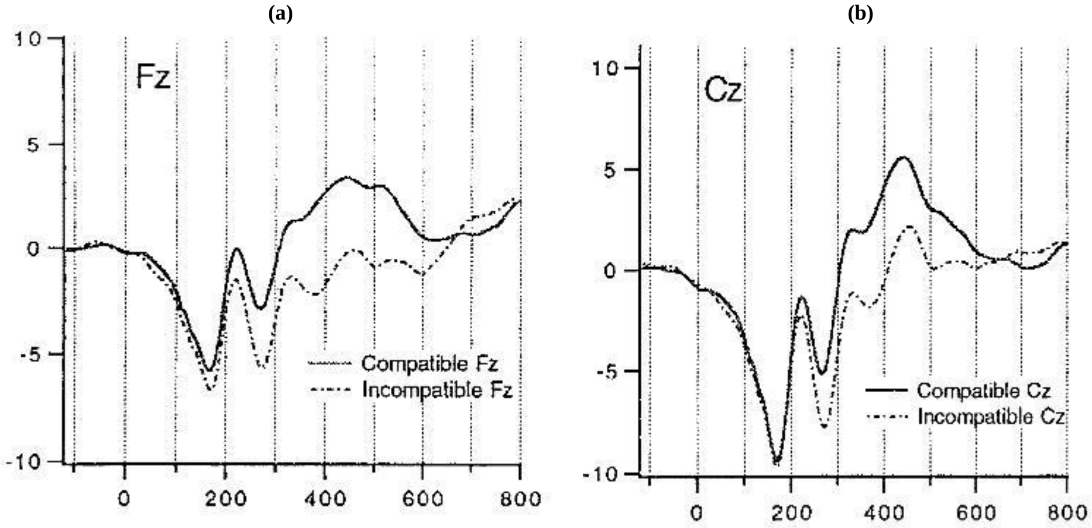


Figure 2.02. Waveforms at Fz and Cz illustrating greater N2 (200-400 ms) followed by reduced amplitude for the LPC (including P3b peak) for difficult (Incompatible) versus less difficult (Compatible) conditions in a stimulus-response compatibility task (negative is plotted down; From Christensen, Ivkovitch, & Drake, 2001).

The key issue here is that these two explanations of P3b amplitude reduction under processing load—separation of two previously co-occuring, summating positive peaks (processes), or concurrent activity of a negative going wave reflecting WM activity—are not necessarily mutually exclusive. In fact, over a range of studies where processing load increases and conceivably taxes WM resources *over the course of a single processing episode* (e.g., Christensen et al., 2001; Isreal et al., 1980; Strayer & Kramer, 1990; Wickens et al., 1983; Wijers, Mulder, Okita, Mulder et al., 1989; Wijers, Otten et al., 1989) it is common to observe a waveform morphology consistent with that described in section 2.1.5 as reflecting the devolution of the unitary P3b into the LPC, where an early positive going peak, the P2, and the later P3b, are interceded by a negative going deflection in the window 200-400 ms post-stimulus, that is, an N2, with magnitude of the N2 related to amplitude decrease for the subsequent portions of the LPC including P3b peak. This pattern is illustrated in Figure 2.02.

The N2 component (notably the *fronto-central* N2) may reflect functional behaviour associated with load dependent WM activity. Details suggesting that this may be the case are outlined in the following section, 2.1.7. In this explanation the separation of the processes generating the positive peaks is determined by intercession of the negative going, processing related activity. In terms of the activation model, this possible expression of WM load related activity could reflect extra binding operations required for effective context updating, and the duration of the processes accomplishing these operations thus delays context updating and so prolongs P3b peak latency. So by this argument formation of the LPC by *devolution* of the P3b into P2, and P3b peaks interceded by the negative deflection would be a reflection of load related taxing of WM resources that results in prolonged (serial) processing prior to context updating being achieved.

This position is one which the research program presented in this thesis addresses. In theory it is supported by the possible functional significance of a class of *anterior negativities*, which include the ‘generic’ N2 so far referred to, that occur in a range of paradigms including those that demand the processing of response selection rules, conflict monitoring, object identification, or memory search, and which are most often also followed by a late positive going deflection that could be seen as a P3b. Most often these negative going deflections are also preceded by a positive peak, which has been referred to so far in this Chapter as a P2, and which in fact defines the negative deflection that follows it—the anterior negativities exist in contrast to this peak which precedes them but in many of the studies addressing N2 like deflections little or nothing is said about the preceding P2 (see Figure 2.02: the positive peak at around 220 ms is where the negative going activity of the N2 begins. The positive-negative modulation defines both components in the morphology of the LPC and is an important point of discussion shortly in section 2.1.8). The following section provides details concerning these anterior negativities that fall in the N2 window, after which the P2—and it relationship with the N2—is discussed.

2.1.5.2 Fronto-central Negativities: The N2

Reference to the N2, or N200 in the ERP literature refers to components which are generated from different paradigms (Folstein & Van Petten, 2008) but occur in roughly the same time window (200-400 ms post-stimulus). As mentioned earlier, nomenclature in the ERP literature is varied, and this is the case for the N2 component as well. Several subcomponents of the N2 have been proposed, varying in their topography and eliciting conditions (Pritchard, Shappell, & Brandt, 1991). Examples are the fronto-central N2a, or mismatch negativity (MMN), which might index automatic detection of novelty or deviance from prevailing context, and does not require attention to the eliciting stimulus to be elicited (Näätänen, 2001); the N2c, a posterior component which might index selection of a target item and requires attention to the eliciting stimulus to occur (Luck, 1995b); the N2b, or Go/NoGo N2, which occurs in a range of circumstances involving conflict monitoring or the requirement for response inhibition and also requires attention to the eliciting stimuli to occur (Kok, 1986), and has a fronto-central distribution (Deacon, Breton, Ritter, & Vaughan Jr., 1991; Patel & Azzam, 2005); or the frontal processing negativity, a fronto-central component which might reflect complex task relevant feature integration, and which also requires active attention to the eliciting stimulus (Karayanidis & Michie, 1997). Given (a) the appearance of a fronto-central N2 in the few item enumeration studies reviewed above, (b) the association of fronto-central negativities with processing related to active task demands, and (c) the active attentional deployment inherent in the response contingent item enumeration task used in the research reported in this thesis, any N2 seen in relation to this task is likely to be more functionally similar to the N2b, or the frontal processing negativity than to the posterior selection related negativities, or the automatic MMN. Thus, the remainder of this section presents information regarding the eliciting circumstances and behaviour of fronto-central negativities in tasks that demand intentional attentional deployment.

Fronto-central N2 components referred to in the literature have been generated from a range of paradigms manipulating variables such as stimulus unfamiliarity, or encoding difficulty (Daffner et al., 2000; Nittono, Shibuya, & Hori, 2007), modal attention (Eimer & Schroeger, 1998), ease of visual object identification and memory search (e.g., Schendan & Kutas, 2002; Schendan & Kutas, 2003), response conflict and response inhibition in stroop (Tillman & Wiens, 2011), flanker (Kopp, Rist, &

Mattler, 1996), and stop-signal tasks (Kok, Ramautar, De Ruiter, Band, & Ridderinkhof, 2004), as well as the types of information that need to be processed for appropriate response selection in Go/NoGo paradigms (Bokura et al., 2001; Bruin et al., 2001; Fox, Michie, Wynne, & Maybery, 2000; Jonkman, Lansbergen, & Stauder, 2003; Walter Ritter et al., 1983; Walter Ritter, Simson, Vaughan, & Macht, 1982; Rodriguez-Fornells, Schmitt, Kutas, & Münte, 2002; Schmitt, Rodriguez-Fornells, Kutas, & Münte, 2001; Wijers, Mulder, Okita, & Mulder, 1989), and selective attention in selective search tasks (Wijers, Mulder, Okita, & Mulder, 1989). A key presenting feature of this fronto-central N2 is that relative amplitude *decreases* (i.e., becomes more negative going) with complexity of information required for discrimination, and it invariably is followed by a P3 component, such that early examinations of it were often done in terms of an ‘N2-P3’ complex. This sets the N2 like components apart from other fronto-central negativities such as the N400 which often have a later onset, but also a longer duration and often appear in the absence of a following P3 (Marta Kutas & Federmeier, 2009).

The generation of N2 like components in diverse kinds of tasks has led to speculation over what sort of functions the component(s) might reflect (e.g., Deacon et al., 1991). Robust presentation of a component referred to as the N2b in Go/NoGo paradigms has seen it commonly termed the Go/NoGo N2, amidst claims that it reflects the operation of processes of response inhibition (Kok, 1986), but there is evidence to broaden this interpretation at least to one of conflict monitoring. In a Go/NoGo paradigm a strict reflection of response inhibition would see the N2 occur only in conditions where a response is clearly to be withheld, but when sequence effects are introduced the component is evident also when expectancies are violated even when no response inhibition is required (Smith, Smith, Provost, & Heathcote, 2009). Additionally, in conditions where some non-targets have varying degrees of perceptual overlap with targets, larger N2 components have been generated to the perceptually similar non-targets than to the targets, suggesting that key task relevant features of the stimuli have been processed and have generated an imperative towards action that conflicted with the appropriate NoGo response (Azizian et al., 2006). Interestingly, these have been of greater amplitude than those generated to the actual targets, suggesting greater processing effort was required to eventually discriminate them from targets. Others have found results agreeing that this monitoring may include active selection of complex task relevant features required for response rule implementation (Fox et al., 2000), which may include rule assessment and then transmission of the outcome to later processes to implement response inhibition, or not (Bruin et al., 2001). These results suggest a level of executive control—that is, WM activity—involved with the processes reflected by the fronto-central N2, a conception that is strengthened when other manipulations of the Go/NoGo paradigm are examined.

The Go/NoGo paradigm has been used to vary the nature of the information that is necessary to either respond or withhold response depending on the class of stimuli displayed, and thus assess the time course of availability of different types of information. The logic behind this approach is that, firstly, so long as the appropriate response is emitted (or withheld) any N2 generated to a class of information indicates that that information was available to be processed *and was processed satisfactorily*. Secondly, differences between different classes of information in the onset latency, peak latency, or amplitude of the N2 indicate differences in terms of *when in the processing stream* it becomes available, *how long* processing may take, and the relative *intensity of processing*. For example, Rodriguez-Fornells et al. (2002) showed subjects pictures or words depicting common animals or objects and required them to respond, or withhold response, on the basis of either semantic (object or animal) or phonological (first letter being vowel, or consonant) information. At frontal electrodes NoGo trials based on the semantic information generated N2s with both onset and peak latencies up to 70 ms earlier than in the phonological condition. No amplitude differences were reported. The authors concluded that semantic information was available earlier in the processing stream than phonological information, findings replicated by Schmitt et al. (Schmitt, Münte, & Kutas, 2000). With the same methodology similar findings have emerged regarding time course of availability of semantic (~80 ms earlier) versus syntactic information, contrasting animal/object classifications with syntactic gender information about the names of the animals, and objects (in German, which has an established gender syntax system; Schmitt et al., 2001).

Besides reflecting time course of availability, N2 peak latency has also been shown to increase across a range of circumstances where processing of task relevant stimuli becomes more difficult, suggesting that, as part of the LPC, it also may provide information about the relative duration of evaluative processes as a function of processing load (Folstein & Van Petten, 2008; Patel & Azzam, 2005; Pritchard et al., 1991). N2 latency has been shown to be prolonged for deviant pictorial stimuli which were difficult to classify, and thus encode, as well as to target stimuli which required attentional processing, relative to common, simple stimuli (Daffner et al., 2000); for elemental stimuli which required assessment in terms of solving the *exclusive or problem* in logic (where either element is true but their conjunction is false; e.g., A+, B+, but AB-) over simple configural stimuli in a Go/NoGo task (Fox et al., 2000); and for multiple feature conjunctions versus simple features in visual search tasks (Karayanidis & Michie, 1996, 1997), and similar findings have emerged from the auditory domain where N2 latency has been found to covary with RT in response to difficult discriminations in a vigilance task (of note, both also covaried with P3 latency; Ritter, Simson, Vaughan, & Friedman, 1979). Additionally, in stop-signal tasks unsuccessful stop trials generate longer latency N2s than successful ones, suggesting more lengthy processing that ended up being too late to inhibit the incorrect response (based on a ‘horse race’ model; Kok et al., 2004; Ramautar, Kok, & Ridderinkhof, 2004), and N2 latency has been shown to increase in aged people when conjunctions needed to be identified, probably due to decline in WM function (Patel & Azzam, 2005).

These findings show that the N2 reflects not only processes involved with response inhibition, but also with task relevant feature processing, and rule based discrimination. They show clearly that the N2 is sensitive to the availability and utilisation of task relevant information, and that both amplitude and latency can increase in response to a range of different types of task difficulty. This conclusion is also supported in other tasks that don’t require response selection that might involve inhibition. For example, in a visual search task which also directly loaded WM resources, for each block subjects had to memorise a set of either one or four letters and were shown a stream of letter pairs where letters were either small or large (equiprobable). They had to attend to one class of letter (in each block) and respond to any letter that was part of the memory set letters with a button press (Wijers, Mulder, Okita, & Mulder, 1989). N2 components were generated to target letters, and these were larger for the four letter memory set condition than when there was only a single letter memory load. There were also conditions where the attended stimuli were identified by colour, or by conjunction of colour and letter size. The largest amplitude N2s were generated to the conjunction condition with the high memory load.

Thus, in task conditions where response inhibition is required, as well as in those where it is not, task relevant information is shown to modulate properties of the N2. So, independent of whether the component indexes response inhibition processes in certain situations, response inhibition is obviously not the only cognitive processing that the N2 reflects. Where conflict with task relevant requirements is possible due to nontarget stimulus properties which are similar to those of target stimuli the N2 becomes apparent, thus illustrating processes which detect partial matching with task relevant information which, for this to occur, must be held active to be utilised in such a process. Where semantic information is critical for task performance the N2 appears earlier than it does for either phonetic or syntactic information but the component is evident for each type of information, and successful task performance indicates that within perhaps 200-250 ms of a stimulus becoming available for processing all these different types of information can be active and available for use in assessing appropriate task relevant response (i.e., parsing task rules). Greater complexity of a rule determining the correct way to respond can increase negative going amplitude of the N2, an effect which can be magnified by the extent of current memory load, and difficult discrimination can increase both amplitude and latency. Together these demonstrated properties of the fronto-central N2 indicate that the component can reflect the activation of task relevant representations, which include current task rules as well as semantic information from LTM, and the operation of control processes over these active representations in the service of correct task outcome. In terms of the activation model of WM and the range of attention, these properties clearly suggest that the fronto-central N2 component reflects integrative WM activity which is evident from ~200 ms post-stimulus onset, and therefore that amplitude and latency changes in the N2 could be informative about WM activity in this window.

This interpretation begs the question as to the functional significance of the N2 component evident in the enumeration studies reviewed in section 2.1.4, particularly the Kemijo and Takeda (2009) study where it was evident even for one item, and for the older but not the younger participants became more pronounced as item load increased. Kamijo and Takeda's suggestion that this reflected an age related failure in automatic processing during subitizing which meant that older participants had to utilise WM related processes to subitize is well supported by the information just presented regarding the N2. If it is the case that the N2 is evident during subitizing of the response contingent displays presented in the current research program then it would be plausible to interpret it as reflecting some sort of load based WM activity. Of significant interest would be any systematic increases in latency or negative going amplitude as a function of item load, which could be used to infer increased WM involvement.

2.1.5.3 The P2

When the ERP waveforms are examined from a number of studies which find N2 and P3b components in visual paradigms there is often the presence of a positive going component preceding the N2, which typically shows an onset around 100 ms and a peak somewhere around 180-220 ms post-stimulus. This *P2* component is very often ignored completely in these papers, or at best given cursory treatment with little interpretation as to what its functional significance might be, with the focus invariably being on the following N2, and / or P3 components (e.g., Azizian et al., 2006; Bartholow et al., 2005; Bokura et al., 2001; Bruin et al., 2001; Daffner et al., 2000;

Han, 2000; Jonkman et al., 2003; Nittono et al., 2007; Tillman & Wiens, 2011). Example waveforms from some of these studies are presented in Figure 2.03.

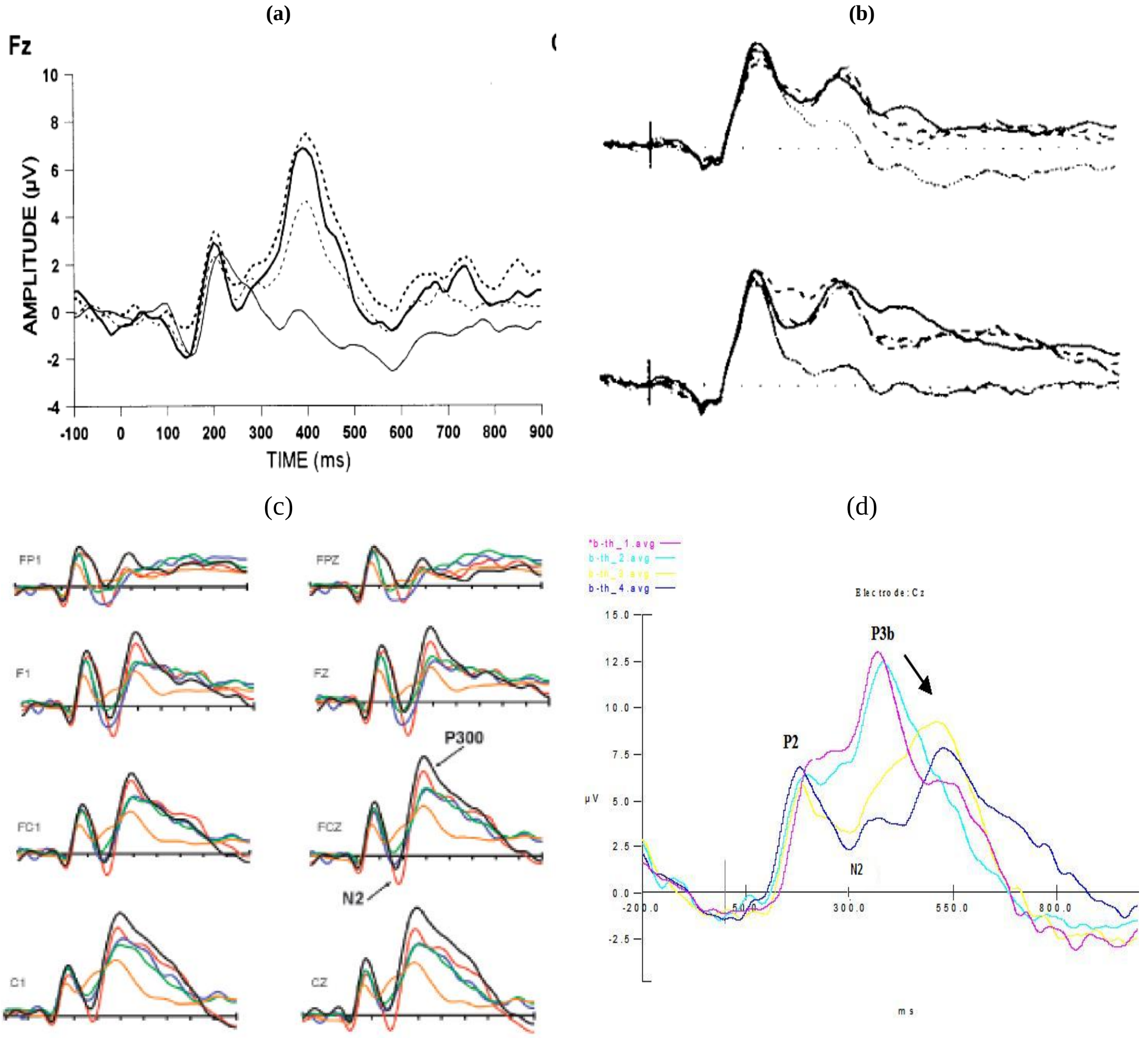


Figure 2.03. Examples of waveforms showing the P2-N2-P3b complexes comprising the LPC, from studies that have ignored the P2 (Panels (a), (b), and (c)). Panel (d) shows unitary P3b (pink trace; low load) devolving to LPC (dark blue trace; high load), comprising P2-N2-P3b. Positive is plotted up. (From (a) Bruin et al., 2001; (b) Wijers et al., 1989 (c) Azizian et al., 2006; (d) Quain, unpublished pilot data, 2004)

The waveforms in Figure 2.03 make clear the point raised at the end of section 2.1.5.1 regarding the defining relationship in the LPC between the P2 and the N2 which follows it, and panel (d) illustrates clearly the devolution of the unitary P3b into the LPC comprising the P2, the N2, and the P3b, that is described in that section, and evident in each of these other example waveforms.

While the P2 has often been ignored, where it has been examined in its own right in relation to visual stimuli it has been associated with functional significance in terms of early selective attention involving salient feature analysis (Luck, 1995a), or target detection behaviours as part of a network that manipulates task salient representations (Potts, 2004; Potts & Tucker, 2001), or in indexing selection of relevance (Wolach & Pratt, 2001). In target detection tasks P2 tends to have longer latency when object identification rather than spatial location is the response cue (Potts, 2001), or when target identification cannot be done automatically, such as in a varied target mapping condition (Luck & Hillyard, 1994), suggesting that the component indexes top-down selection processes that are resource demanding. P2 amplitude is increased to relevant targets which require some response, either covert or overt (Potts, 2004), and amplitude also increases in response to emotional salience, showing enhanced amplitude to stress related stimuli (Michalski, 1998), and to facial expressions of fear (Ashley, Vuilleumier, & Swick, 2004), which has been interpreted as attentional readiness to respond to environmental threats which thus have survival relevance. Thus, the processes that the P2 reflects take longer when identification of relevant information is made more difficult, and operate at greater intensity when the information is more relevant to immediate goals. The P2 also shows amplitude attenuation to stimuli presented during the attentional blink, but not for stimuli presented while there is a memory load. In contrast, both the N2, and P3b are attenuated by both high memory load, and the attentional blink (Akyürek, Leszczynski, & Schubo), which confirms the P2s involvement with attentional resources, but suggests that these are resourced at a selection, not a maintenance or processing stage.

Thus, the P2 can be seen as an index of selective attention that is sensitive not to just visual features but also to top down influences such as salience, and so must fall within the range of attention. In paradigms which are processing intensive it robustly co-occurs with the following N2, a component shown to be associated with indexing task relevant processing involving executive control. Therefore, in terms of the functional significance of the P2 in processing episodes reflected by the unfolding of the LPC it is plausible that the P2 might fill a role along with the N2 in early WM processing preceding higher level categorisation and context updating activities reflected by the P3b.

2.1.5.4 The P2-N2 Complex

A similar functional role for a P2-N2 complex as just described has been proposed by Missonnier et al. (2003). They suggested that a positive-negative deflection during approximately 140-280 ms post stimulus which they identified during an n-back task indexes storage uptake by WM as a precursor to higher order processing operations proceeding. Interestingly, latency of both the positive and negative peaks was *reduced* by increased WM load (a 2-back condition versus a 1-back, or no load condition), which is the opposite to the P2s behaviour in target detection tasks, where greater difficulty increased latency. This suggests that even though P2-N2 deflections are observed in a number of paradigms, interpretation of their functional significance needs to carefully take into account the eliciting circumstances. The n-back task increments WM demands by requiring maintenance and matching operations that increase in complexity with each additional step ‘back’ in the stimuli stream the target item must be maintained from. In contrast, target detection tasks do not tax maintenance demands in the same way, but do require matching operations when items need to be selected in terms of their identification, so it is likely that they place less demands on WM in terms of dual-task requirements.

It should be noted here that Missonnier et al. (2003) interpretation of the P2-N2 complex as indexing storage processes only is open to question on the basis of the review in section 2.1.5.2 of the varied circumstances that affected the N2 component. The key suggestion was that it may reflect *integrative* activity related to top-down processing and executive functions. If this is the case then Missonnier et al. interpretation of the entire P2-N2 complex as reflecting the loading of information into WM *storage* might be too simplistic, making the mistake of treating the entire complex as a discrete unit having a discrete function. In fact, the complex occupies around 200 ms of a processing episode, bridging the period from the end of sensory processing (perhaps 80-100 ms; Luck & Hillyard, 1994; Rugg & Coles, 1995a; Sagi & Julesz, 1985; Schendan & Kutas, 2003) to around 300 ms, when higher order functions such as categorisation and context updating in WM are well underway, and in some cases completed (Brookhuis, Mulder, Mulder, & Gloerich, 1983; Christensen et al., 2001; Verleger, 1997). During this lengthy period (in terms of a processing episode) there is no doubt that task relevant information would need to be ‘loaded’ into WM storage (i.e., activated). However, as processing activities proceed over those representations it is likely that additional representations such as related semantic representations and rules determining future processing priority would need to be integrated with them as progressively higher order processing proceeded until categorisation and ultimately context updating was accomplished effectively. In this sense the P2-N2 complex might reflect not simple storage related activation, but integrative processes as well. In fact, this postulation characterises *momentary* (as opposed to static, or ongoing) storage and processing in WM as a dynamic, integrative process, and not a static discrete one. Under these circumstances, the P2 component might reflect selective attention to (i.e., activation of) task relevant representations, and the N2 might reflect load related integrative processing over these representations.

In terms of the current research program, if this is the case then if there is load related WM activity occurring during subitizing we might expect this activity to be indicated by changes in the P2-N2 complex, with the key issue being the number of items displayed when any changes occur. However, it is not exactly clear what changes to the *complex* might be expected to be clear indicators of WM activity. While the review of N2 behaviour under conditions that demanded some WM involvement suggested that increases in N2 latency and amplitude would likely be indicators of increasing demands on WM resources, and the P2 also showed some latency increases to more demanding search tasks, the findings from Missonnier et al. (2003) n-back task point in the opposite direction, suggesting that latency decreases in both components are associated with greater WM load. This finding could result from some idiosyncrasy of the n-back task, notably that the task is presented as a continuous performance task (i.e., a stream of stimuli) with a constant requirement to update a memory set (one item for a 1-back condition, 2-items for a 2-back condition) as well as maintain that set active while engaging in matching operations comparing the current stimulus to the memory set. This makes it a far ‘busier’ task in terms of *maintenance* than many of the tasks which generate N2s, and certainly than the item enumeration task. However, the matching component is a simple template comparison which is unlikely to place the same type of demands on *integrative* processes as tasks with more complex decision rules for example, or those that require manipulation and integration of semantic information. Therefore it might be expected that tasks which demand more integrative activity would generate increases in N2 latencies, and amplitudes.

In terms of the item enumeration task, the serial counting explanation presents an argument (see section 1.2.3) suggesting that it is integrative demands that drive the subitizing slope, and certainly during counting (see section 1.2.1) great demands are made on integrative processing. Given the weight of evidence showing N2 increases in both latency and negative going amplitude to higher demands on WM, it is likely that this is the pattern of N2 behaviour which would indicate any increasing WM load contributing to the subitizing slope. It is, however, unclear what influence differences between the n-back task, target detection tasks, and the item enumeration task might imply for predictions concerning P2 behaviour during item enumeration. This is a question that will be answered through empirical investigation in the first experiment, reported in Chapter 4.

2.2 Summary and Progression

2.2.1 Summary

This chapter has introduced the concept of ERPs, and the components of interest which comprise the LPC—the P2, the N2, and the P3b. A key feature of the LPC is that it comes about through the *devolution of the unitary P3b*, and this occurs in response to increasing processing load, a function of task difficulty, which can be seen as reflecting increased demands on WM over the course of a processing episode. Increased task difficulty is confirmed by prolonged RT, and *when the task requirements can still be completed within a processing episode*, this prolonged RT is often accompanied by an increase in P3b peak latency which in visual search or Sternberg tasks (which have some task components that may be similar to those required for item enumeration) shows medium to high *sensitivity* to the increases in RT. These processing load related increases in P3b peak latency are accompanied by a reduction in P3b amplitude, and in demanding tasks often by the appearance of a preceding P2-N2 complex. There is a school of thought that the large amplitude of the unitary P3b for relatively simple tasks results from the summation of two closely occurring processes reflected by two positive components, perhaps the P2 and the P3b, and that under processing load the first process, and the P2, remains static while the second process, reflected by the P3b, is prolonged as context updating takes longer. Thus, the two processes no longer summate resulting in amplitude reduction over the entire latter portion of the LPC, and of the P3b peak. This load related separation of the two peaks may be influenced by the intercession of WM related activity reflected by a negative going wave, perhaps the N2, that contributes to prolonging context updating, and P3b latency, as well as to the noted amplitude reduction.

In terms of item enumeration it has been noted that there is no published information regarding ERP behaviour during response contingent displays, and it is the purpose of this research program to examine this issue. In the tachistoscopic item enumeration studies reviewed in section 2.1.4, the P3b was generated in each study, with two of them reporting an amplitude reduction between the subitizing range (1-3) and the counting range (4-6), which might be expected as processing load increased with the onset of counting. However, only one of these studies reporting an associated latency increase. In contrast, in the only study to provide item level waveform data during the subitizing range (Kamijo & Takeda, 2009) no latency data was addressed, but an increase in amplitude was reported at four items. Thus, there is insufficient information from these studies to generate a clear picture regarding P3b behaviour during subitizing, and across a transition to counting as item load increases.

The N2 component was evident during subitizing in each of the tachistoscopic item enumeration studies reviewed in section 2.1.4. Bruder et al. (1998) did not provide data per item level, but suggested that the N2 was indexing allocation of conceptual resources to stimulus classification, and related it directly to differences in WM activity between schizophrenic patients and controls. In the Kamijo and Takeda (2009) study, and that of Nan et al. (2006), enough information was provided to indicate that the N2 became more pronounced as item level increased within the subitizing range. However, in neither paper was a detailed argument made as to the functional significance of the component. In terms of the range of eliciting circumstances outlined in section 2.1.5.2, and of speculations made by the serial counting explanation in section 1.2.3 regarding quantification operations which might take place during subitizing it is possible that the N2 might index integration of task relevant representations which could include assigning constituent identity to the items, and binding task relevant semantic information such as addition rules or number words to the item representations. In short, it is possible that the N2 reflects quantification operations. Given the noted association of the N2 with integrative processing in WM, and that in tasks where increased difficulty places demand on WM resources occurrence of an N2 is associated with increases in context updating time (i.e., P3b peak latency) and reduced P3b amplitude, the item level at which any N2 develops in this research program will be of significant interest in making inferences regarding the onset of serial WM activity in item enumeration.

The N2 was of course preceded by a P2 component in each of the tachistoscopic item enumeration studies but only Kamijo and Takeda (2009) examined it, finding increasing amplitude as item number increased. As noted above they also found an amplitude increase in the N2 with increasing item load. This P2-N2 pattern could reflect increased demands on WM resources, in line with Missonnier et al. (2003) speculations regarding this P2-N2 modulation, with the expansion on this position detailed at the end of section 2.1.5.3 based largely on the functional significance of the N2 argued for in section 2.1.5.2, and the dynamic nature of information flow in the course of a processing episode. However, given the conflicting behaviour of the P2, particularly, between target detection paradigms, and the supposedly WM demanding n-back task, predictions regarding the role and possible behaviour of the P2 in item enumeration will have to wait until it is empirically examined in Experiment 1.

2.2.2 Progression

The broad aim of the research program reported in this thesis is to examine the role of WM activity during item enumeration of response contingent displays, with the specific aim of using ERPs to assess whether the characteristics of subitizing are determined by preattentive, or parallel processes, or by capacity constraints on serial WM activity. Because there has been no previous ERP research examining item enumeration using response contingent displays the research program begins by examining the ERP profile generated by enumerating items under these conditions, with a specific focus on examining possible causes of the subitizing slope. In terms of the ERP components of the LPC reviewed in this chapter, and the (limited) findings from studies addressing item enumeration using tachistoscopic displays, a number of predictions can be made as to what sorts of LPC behaviour might reflect serial WM activity during item enumeration, with the crucial point being whether any of these indicators develop during subitizing. These predictions are developed in detail in the rationale for Experiment 1, reported in Chapter 4, with extensive reference back to this Chapter, and to Chapter 1.

CHAPTER 3

Technical Considerations and Methods

3.1 Laboratory Configuration

All testing was carried out in the Psychophysiology Laboratory at the University of New England. The laboratory consists of an outer room from which the experiment is controlled, and an adjoining room where testing takes place. Both rooms are climate controlled, providing consistent temperature and humidity across testing sessions, a desirable circumstance because fluctuations in these variables can alter the scalp electrode interface, thus altering the conductive properties of the electrodes across testing sessions (Kappenman & Luck, 2010). The laboratory is also electrically shielded, though this is now unnecessary given the high input impedances of modern EEG amplifiers (Ferree, Luu, Russell, & Tucker, 2001). Sound attenuation is adequate in the testing room, which is video monitored, and has capacity for bi-directional audio communication with the participant.

As is appropriate to avoid ground-loops, all AC outlets in the laboratory share a common, true ground (Cadwell & Villarreal, 1999). Mains power supply conforms to the Body Protected (BF) standard of AS/NZS 3003:2003. A BF environment protects participants against shock if any contact is made with a live conductor and earth (either directly or as a function of leakage currents), and is suitable for the procedures conducted in the Psychophysiology Laboratory at UNE.

Two separate recording equipment configurations were used. As no audio stimuli are used in this research program, the laboratory’s audio capabilities are not described beyond voice communication between the control and testing rooms.

3.1.1 Biopac System

Experiment 1 was conducted using a Biopac MP100 data acquisition unit (Biopac Systems, Inc., Goleta, California), with an array of 7 EEG amplifiers, and 1 EOG amplifier. The amplifiers are A/C coupled and have two high- and low-pass anologue filter settings (0.1Hz or 1Hz; 35Hz or 100Hz, respectively) and a selectable 50Hz notch filter. The MP100 performs analogue-to-digital (A/D) conversion, and binds trigger information from the stimulus presentation computer into the EEG data stream sent to the acquisition computer for storage. The unit had adequate A/D resolution of 16 bits, giving voltage resolution of 0.3μv, but relatively low input impedance of only 1 MOhm, a problem if inter-electrode impedances are not kept minimal (Ferree, et al., 2001; Kappenman & Luck, 2010). The Biopac system does not include an impedance testing routine, therefore impedance testing was accomplished using an AC impedance meter modified from an old Beckman Accutrace headbox. A Pentium PC running Windows 98 and Biopac’s AcQ software controlled the acquisition.

The *control computer,* performing stimulus presentation and response collection, was a PC running E-Prime software (details in section 3.2.1), using a 17 inch CRT monitor set at 85Hz vertical refresh. The keyboard number pad was used to collect response time. Transistor-to-transistor Logic (TTL) pulses were transmitted to the MP100 by a PCI I/O card through an optically isolated path.

3.1.2 NuAmps System

For subsequent experiments EEG was collected using a NuAmps EEG amplifier unit (Compumedics NeuroScan, 2003). The NuAmps unit is a true DC differential amplifier array with the capacity to record from 40 monopolar channels, all physically referenced to the Ground electrode, at sampling rates up to 1000Hz.

Subsequent re-referencing and bandpass parameters are accomplished in software. The Nuamps electrode cap configuration is illustrated in Figure 3.01. The unit has good (22 bit) A/D resolution, and excellent input impedance of not less than 80 MOhm. Thus, inter-electrode impedances higher than 5kOhm are not problematic (Ferree, et al., 2001). A PC running Windows XP Professional and the Acquire module of Neuroscan's Scan software controlled acquisition. Gain for all channels was fixed at 19 by the NuAmps hardware, corresponding to a voltage resolution of 0.063μv at the 22 bit A/D resolution of the NuAmps. Stimulus presentation was performed by the Windows XP controlled desktop PC running E-Prime software, using a 17" CRT monitor set at 85Hz vertical refresh The keyboard number pad was used to collect response time. TTL pulses were transmitted to the Nuamps by parallel port.

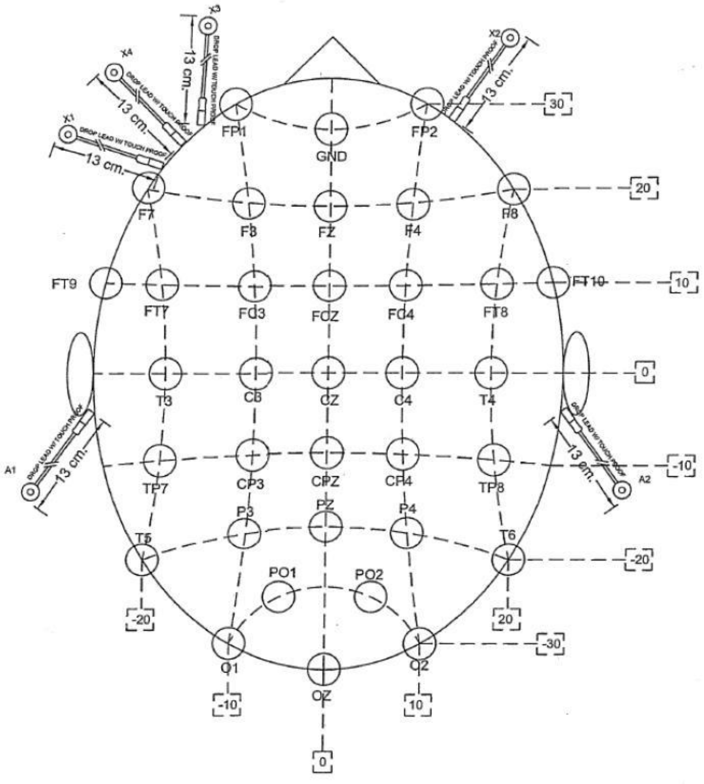


Figure 3.01. Electrode Montage – Nuamps System

3.2 Stimulus Presentation

3.2.1 E-Prime Software

Stimulus presentation and response collection for all studies was controlled by paradigms constructed in E-Prime software (v1.2; Schneider, Eschman, & Zuccolotto, 2001). E-Prime is an object oriented programming environment specialised for the design and implementation of a wide range of psychological paradigms requiring stimulus presentation and response collection. The program has a graphical user interface (GUI) which allows basic components of a paradigm to be created and manipulated without the necessity to write code. These components include procedure definitions, ordered lists of block and trial parameters, text, slide, and audio stimulus presentation objects, and feedback objects. Each object possesses a set of properties – display duration, input masks, and time to wait for responses, for example – that can be set through the GUI. In addition, E-Prime also incorporates a Visual Basic for Applications (VBA) programming language which makes it possible to manipulate the properties and methods of these objects by inserting user written code within the procedure timeline. This allows complex contingent branching procedures to be implemented, providing great flexibility and control over the content and progression of the elements comprising each trial, and the overall progression of the procedure. The VBA interface also gives precise capacity to generate stimuli, to create and draw screen elements using a code only object, the Canvas, to use port communication with other equipment, and to log important information, including custom time stamping of critical instants in code execution. All E-Prime GUI objects also have the capacity to log in detail the timing of their critical actions, and to accept and log responses from a range of input devices. Thus, while it is accessable to researchers with little or no programming experience, E-Prime also provides means to generate complex paradigms, and to account for and report the timing of key elements of experimental designs, such as stimulus duration, and interstimlus intervals, as well as to track the performance of important aspects of code execution during the procedure. This capacity allows the researcher to verify not only that the paradigm is constructed in a way that integrates effectively with the control computer’s hardware, thus ensuring accurate timing of trial elements, but also to report the timing integrity of the experimental manipulations.

3.2.2 Timing Integrity

The issue of timing integrity in cognitive research is an important one. High performance desktop PCs running multi-tasking operating systems (OS; predominantly Windows) are readily available and have become the tool of choice for conducting cognitive experiments, in conjunction with purpose built software routines, or commercially available packages such as E-Prime. Fast central processing units (CPU), ample random access memory (RAM), and excellent graphics and audio hardware combined with the multi-media capabilities of these OS allow for great flexibility in stimulus presentation, and the GUI allows for convenient access to these capabilities. However, PCs come in thousands of different hardware configurations, with many different software drivers. Not all software / hardware combinations are capable of interfacing with stimulus presentation programs in ways that allow the timing parameters of a paradigm to be implemented as written. Furthermore, modern OS take arbitrary control of non-system applications, sometimes for hundreds of milliseconds, disrupting their capacity to read the computer crystal clock and thus to deliver accurately timed stimuli. The fundamental purpose of E-Prime is to provide a software platform capable of delivering timing accuracy in experimental paradigms implemented on widely varying modern PC hardware / software configurations running Windows OS.

3.2.3 Millisecond Precision

The overiding concept driving E-Prime, and the appropriate development and implementation of cognitive experiments in general using PCs, is that of *millisecond precision*. This does not mean that no measured timing precision is aberrant by more than a single millisecond, which is impossible in the PC – OS environment. Rather, a broad operational definition of millisecond precision is that timing errors should not increase measurement variance by more than 1ms^2, an achievable outcome on appropriately configured hardware, running quality software routines.

We know that cognitive processes unfold over a matter of milliseconds (e.g., Brookhuis et al., 1981; Corbetta, Shulman, Miezin, & Petersen, 1995; Deutsch & Deutsch, 1963; Dien, Spencer, & Donchin, 2004). Consequently, the key manipulations in research addressing cognition are the types of information presented, and often the display durations of this information, and the intervals between information elements – inter-stimulus intervals (e.g., Bodner & Masson, 1997; Ferrand, Grainger, & Segui, 1994; Forster, 1998; Pesciarelli et al., 2007). It is manipulations such as these that allow inferences to be drawn regarding the timing and interrelation of cognitive processes, and in cognitive neuroscience the relationships between these processes and the underlying neural activity . Given the nature of cognitive research, and particularly that using EEG as a dependent variable, a focus on millisecond precision is essential so that error variance is minimised by reducing excessive machine variance, thus ensuring valid data that is reflecting the cognitive processes of interest.

3.2.4 Human Variance

In any EEG recording the total variance is the sum of human variance and machine variance. Human variance can be illustrated by *latency jitter* of ERPs. As explained in Chapter 2, because of signal to noise ratio issues ERPs are generally extracted through averaging many epochs of EEG, thus deriving an average waveform. However, when single trial ERPs are compared with average ERPs for an individual it is a common finding that the latencies of the single trial ERP components vary (jitter) around the average latencies {Coles, 1995 #207;Rugg, 1995 #206;Rugg, 1995 #208}. If we suppose that an ERP peak represents a cognitive process, then latency jitter implies that the timing of that process may not be entirely constant across trials. Of course, this jitter is also reflected in other indices extracted from the EEG, such as frequency domain measures. This variation may be a matter of fluctuating attention, which can be controlled for to some degree with sound methodology, but latency jitter is a fact of EEG research .

3.2.5 Machine Variance

In an EEG recording environment there are two ways in which timing aberrations and thus machine variance may occur. The first, touched on earlier in section 3.2, is through trial element timing errors by the control computer. The second source of possible error is in the TTL communication between the control computer and the EEG amplifier unit.

3.2.5.1 Control Computer

In order for E-Prime (or any stimulus presentation package) to be able to provide accurate timing of trial elements, it must be able to interface properly with it’s PC in four respects. Firstly, the program needs to be able to accurately identify the vertical blanking signal from the CRT monitor. This is vital, because all screen events are synchronised to screen refreshes, and thus the durations of all time critical trial elements need to be multiples of the refresh duration. If the PC configuration precludes accurate, consistent identification of the blanking signal then screen refreshes will be missed and consistently accurate timing of trial elements is impossible. Secondly, it is critical that E-Prime have excellent read access to the computer’s crystal clock. While the OS will regularly suspend for some milliseconds any program’s execution and thus access to the clock, E-Prime’s authors present data showing that they implement effective code routines to minimise the affect of such time-outs on timing accuracy. However, again some PC configurations do not support accurate, consistent access to the hardware clock, and on such machines the critical timing of code execution, and thus trial progression, will be disrupted periodically when the hardware signals from the clock are unavailable. The third constraint is that E-Prime needs to have effective communication with the video-card through it’s software driver, in order to cache and to draw stimuli effectively. If this communication is not consistent then problems can arise in the timing of critical screen events, particularly if the code performing the graphics operations in E-Prime is situated within a time-critical portion of the trial sequence. This bears on the fourth constraint, which is appropriate code organisation. It is imperative that the paradigm be written in such a way that stimulus preparation code either is executed during non-time critical portions of a trial, or, if this is not possible, the host PC must be capable of performing the necessary operations without disrupting timing accuracy. Given an appropriate host computer and good code organisation, E-Prime claims to be able to deliver millisecond precision, a claim supported by timing data from our system.

3.2.5.2 TTL Communication

In order for EEG research into cognition to proceed appropriately, accurate *time locking* of stimulus presentation to the EEG must occur. To achieve this the paradigm must include instructions for the control computer to send a condition specific signal (TTL pulse) to the EEG amplifier at the moment of stimulus onset (generally the beginning of the screen refresh where the stimuli is drawn), and the EEG amplifier must bind that information accurately to the continuous EEG, which is then streamed to the acquisition computer for storage. If this organisation is effective time locking has been achieved - the EEG following the event-markers will contain electrical activity generated by brain regions involved in task related cognition, and this signal may be extracted using the marks as reference. However, consistent inaccuracies in TTL communication would obviously be problematic because in epochs where errors occurred the time course of the task related processes would be early or late relative to accurately marked epochs. Problems with TTL communication could occur because of badly written code, or poor hardware / software configuration on the control computer, thus causing delays between execution of the port communication code and sending of the TTL pulse, or because of faulty routines in the EEG amplifier for binding the TTL information to the continuous EEG.

3.2.6 Considerations

It is easy to see how the addition of machine variance to an EEG recording could result in excessive error variance, and useless data. In a worst case scenario where the control computer is unsuitable, with poor code organisation, and excessive slop in TTL communication any differences in the EEG due to experimental manipulations could well be completely distorted. This can be illustrated by referring to Figure 3.02. The leftmost pane shows an example of the N1 – P1 – N2 complex, where peak amplitudes for these fast components differs across conditions. The central pane represents the prototypic N1 – P1 – N2 complex. The top right pane represents the occurrence of this complex over 6 trials where time locking is incorrect, either through stimulus onset aberrations or TTL slop, with the bottom right pane being the hypothetical average of these 6 trials. It is obvious that poor time locking could lead to the ‘process’ represented by the complex being averaged out of the signal entirely

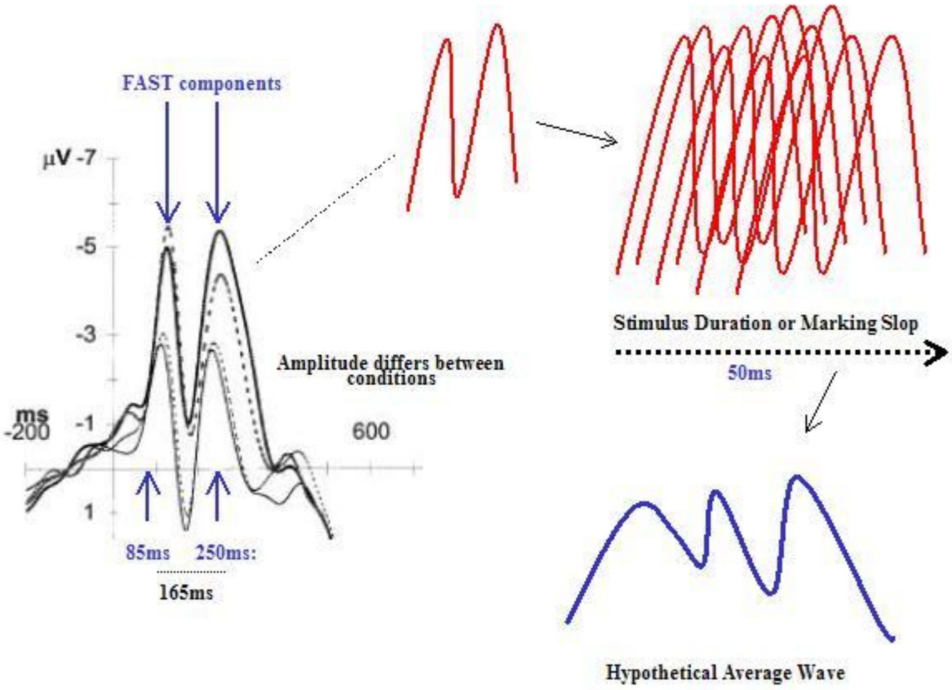


Figure 3.02. Hypothetical of the effect of machine variance.

3.2.7 Implementing Millisecond Precision

In order to implement millisecond precision it is necessary to assess and control for timing aberrations in the EEG recording environment during the course of each trial by recording timing information which allows monitoring of the durations of critical events on the control computer, including tracking of screen refreshes, and sending of TTL pulses to the EEG amplifier unit. This is accomplished within the code describing the paradigm. To assess the accuracy of the TTL communication it is also necessary to extract the event-marker latency information from the EEG file, a facility provided in the proprietary software (Scan version 4.3, Neuroscan, 2003). Once the appropriate information has been made available from the two PCs in the recording system, any trials where timing aberrations might have occurred can be identified and flagged for exclusion from later processing.

All paradigms used in the research presented in this thesis have been written entirely by the thesis author. They all have included commands that time stamp the beginning and end of execution of blocks of code which implement time critical segments of each trial. Scripts (also written by the thesis author) written in the SPSS basic language (SaxBasic when these scripts were written, WinWrap basic now after IBM bought SPSS) then took these values and computed critical durations and intervals, identified any aberrant trials and flagged them for exclusion from later processing, and produced trial by trial and summary output detailing the performance of the control computer. The script also took the latency values of the event-markers in the EEG file and identified epochs where the control computer and the EEG amplifier system may have failed to synchronise a TTL pulse with an event marker, then flagged any such epochs for exclusion and produced diagnostic output. This approach ensured timing integrity. Prior to testing taking place the capacity of our EEG recording system for millisecond precision was assessed using this procedure, an undertaking that proved valuable. The graphics subsystem of our initial control computer proved unequal to the demands of millisecond precision. We changed computers, finally settling on a Radeon graphics solution that provided perfect graphics performance.

3.3 Data Handling and Interpretation

3.3.1 Data Extraction

Recording parameters for all of the EEG data presented throughout this thesis are detailed in the various Method sections. Scripts written (by the thesis author) in the *TCL* language interface available in the Neuroscan Scan 4.3 software were used to automate filtering of the continuous EEG, extraction of epochs around critical events, artifact assessment and rejection, averaging of appropriate epochs, and extraction of the key values for subsequent statistical analyses. These values were picked up by SaxBasic scripts in SPSS and formatted for subsequent repeated measures analyses.

3.3.2 Effect Sizes

It is common practice these days to move away from sole reliance on hard and fast significance testing and take considerable account of the *size* of any effect when interpreting the findings reflected in statistical output {Cohen, 1994 #376}. *Partial eta-squared* (denoted as η_p^2) is used as a measure of effect size throughout this thesis. Partial eta-squared is a measure of variance, like r-squared. It partials out the effects of other variables, and from the remainder tells us what proportion is explained by the effect, and its error variance. Rules of thumb for size of effects suggest that η_p^2 of .01, .06, and .14 correspond to small, moderate, and strong effects, respectively. However, a more conservative approach is taken in this research program. Throughout this thesis ‘non-significant’ effects are often discussed in terms of their effect size, with η_p^2 of around or less than .2 referred to as a *small* or *weak* effect, and these are not often discussed as meaningful. However, those effects with a η_p^2 from around .25 to .4 are referred to as ‘*moderate’*, while those around .5 or above are referred to as ‘*strong’* effects, and a number of such effects are discussed at length in interpreting the experiments which follow.

CHAPTER 4

Experiment 1: ERP Activity During Item Enumeration

4.1 Rationale, Aims, and Hypotheses

In Chapter 1 it was noted that dominant explanations of subitizing suggest that it is determined by the operation of preattentive, or parallel processes. This observation is based on the supposition that the numerosity of small collections of items is apprehended seemingly immediately, and without apparent effortful cognitive processing. However, as noted in sections 1.2.1 and 1.2.2, a key finding suggesting that subitizing does not reflect *immediate* apprehension of numerosity is that there is a subitizing slope that is described by a non-zero function. Thus, any explanation of subitizing has to address this fundamental characteristic. If indeed subitizing is determined by the limited capacity of some mechanism that performs operations critical to quantification (but not necessarily quantification itself) on multiple units of visual information in parallel—whether preattentive or not—then prior to this capacity being exceeded there would be no reason to expect increases in enumeration time and so the slope should be zero. A common explanation for the slope has been that serial WM activity is involved in binding the number word to the quantified collection, an approach based on Sternberg's (1966) work on memory scanning, that suggests the numerical increase in RTs at each of the two steps within the slope reflects the additional time taken to scan one extra slot through a sequential list of number names in order to access the appropriate number word (Gallistel & Gelman, 1991; Klahr & Wallace, 1973; Mandler & Shebo, 1982; Trick & Pylyshyn, 1994).

For this explanation to be valid scan times across the item levels in the subitizing slope should be of equivalent increments, as for each extra item in an array only a single extra number word list entry needs to be scanned—in which case the subitizing function should be linear. However, over the years a common finding in enumeration studies, noted in section 1.2.2, has been that the RT increments between successive item levels comprising the subitizing slope are generally not numerically equal (by definition of course they are *statistically* non-different), with the increment from two to three items often being more than twice as large as the increment between one and two items (described in section 1.2.2 as the non-linear subitizing slope). This observation is inconsistent with claims that the subitizing slope *solely* reflects serial list search by WM for the appropriate number name, and so with preattentive or parallel processing explanations of subitizing. Existence of the non-linear slope begs the question as to whether, at the critical step to three items that defines the non-linear subitizing slope, and still within the subitizing *span* of almost all people, additional serial processing in WM beyond that required for list search alone may be determining the subitizing slope. This supposition is consistent with the serial counting explanation, outlined in section 1.2.3, which suggests that serial processing in WM is involved with enumerating even two or three items, and that the subitizing slope results from an increase in activation and binding demands with the addition of each extra item, that taxes the limited resources of attention. This position is supported by a number of recent findings, reviewed in section 1.2.4, where disruption of attentional deployment during subitizing has disrupted subitizing efficiency.

As noted in Chapter 2, there has been no published ERP research addressing item enumeration of displays where stimulus exposure time has been subject determined—that is, response contingent displays. Therefore, the aim of this first experiment was to contrast preattentive or parallel processing based explanations of subitizing with the serial counting explanation by presenting subject terminated dot arrays and relating features of the morphology of the LPC across item levels to the reaction time profile describing the RT subitizing slope. The primary goals were to identify any components of the LPC that might be informative about whether or not serial WM activity is involved in determining the subitizing slope, and assess whether chunking capacity (see sections 1.1.2 and 2.1.5.1.1) may determine the subitizing range.

The focus of this research is on changes in the morphology of the LPC that might occur firstly across item levels one to three, which comprise the subitizing slope, and then across four, or five items—where, as information load increases, most people would be expected to make the transition from subitizing to counting as indicated by significant RT costs (see section 1.2.1). So, while the task in this experiment included arrays of from one to ten items (thus controlling for density judgements at six and seven items), only the data from levels one to seven are analysed, with a clear explanatory focus on item levels within the subitizing range of one to four items, which encompasses the subitizing slope of one to three items, and the subitizing span of the entire sample. As noted in section 1.2.1, individual subitizing spans fall predominantly in the three or four item range, but they can vary up to five or (very rarely) even six items. Thus, analyses across item levels one to seven should show any changes in ERP activity that might occur within the subitizing slope (one to three items), and also that might indicate transition from subitizing (one to three or four items) to counting (beyond four items).

Individual differences are obviously important to take into account when examining the possibility that serial WM activity might be involved in subitizing, and the closely related idea that limitations in the capacity for activation in WM during an episode of enumeration might determine the individual subitizing span. In section

1.1.2 the concept of a processing episode was introduced, along with its proposed relationship to chunking capacity, and in section 2.1.5.1.1 these concepts were explicitly related to duration of an episode of context updating in WM, with context updating occurring across the course of the LPC and the endpoint of this processing episode being defined by the P3b peak. Now, if (a) the LPC is informative about chunking capacity, and (b) chunking capacity is related to the limits that determine the subitizing span, and (c) individuals differ in their subitizing spans (they do; see section 1.2.1), then any key indicators in the LPC of chunking capacity being reached or exceeded should vary as a function of item level across participants with different subitizing spans.

This variation could appear as either a *discrete* effect, a *distributed* effect, or a combination of the two. For example, in terms of a group with an RT derived sibitizing span of three items, and another with a span of four items, a discrete effect would manifest as a specific change in behaviour of an ERP component, or the appearance of a novel component, that is associated systematically with the RT derived subitizing span such that the change appears one item later for those with a span of four items than for those with a span of three items. In contrast, a distributed effect would appear as change that is apparent across more than a single increment in number of items, and that is evident in both groups at at least one shared item level. Because there have been no previous ERP studies addressing response contingent item enumeration it is not possible to make stark predictions as to just how the LPC components identified in Chapter 2 might behave in terms of discrete versus distributed changes as information load is incremented. Firstly, an ERP profile associated with item enumeration needs to be established—the goal of this first experiment.

To this end, prior to examining any span group differences, the initial analyses reported for this experiment examine the ERP and RT data for the entire sample as *one group*. This maximises the power to address questions about the types of processing that might be occurring during subitizing, and within the subitizing slope, since all participants were still subitizing (by RT) at three items. Additional questions about possible ERP changes related to subitizing span can be seen in terms of weighted contributions of the span groups to either discrete or distributed changes in the waveforms. The hypotheses presented below address the waveforms from this perspective.

In Chapter 2 it was suggested that related modulations in three components of the fronto-central LPC—the P2, the N2, and the P3b—might reflect load related WM activity during enumeration of increasing numbers of items, and that these modulations could be indicative of serial WM activity that might determine the subitizing slope, as well as the individual subject's subitizing spans. In developing the Hypotheses that follow, a chronometric approach is outlined, taking advantage of the documented sensitivity of latency of the P3b peak to task manipulations that ostensibly load WM, and which prolong RT (see section 2.1.5.1.1). The information regarding the components of the LPC that was detailed in Chapter 2 can be summarised around a series of hypotheses that reflect the logical progression of the analyses as they are presented later in this chapter. These analyses begin with the P3b peak, as it is the endpoint of the LPC and represents the culmination and measure of duration of an episode of context updating. Four Hypotheses address the P3b. Beginning with the P3b is essential because the duration of the LPC and thus an episode of context updating (and a processing episode) as defined in this research project, is determined by the P3b peak, and the analyses assessing the behaviour of the preceding N2 and P2 examine their relationship to any changes in duration of context updating. The P2 is addressed next, as it represents the first landmark in the waveform, starting around 100 ms and peaking around 200 ms, and changes in the P2 relative to the P3b will determine how any subsequent changes in the N2 are interpreted. Because of conflicting possibilities of the P2s behaviour in the item enumeration task the component's latency and amplitude are assessed without any clear predictions to confirm or not, however a single Hypothesis assessing its relation to P3b latency is presented. Finally, two Hypotheses address the negative deflection, measured here as the N2 peak.

Firstly, as detailed in section 2.1.5.1.1, across a range of paradigms where an immediate response is demanded and thus the processing required to accomplish the task (i.e., select a response) can occur within the duration of the LPC (i.e., within a processing episode), P3b latency has been shown to increase as processing load and complexity are increased, which presumably leads to more time consuming processing in WM—that is, serial processing—in order for context updating to be achieved. As noted in section 2.1.5.1.1, once processing demands can no longer be accomplished within the duration of the LPC (signaled by RTs occurring well beyond the latency of the P3b peak), further conditions of increasing difficulty, indicated by increased RT, do not result in P3b peak latency increases. As described in section 2.2.2, in the item enumeration task the item level at which this 'saturation' might occur could have a relationship with the RT derived subitizing span, indicating that a capacity limit on the amount of information that can be processed over the course of the LPC may be implicated in determining the subitizing limits. Thus:

Hypothesis 1: If there are increasing serial WM demands as item level increases within the subitizing *slope* then latency of the P3b peak should increase with the addition of each extra item—that is, between successive item levels one and two, and two and three.

Hypothesis 2. If capacity of the LPC is associated with RT derived subitizing capacity, then there should be a relationship between the item level where P3b latency ceases to increase, and the group's RT derived subitizing *span*.

Data supporting Hypotheses 1 and 2 would establish a *general* relationship between P3b peak latency, and thus context updating duration, and the capacity limitations reflected by the RT derived subitizing span. Such a relationship could then be examined further to assess whether any serial WM activity reflected by increases in P3b latency between successive item levels in the subitizing slope is associated with corresponding RT increases, and thus with determining the subitizing slope of this group, allowing the list scanning explanation of the subitizing slope to be assessed.

The logic of this approach can be explained by initially referring to the P3b latency values alone. As already noted, the list scanning explanation of the subitizing slope suggests that the time cost in WM to scan each extra list entry should be equal. The P3b latency data can be used to make inferences about the involvement of serial binding activity *beyond that required for list search alone* at the critical step from two to three items, based on the following logic: Any P3b latency increase between one item and two items would provide a base measure of the time cost to scan one extra list entry, assuming that this is a critical step in WM context updating to achieve enumeration. If the only functional difference between enumerating one item or two items is search time through an ordered list of number names (i.e., a two item array is subitized by the same specialised mechanism which quantifies a single item) then the peak latency difference between these conditions should reflect the scanning time per extra list item. Thus, if quantifying three items is functionally equivalent to quantifying two (i.e., a three item array is subitized by the same specialised mechanism which quantifies one or two items) there should be an equivalent latency increase between two and three items, as a single extra list entry needs to be scanned. However, to provide compelling evidence that any processing costs reflected by P3b latency increments are implicated in determining the response costs defining the subitizing slope (whether of equal or unequal increments), a relationship would first need to be shown between the P3b increments and any corresponding RT increments.

This can be addressed by examining another property of P3b latency, its *sensitivity* to manipulations which increase RT, which was detailed in section 2.1.5.1.1. In some paradigms which present conditions with increasing processing load, notably visual search and Sternberg paradigms which might share some processing demands with the response contingent item enumeration task (see section 2.1.5.1.1), increases in P3b latency between successively more difficult conditions show medium to high sensitivity to the same task manipulations that prolong RT between the conditions. This property may be useful to assess the degree to which increased processing time between successive item levels from one to three items, reflected by P3b latency increases (e.g., latency at 2 items minus latency at 1 item; latency at 3 items minus latency at 2 items), is associated with corresponding RT increases, and thus provide a basis for assessing the list scanning explanation of the RT subitizing slope.

Hypothesis 3. If time taken to achieve context updating is related to the numeric RT increments within the subitizing slope (which by definition are not statistically significant) then any P3b latency increases between one and two items, and two and three items should show medium to high sensitivity to any corresponding RT increments.

Data supporting Hypothesis 3 would establish a *specific* relationship between incremental increases in the duration of context updating during subitizing, as reflected by increases in P3b latency, and RT increments over the subitizing slope. Thus, if the RT slope of this sample is linear—that is, the increment between two and three items is of equivalent magnitude to that between one and two items—the list scanning explanation of the subitizing slope would be supported. However, if the RT slope of this sample is *non-linear*—that is, the increment between two and three items is of greater magnitude to that between one and two items—then *significant* P3b latency increases that were also sensitive to the corresponding RT increments would support the inference that serial processing in WM *beyond that required for list scanning alone* is implicated in determining the subitizing slope.

In section 2.1.5.1.2 another robust feature of the P3b was noted. Load related P3b latency increases are accompanied by a *decrease* in amplitude, a pattern consistently seen in the P3b in tasks where conditions increase in difficulty (see Kok, 2001) and thus place an increased load on WM. A similar finding in the subitizing slope in the current data would provide convergent evidence that serial WM activity may be implicated in determining the slope, and may also serve as a more general indicator of processing capacity being taxed. Thus:

Hypothesis 4. If increased processing load demands on WM are involved in prolonging P3b latency, increases in P3b latency should be accompanied by a decrease in P3b amplitude as item load increases.

The next component of interest is the P2. As noted in section 2.1.5.1.2, one interpretation of P3b amplitude decrease under processing load has been that it reflects the temporal separation of previously concurrent, and thus summating processes, with timing of the first component staying relatively static while the second, the P3b, occurs later as context updating is delayed by the intercession of extra processing requirements, which may be reflected by slow negative waves in the 200 to 400 ms range that could indicate load related activity in WM. This morphology was explained in section 2.1.5.3 and in 2.1.5.4, where development of the P2-N2 complex and its possible functional significance as an index of load related WM activity was discussed. In terms of the possible task demands of the response contingent item enumeration task it is unclear what to expect of the P2. Arguments regarding the devolution of the P3b would suggest that latency of the first positive peak, what is being suggested here as the P2, would stay static (see section 2.1.5.1.2), while the latter peak is delayed. On the other hand, target detection paradigms sometimes see the P2 latency become prolonged when object identification is difficult (see section 2.1.5.3), yet in a WM demanding n-back task a P2 component was shown to *decrease* in latency as WM load increased (see section 2.1.5.4). Similarly with amplitude, the behaviour of P2 like components has been mixed under differing paradigms which obviously demand differing types of processing (see section 2.1.5.3). Given the uncertainty of just what type of processing subitizing in the item enumeration task will recruit during this early part of a processing episode it is not possible to make clear hypotheses as to how P2 latency and amplitude might behave. Therefore P2 latency and amplitude will be assessed to provide a profile of P2 behaviour within the LPC as item level increases. However, the influence of the timing of any processes reflected by the P2 on prolongation of the following P3b peak within the subitizing slope is important in interpreting the behaviour of the following N2, so it will be formally assessed:

Hypothesis 5. If timing of whatever processes generating the P2 influences directly any delays in context updating, then any latency changes in the P2 between successive item levels across the subitizing slope should be related (sensitive) to any corresponding changes in P3b latency.

On the basis that there are increases in context updating time across the subitizing slope, support for Hypothesis 5 would cast doubt on theories suggesting that prolongation of P3b latency is a result of time consuming WM processes in the 200-400 ms window interceding between relatively time invariant processes reflected by an early positive peak, the P2, and the later P3b. Instead, this would suggest that the locus for load related delays in context updating could be earlier in the processing stream, starting perhaps 100 ms post-stimulus, and may involve not integrative WM processes, but more primitive processes related to item selection.

The final component of interest is the N2. As noted in section 2.1.5, and 2.1.5.2, one view is that negative deflections in the 200-400 ms range such as the N2 are involved in the temporal separation of the two positive components of the fronto-central LPC, because they reflect additional load related activity in WM that prolongs the time necessary to achieve context updating—thus, the P3b takes longer to peak because of the duration of additional processes that the extra load demands, and these demands occur relatively late in the processing stream, after earlier selection processes. In section 2.1.5.2 evidence was presented showing that across a range of paradigms these fronto-central negative going deflections tend to show drops in relative amplitude and increases in latency as a function of task difficulty, and thus increased demands on WM. If the N2 reflects serial WM activity, in this case the operation of quantification processes as suggested in sections 2.1.5.2 and 2.1.5.4, then it is likely to extend for longer, and to drop in amplitude with increasing item load. Thus:

Hypothesis 6. If the N2 reflects load related serial WM activity that precedes and prolongs the duration of context updating reflected by latency of the P3b peak, then a) N2 latency should increase with increasing item level across the subitizing slope, and b) the increases should be sensitive to any corresponding P3b latency increases.

Hypothesis 7. If the N2 reflects load related serial WM activity, then N2 relative amplitude should drop as WM load increases sufficiently to require serial activity, and this should be coincident with P3b amplitude reduction.

4.2 Method

4.2.1 Participants

Seventeen adults from Armidale, Australia, participated. Two terminated their sessions early, and three provided insufficient artifact free trials at each condition, leaving a final sample of 12 adults (4 male; mean age = 25 years, SD = 7.4). All had normal or corrected to normal vision, and reported being right handed.

4.2.2 Task

The item enumeration task used for this experiment consisted of repeated trials of individual item arrays comprising from 1 to 10 items, presented in random order on a computer monitor in a 5 * 5 centred grid measuring 4 * 4 cm on screen. The grid thus subtended less than 1 degree of visual angle at the viewing distance of 1.2m. The grid lines were not visible. Items were the solid square symbol from the Windows Map Symbol font set (U+0021), presented in black on a gray background, at 9mm^2. For each trial the stimuli were randomised within the grid. There were no constraints on either configuration, or on successive trials being from the same condition (item level). A speeded button press response was required once subjects had enumerated correctly the number of items in the array. To reduce ocular and movement artifact, each trial was initiated by the participant when they had judged that they were ready to concentrate during the trial without needing to blink or move in the chair. Once the trial was initiated a centred black fixation cross was displayed for 1500 ms, followed by 235 ms of gray screen[1], then the item array was drawn to screen, remaining visible until the participant responded by pressing the '+' key on the keyboard numberpad with the right index finger, at which time RT was recorded. After a gray screen for 1000 ms a centred pink question mark appeared prompting the participant to verbally report the numerosity of the array. The experimenter keyed in the response and a fixation cross was then displayed with feedback – 'correct' (in yellow), or 'wrong' (in red) – displayed above it. This cross and feedback remained visible until the participant pressed the '+' key on the keyboard numberpad when ready to initiate the next trial. This cleared the feedback, leaving the fixation cross to begin the next trial.

4.2.3 Procedure

The procedure required attendance at two sessions. The first, lasting approximately 45 minutes, was a familiarisation and practice session where no EEG was recorded. This session was considered important because prior to EEG data collection it allowed participants to learn the critical task requirements, and thus develop practiced enumeration strategies within the artifact minimisation constraints, described below. It was hoped this approach, along with the self-paced paradigm,

---

[1] It was originally proposed to have a 35 ms (3 x screen refresh) delay between fixation cross offset and array onset. However, before moving to the Scan software we used custom built data extraction routines that required a TTL pulse width of 200 ms. The E-Prime code holding the pins high over this duration delayed execution of the stimulus presentation code, thus dictating an additional 200 ms between fixation offset and stimulus onset. This timing organisation was only confirmed halfway through the data collection for this experiment, so the paradigm was left unaltered.

would maximise attention to the task, and minimise epoch attrition through artifact during the relatively lengthy data collection session.

Upon arriving at the lab for this session participants were initially given basic information on the experimental procedure, and informed consent was gained. They were then seated comfortably in the testing room approximately 1.2m from the computer monitor and informed that they would be monitored by closed circuit TV throughout the procedure. Bi-directional voice communication was then explained, followed by a briefing on the importance of minimising ocular and movement artifacts during an EEG recording, and an explanation of a trial sequence. The self-paced nature of the paradigm was explained, with instructions to delay cueing each trial until they were sure their eyes were relaxed and they were settled and comfortable in the chair, and to avoid blinking and gross movements during the *critical period* – from onset of the fixation cross through to the numerosity prompt. Participants were instructed explicitly to enumerate the arrays and respond *as fast as possible, while maintaining accuracy*. Thus, both speed and accuracy were stressed, with the admonition to use the following practice trials to freely develop enumeration strategies within these constraints. The testing room lights were then dimmed and participants completed 100 practice enumeration trials—ten at each item level.

Session two took place between 1 and 10 days following session one, and lasted approximately 2.25 hours, depending on the self-pacing of each participant. Once participants were seated in the testing room the session began with a rebriefing on the task, and the importance of using the self-pacing to minimise artifacts. The electrodes were then attached and the task begun. There were 10 familiarisation trials—one at each item level—followed by 420 experimental trials presented in 14 randomised blocks of 30 trials comprising three trials at each of the 10 item levels.

Upon completion subjects were debriefed regarding the strategies they used during the task.

4.2.4 EEG Recording

For this experiment the electrophysiological signals were recorded using the BioPac equipment, detailed in section 3.1.1. The EEG and electrooculogram (EOG) were recorded using Ag/AgCl disposable electrodes (10mm) from 2 scalp sites: Fz and Cz (according to the 10-20 system) referenced to A2, with a single bipolar EOG channel from electrodes at the outer canthi, and below the right eye. This EOG arrangement allows capture of both vertical and horizontal eye movements with a single channel (Deacon, Breton, Ritter, & Vaughan Jr., 1991). All channels had an analogue bandpass 0.1Hz to 100Hz. EEG channels were amplified at a 10,000 gain setting, the EOG at 5000. The continuous EEG/EOG was sampled at 500 Hz at an A/D resolution of 16 bits, and recorded to a hard disk along with stimulus event marking information.

4.2.5 EEG Data Reduction

For each subject the continuous EEG was digitally low-pass filtered offline at 25 Hz, 24db roll-off, using a finite impulse response (FIR), zero phase shift filter. Epochs including a 200 ms prestimulus baseline and extending 1000 ms post stimulus onset were extracted, baseline corrected, and averaged per condition. Epochs where voltage in any channel exceeded +/- 75 micro-volts were excluded from the average, as were trials with an incorrect response, or where RT was either below 300 ms or excessive (see data screening criteria, below). Grand mean waveforms were then generated for each condition, inspected for evident components, and component windows derived. These are detailed below in section 4.3.3. Latency and amplitude values within these windows were then extracted from individual subject's average waves.

4.3 Results and Discussion

4.3.1 Data Screening

As explained in section 4.1, while arrays of eight, nine, and ten items were presented to control for density based judgements at six and seven items, only data from item levels one to seven were used in the analyses, leaving a total of 3528 experimental enumeration trials from the 12 participants. Across these trials there were only 20 errors, with five the highest by any subject, an overall rate of 0.6%, confirming that subjects concentrated on enumerating accurately as instructed. These trials were excluded from all analyses. Error rates will not be addressed further in these analyses.

RT data was screened for each subject by firstly excluding trials where errors were made or where RTs were below 300 ms (arbitrary cutoff for responding prior to accessing the name word, and guessing or using information in an iconic store to complete the enumeration process), then computing a Z-distribution per item level, and excluding all trials +/- 2 standard deviations. As far as possible, this approach controls for trials where concentration lapses extend stimulus evaluation and response execution, as well as for trials where quick responses were made prior to accessing the number word. This criteria led to a further 153 trials (4.3%) being excluded. Additionally, 139 trials (3.9%) with a correct response that were rejected from EEG analyses due to artifact were also identified and excluded from the RT analyses. Thus 8.8% of trials (312) were excluded for all reasons, leaving 3216 trials in total, with an average of 459 trials (38 per subject) at each item level, for the RT and EEG analyses.

4.3.2 Behavioural Data

Mean RTs per condition for the entire sample are presented in Table 4.01. As expected, RTs for one, two, and three items were swift (< 500 ms). The subitizing slope was evident, with an *average* increment of 32.5 ms per additional item. However, the increments per item were not numerically equivalent. The numerical difference in RT between one item and two items was 12 ms, while that between two and three items was 53 ms, indicating (by the argument detailed in section 1.2.2) a non-linear subitizing slope for this sample. The RT cost between three and four items again rose sharply, to 130 ms, while from five items on the increments are all in the range of 300-350 ms. This profile fits the classic enumeration plot, with a relatively flat function describing one to three items, which is typical of subitizing, and a steep slope describing four to seven items, which is typical of counting. This dichotomy is best illustrated by plotting the means, as in Figure 4.02, which shows an elbow in the function beginning at three items. To confirm that subitizing occurred in this sample the subitizing span for each subject was computed (by using orthogonal difference contrasts between item levels). The subitizing span was taken as the item level prior to that at which the contrast was significant. All subjects showed evidence of subitizing. Eight subjects had spans of three, and four subjects had spans of four items. Thus, the average span for the group was 3.33.

Table 4.01. Average RTs, standard deviations, and number of trials for each item level across the 12 participants.

Items	RT(ms)	SD(ms)	N (trials)
1	416	82	445
2	428	74	449
3	481	87	452
4	611	156	458
5	937	288	460
6	1251	337	460
7	1598	388	472

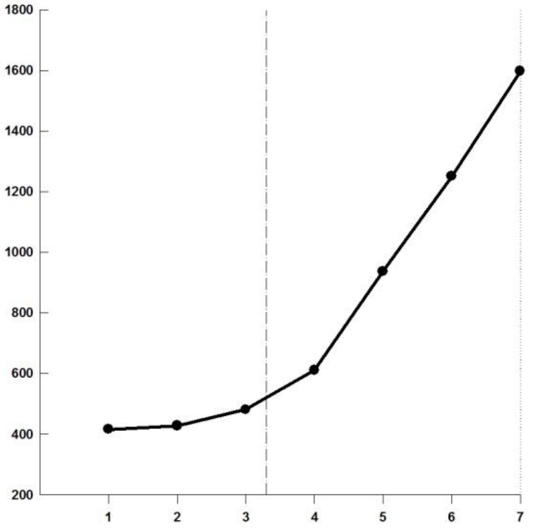


Figure 4.01. Mean RTs for each item level. The vertical dashed line is the subitizing span for the group.

Summary. These data confirm the general RT profile of item enumeration that has emerged from previous studies (Balakrishnan & Ashby, 1991; Gallistel & Gelman, 1991; Logie & Baddeley, 1987; Mandler & Shebo, 1982; Trick, 1990; Van Oeffelen & Vos, 1982; Wolters, Van Kempen, & Wijlhuizen, 1987). There is evidence of a discontinuity in the RT slope that shows a small, numerical increase in RT across one to three items, then a larger numeric increase between three and four items (where eight of the subjects ceased to subitize), after which the typical linear counting slope is evident. The numerical differences between RTs for one and two items (12 ms) and two and three items (53 ms) are consistent with those reported in other studies where a non-linear subitizing slope has been found (Jensen, Reese, & Reese, 1950; Mandler & Shebo, 1982). The consistency of this RT pattern with that found in the wider subitizing literature also suggests that using a button press response rather than a voice trigger to record RT has provided valid RT data. Given that the RT data for this sample indicate that subitizing took place, and thus that a transition occurred between subitizing and counting, the EEG data may provide some insight into the determinants of the non-equivalent RT increments of the subitizing slope, and thus directly address the question of serial WM involvement in subitizing.

4.3.3 EEG Data

Data from Fz and Cz are used to assess the fronto-central LPC behaviour. The EEG data is graphed over two overlapping item ranges, due to the number of conditions (i.e., for ease of viewing), and the distinction between subitizing and counting in this sample (i.e., for functional considerations). The first range, from one to four items, encompasses the RT derived subitizing span of all subjects—the subitizing *range*. The second, from four to seven items should reflect activity after the subitizing span has been exceeded and counting must occur—the counting range. The grand-average waveforms for Fz and Cz, plotted in the two overlapping ranges, are presented in Figure 4.02.

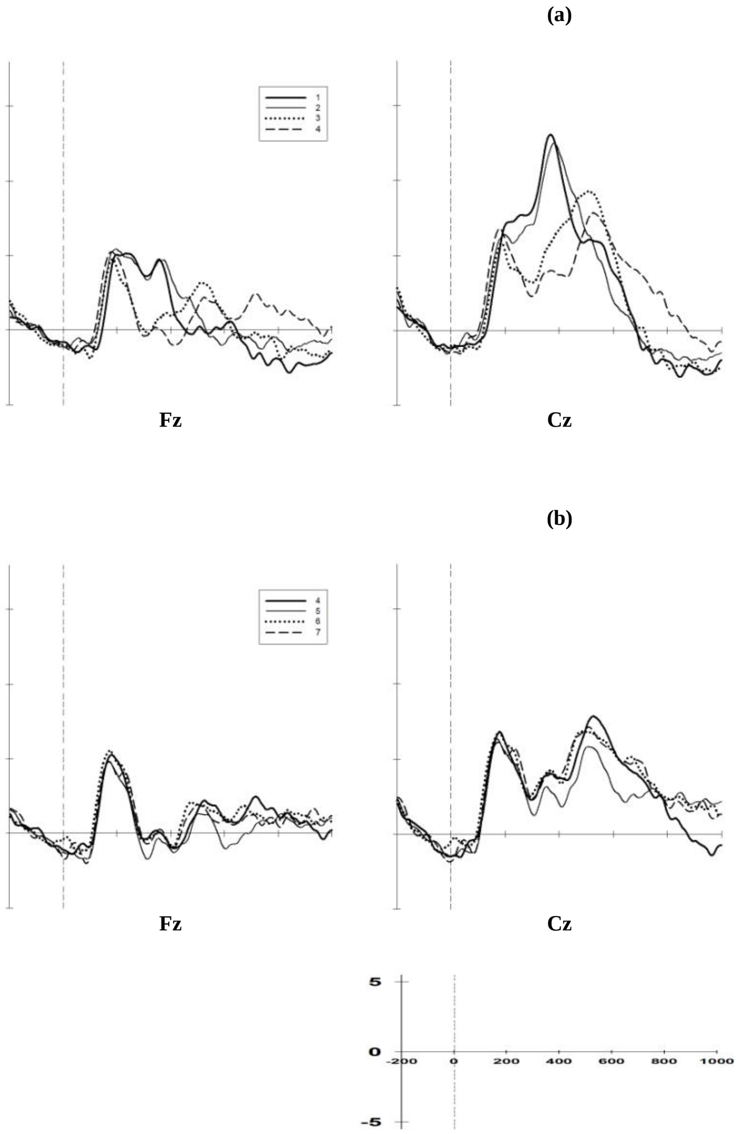


Figure 4.02. Grand average waves at Fz and Cz for 1-4 items (panel a), and 4-7 items (panel b).

In the 1-4 item range there is a distinct change in morphology evident at 3 items, with full development of an early fronto-central positive deflection beginning around 90 ms and peaking around 180 ms. This *P2 component* is defined by development of a following negative deflection with an apparent frontal maximum beginning around 190 ms post-stimulus, and peaking around 300 ms post-stimulus; the *N2*. At Cz development of this P2-N2 complex is suggested at even one item by a sharp flattening of the gradient between 200 ms and 300 ms, and this negative deflection appears more pronounced for two items. At both Fz and Cz it then becomes strongly evident for three items, peaking at four items. Looking at the grand average waves for 4 to 7 items, in Figure 4.02, panel b, it is obvious that the general morphology that developed across three and four items is consistent through the subsequent item levels. To assess these components, peak latency and amplitude values were extracted in two overlapping windows: the P2 was taken as the most *positive* peak between 90 ms to 250 ms, and the N2 was taken as the most *negative* peak between 150 ms to 350 ms. These windows were chosen after examination of the individual subject's average waves revealed some variation in the timing of this positive-negative complex. The N2 deflection then climbs once again to a distinct positive peak between 350 and 550 ms, *the P3b*, measured as the latency and amplitude of the most positive peak in the window from 350 ms to 650 ms. These component windows are summarised in Table 4.02.

Table 4.02. Components and component windows.

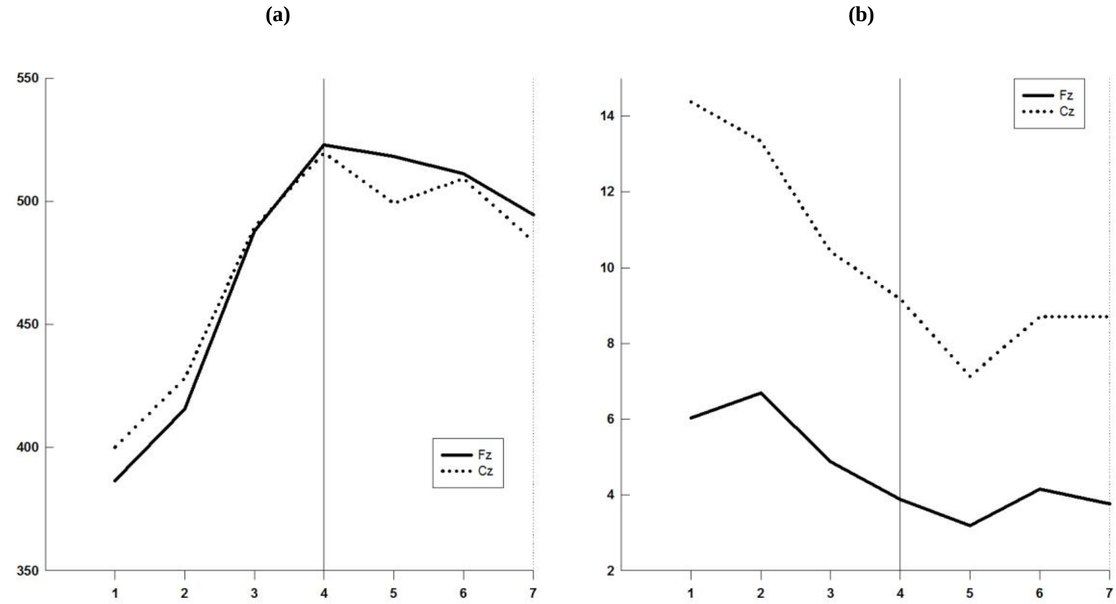
Component	Start(ms)	End(ms)	Type	Distribution
P2	90	250	+ve peak, latency / amplitude	1st peak: Fronto-central
N2	150	350	-ve peak, latency / amplitude	Fronto-central
P3b	350	650	+ve peak, latency / amplitude	2nd peak: Fronto-central

For each window, separate repeated measures analyses of variance (ANOVAs) are reported for the latency values and for the amplitude values. There are two within-subjects factors, site (*2) and item **(*7**). For all ANOVAs, when the sphericity assumption was violated Greenhouse-Geisser correction was used to assess significance. The uncorrected degrees of freedom are reported, along with the epsilon value (Picton et al., 2000). Partial eta-squared (η_p^2) is reported as a measure of effect size. For significant effects, planned comparisons are reported between successive item levels. During interpretation, when it is appropriate items one to four will be referred to as the subitizing *range*, because this includes the RT derived subitizing *span* of all subjects. Items five to seven will be referred to as the *counting range*.

The EEG analyses that follow are performed on pooled data from the entire sample. If there is a relationship between the RT derived subitizing span and the morphology of the LPC, as suggested by the Hypotheses being tested here, then it would be expected that the contributions from the two subitizing spans (3 items, and 4 items) represented in this sample at a ratio of 2:1 would be reflected in a dominant transition between three and four items.

4.3.3.1 P3b

Mean peak latencies and amplitudes across all item levels for the P3b component at Fz and Cz are presented in Figure 4.03. In these and all subsequent plots the thin solid vertical line at four items marks the boundary of the maximum subitizing range, defined by RT, recorded across the 12 participants, and the counting range.



P3b Latency

P3b Amplitude

Figure 4.03. Mean latencies (ms; panel a), and amplitudes (μv; panel b) of the P3b across arrays of one to seven elements, at sites Fz and Cz.

Latency. Examination of Figure 4.03, panel (a) shows a steady increase in P3b peak latency across item levels one to four. There was a medium sized effect for Item, F(6,66) = 8.3, p = .001, η_p^2 = .43. Latencies became longer between one and two items, and between two and three items, a pattern consistent across both sites. Planned comparisons showed that the P3b peak at Fz for two items was 29 ms later than for one item (p = .032), and the peak for three items was 72 ms later than the peak for two items (p = .008). The numerical increase of 35 ms between three and four items was not significant. At Cz, the peak for two items was 28 ms later than for one item (p = .001), and the peak for three items was 62 ms later than the peak for two items (p = .037). The further numeric increase between three and four items of 30 ms was not significant. However, polynomial contrasts on the latencies of one to four items, collapsed across sites, revealed a very strong linear trend, F(1,11) = 30, p < .001, η_p^2 = .732, which suggests that the latency increment to four items may be meaningful.

The latency data support Hypotheses 1. At both sites P3b peak latency (and thus, duration of the fronto-central LPC) increased at each step in the subitizing slope indicating that while subitizing is still occurring in this sample (according to the RT data) context updating in WM takes longer with the addition of *each* extra item.

The relationship between P3b latency and subitizing span is less clear, providing only general support for Hypothesis 2. The increment from two to three items, the final level subitized by two-thirds of the sample, is large and highly significant. The strong linear trend suggests meaningful latency increases up until four items, when latency levels off. Four items is the final level subitized by remaining one-third of the sample. Thus, there are latency increases that could be explained by weighted contributions from low- versus high-span subjects, if there is a definitive relationship between span, and capacity of the LPC. This issue is addressed in the span-group analyses later in the Chapter.

P3b Latency and RT. To test Hypothesis 3, the relationship of P3b latency to RT in the subitizing slope needs to be assessed. Thus, a percentage measure (%RT) was computed representing the change in P3b latency relative to the change in RT between successive item levels within the slope (1-3) and also for the following increment (3-4), using the formula: *(difference P3b latency / difference RT) * 100*. For example, if there was a change in P3b latency of +40 ms, and the change in RT was also +50 ms, then the sensitivity value would be 80%. As explained in Chapter 2, this measure can be seen as an indication of the sensitivity of P3b latency to task manipulations, as a function of the influence of these manipulations on RT, with the constraint (see section 2.3.6.1) that the exact percentage values derived from this measure should be interpreted within three relatively broad degrees of sensitivity: values between 0-32% indicate low sensitivity; those between 33-66% indicate medium sensitivity; while values between 67% and ≥100% indicate high sensitivity (see Verleger, 1997). Peak latency, RT, difference, and sensitivity values are presented in Table 4.03.

Table 4.03. P3b latency (ms), RT (ms), difference (ms), and sensitivity values.

Items	RT-a	RT-b	RT-Diff	P3b-a	P3b-b	P3b-Diff	%RT
(a v b)			**(b - a)**			**(b - a)**	
Fz							
1 v 2	416	428	12	387	416	29*	242
2 v 3	428	481	53	416	488	72*	136
3 v 4	481	611	130*	488	523	35	30
Cz							
1 v 2			12	400	428	28*	233
2 v 3			53	428	490	62*	117
3 v 4			130*	490	520	30	23

The critical relationship here is the step from two to three items because, as argued in section 1.2.2, it is this increment which defines the non-linear subitizing slope. In fact, P3b latency is highly sensitive to the addition of a third item, with increases which closely match the numerical RT increase of 53ms. However, there is no such relationship for the step from one to two items, where P3b latency increases almost twice as much as RT, although it should be kept in mind that the relationship is in the expected direction. Sensitivity decreases to low at four items as the large increase in RT driven by two-thirds of the sample entering the counting slope is accompanied by a much smaller, non-significant P3b latency increase of ~30ms.

These results provide only limited support for Hypothesis 3, however the sensitivity value for the step from two to three items, when processing load would be expected to be increasing, is consistent with the degree of sensitivity between P3b peak and RT that has been found in previous visual search studies (e.g., Luck & Hillyard, 1990; Wijers et al., 1987).

Latency Summary: In conjunction with the RT data the behaviour of P3b latency is consistent with the observations made in section 2.1.5.1.1 regarding the capacity of the LPC in terms of a processing episode, and the relationship that this might have with the transition from subitizing to counting. P3b latency stayed static from four items on, at around 520 ms, while RT increased by 326 ms at the step to five items when the entire sample had ceased subitizing, and continued to increase by similar magnitude at each increment thereafter. However, at three items, when all of the 12 subjects were still subitizing, RT and P3b latency were different by only 7 ms. Additionally, at three items, within the RT derived subitizing span of this group, the duration of the LPC (measured at the P3b peak) was sensitive to the same task manipulations that affect RT, and this sensitivity diminished markedly at four items. These findings establish a general relationship between capacity of the LPC, increases in time to achieve context updating in WM (indexed by P3b latency), and the subitizing span of this sample. Interestingly, the significant peak latency differences between one item and two (~29 ms) and between two items and three (72 and 62 ms at Fz and Cz respectively) were not equivalent, presenting a similar non-linear pattern of increase as the RT slope (where the numerical differences did not reach statistical significance). Thus, in terms of the non-linear subitizing slope, the significant increases in P3b latency, along with the strong sensitivity of this increase to the RT increment at three items, may provide additional information suggesting the involvement of serial WM activity beyond list search in determining the subitizing slope of this group. If quantifying three items is functionally equivalent to quantifying two (i.e., a three item array is subitized), and part of this process involves time consuming memory scanning for the correct number word we should see an equivalent latency increase between two and three items as just one more list item needs to be scanned. This is not the case. There is ~40ms of processing time unaccounted for, which suggests that *at three items*, the final level of the subitizing slope and still within the subitizing span of this group, WM processing beyond scanning an ordered list of number names may be involved in determining the non-linear increase in P3b latency, and the associated non-significant increase in RT which defines the non-linear subitizing slope.

Amplitude. Examination of the amplitudes in Figure 4.03, panel (b) shows an apparently steady decrease in P3b peak amplitude across the item levels. There was a very strong main effect for site, with amplitude at Cz being 5.6μv greater than that at Fz, F(1,11) = 26.9, p = .001, η_p^2 = .728. This pattern is typical of the P3b (see section 2.3.6), and confirms the component's identification in relation to the response contingent item enumeration task used in this experiment. There was a medium size effect for item, F(6,66) = 6.5, p = .001, η_p^2 = .352, accounted for by a very strong site by item interaction, showing greater decrease at Cz, F(6,66) = 7.8, p = .001, η_p^2 = .592. Planned comparisons showed that from two to three items, the same item level where the major P3b latency increase occurred, there were significant decreases at Fz (of 1.8μv, p = .036), and at Cz (of 3μv, p = .004). No subsequent decreases were significant at Fz, however, at Cz there was a further significant decrease from four to five items (2μv, p = .049), the step at which the last of the sample ceased to subitize. These results provide conditional support for Hypothesis 4. At the critical step to three items where P3b peak latency showed major and final increase, and still within the subitizing span of this group, amplitude decreased substantially at both sites. As detailed in section 2.1.5.1 this combination of P3b properties is indicative of load related, effortful processing in WM during subitizing.

P3b Summary: There were significant latency increases at each step within the subitizing slope, and at three items the substantial latency increase was accompanied by a notable decrease in amplitude. Each significant latency increase within the slope was sensitive to the corresponding nonsignificant RT increase. Beyond the subitizing slope latency did not alter significantly, however amplitude at Cz showed a further decrease at the step to five items, when the last of the sample ceased to subitize, and the effortful processing associated with counting was being undertaken by the entire sample.

4.3.3.2 P2

Examination of the grand-average waves in Figure 4.02 shows that the P2 does not become strongly evident at either site until three items, the same item level at which the major latency increase in the P3b occurs. However, the P2 (and the following N2) are evident in clear vestigal form at Fz for even one item, and at two items at Cz there is also a vestigal P2 followed by an obvious ‘flattening’ of the waveform in the N2 window. Thus, analyses in the P2 window over the first three items are likely to reflect the development of the P2 (and the following N2 component) as the morphology of the LPC develops. Mean peak latencies and amplitudes across item levels for the P2 component at Fz and Cz are presented in Figure 4.04.

Latency. Examination of Figure 4.04, panel (a) shows an apparent decrease in latency of the most positive point in the P2 window between two and three items, most notable at Cz. There was a medium strength effect for item, F(6,66) = 5.1, p = .001, η_p^2 = .318. Planned comparisons showed that at Fz there was no difference between one item and two, or between two items and three. At Cz, the peak for three items fell 31 ms earlier than for two items (p = .034). There was no difference in latency between three items and four. Thus, latency of the most positive point in the P2 window reduced between two and three items at Cz, indicating the clear development of the component, and then remained stable, while at Fz the clear development of the peak in terms of latency movement is not reflected by significance testing, although the component is clearly evident at three items in the grand-average wave. Once established, the P2 peak latency remained stable across the subsequent item levels at both sites.

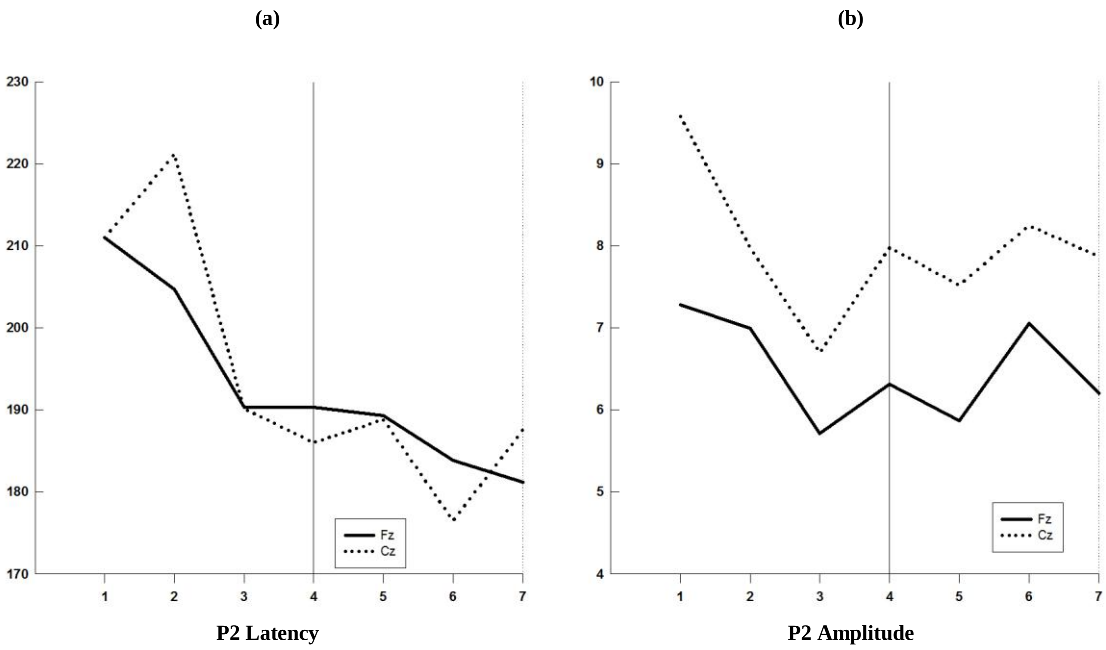


Figure 4.04. Mean latencies (ms; panel a), and amplitudes (μv; panel b) of the P2 across arrays of one to seven elements, at sites Fz and Cz.

In section 2.1.5.3 it was noted that P2 latency is longer when objects require individual identification or discrimination to assess relevance, or target status.

Because the simple dot stimuli used in the item enumeration task require no discrimination to be identified as targets—each dot is as relevant as the others and all dots have target status—the lack of any latency increase as a function of increasing number of items is understandable, and it indicates that the P2 does not reflect processes related to saliency, or discrimination. It also suggests that the timing of the process(es) generating the P2 in this paradigm are not prolonged by increasing perceptual, or processing load, an observation that will be discussed in more detail shortly when the P2-N2 complex is examined.

Another notion of the functional significance of a P2 component in WM demanding paradigms, introduced in section 2.1.5.4, was that (along with the following negative deflection) the P2 indexed some sort of load related WM processing. Under higher load P2 latency was shown to *decrease*. However, caution should be exercised interpreting (in isolation from the following N2) the decrease seen here at three items as an index of increasing WM load. In Missonnier et al. n-back paradigm the distinct P2 was evident in all conditions, including the 0-back condition when there was no WM load at all. Here, in response to the increasing item load during the item enumeration task, the *distinct* P2 was not evident until a load of three items, and the latency reduction at that level may be more a reflection of the devolution of the LPC because of the negative going activity of the following N2 then earlier occurrence of the process(es) generating the P2.

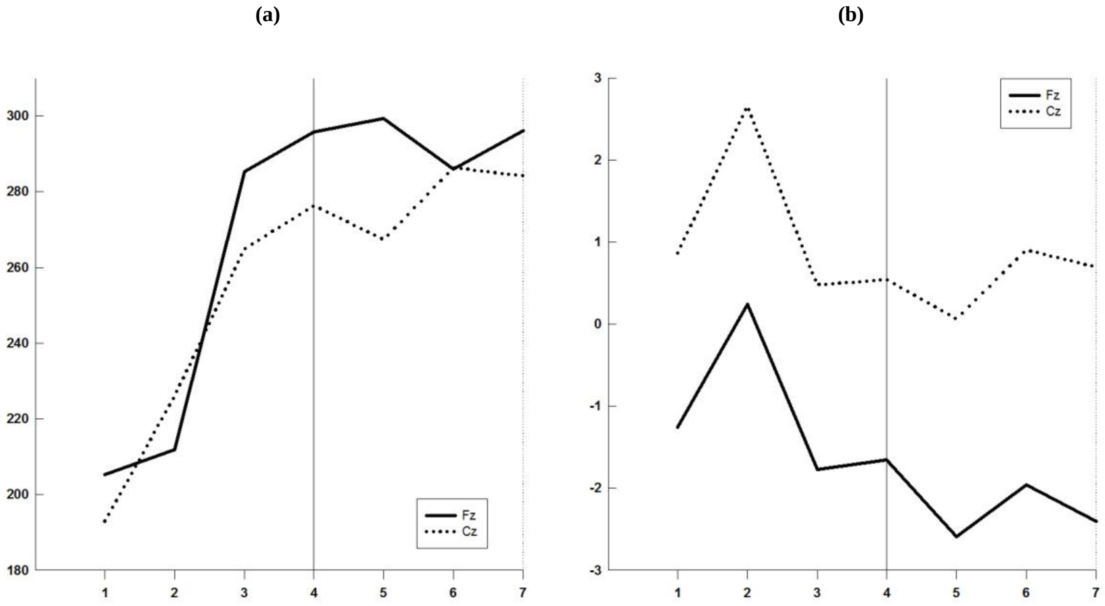
Finally, Hypothesis 5 suggested that if the duration of the process(es) generating the P2 were determining any prolongation in context updating across the subitizing slope then P2 latency should be sensitive to P3b latency increases. In fact, while P3b latency values increased significantly across the slope, the only change in P2 latency was the decrease at three items. Sensitivity measures revealed no relationship between P2 latency and P3b latency changes across the slope. Therefore, Hypothesis 5 was not supported. There is no evidence that timing of the early processing reflected by the P2 has any influence on subsequent delays in context updating which might influence delays in RT that constitute the subitizing slope.

Amplitude. Examination of Figure 4.04, panel (b) shows an apparent decrease in amplitude of the most positive point in the P2 window between two and three items, however at both sites amplitudes in the P2 window did not differ significantly across any item levels. Means and standard deviations are presented in Table 4.07.

P2 Summary: P2 latency decreased significantly at Cz at three items. Latency was not sensitive to P3b latency steps in the subitizing slope. There were no significant amplitude changes at either site.

4.3.3.3 N2

Examination of the grand-average waves in Figure 4.02 shows that, as for the P2, the N2 component becomes strongly evident at three items at both sites, although there is the definite suggestion of a vestigal N2 at Fz for both one and two items, and.at Cz for two items. Mean peak latencies and amplitudes for the N2 are presented in Figure 4.05.



N2 Latency

N2 Amplitude

Figure 4.05. Mean latencies (ms; panel a), and amplitudes (μv; panel b) of the N2 across arrays of one to seven elements, at sites Fz and Cz for each item level.

Latency. Examination of the latency plot in Figure 4.05 (panel a) shows a large apparent increase in latency at both sites between two and three items. There was a very strong effect for Item F(6,66) = 16.5, p = .001, $\eta_p^2 = .6$. Planned comparisons showed that at Fz the latencies for one and two items did not differ. However, at three items the peak was 74 ms later than for two items (p = .002). At Cz the peak for three items was later than for two items by 39 ms (p = .035). No other differences were significant. Thus, there is some support for the first part of Hypothesis 6. At both sites the N2 component became sharply defined at three items, reflected by the large latency increases. These significant changes occurred at the last step of the subitizing slope, still within the subitizing span of this group, and importantly, at the same item level where the P3b showed a large latency increase that was sensitive to the corresponding increment in RT.

Sensitivity to P3b. The significant N2 latency increase at three items at both sites is similar to the pattern of P3b behaviour at three items, indicating that the N2 may reflect the operation of extra, time consuming processing requirements which intercede between the processes reflected by the P2, and the culmination of context updating reflected by the P3b. If this is the case, the timing of this processing episode should be sensitive to the impact task manipulations have on P3b latency (which has already been shown to be sensitive to the prolongation of RT that occurs with the addition of a third item). To assess the relationship of N2 latency to P3b latency a sensitivity measure was computed (%-P3b): *difference N2 latency / difference P3b latency * 100*. The relevant values are presented in Table 4.04.

The sensitivity pattern differs between the two sites. For the step from one to two items the change in N2 latency at Fz showed little relationship to the corresponding change in P3b latency. In contrast, however, at Cz the N2 peak—vestigal as it was—was highly sensitive to the corresponding P3b latency increase, accounting for it entirely. At the step from two to three items, the N2 latency increase at Fz became highly sensitive to the corresponding prolongation of the P3b peak, while at Cz the N2 latency increase remained sensitive to the P3b increase. Thus, N2 latency at Cz shows considerable sensitivity to the P3b latency changes at both steps within the subitizing slope (and the span of this group). In contrast, N2 peak latency at Fz showed minimal sensitivity to the P3b at the step to two items, but the pronounced sensitivity increase at three items suggests a close relationship between the activity reflected by the N2 and the subsequent delay in context updating reflected by the increase in P3b latency. These findings support the second part of Hypothesis 6, suggesting that during item enumeration the N2 may index time consuming processing in WM that leads to increased context updating time.

Table 4.04. N2 latency (ms), difference (ms), and sensitivity values. P3b numeric difference values (ms) are included for comparison. Note: * denotes significant.

Items	P3b-Diff	N2-a	N2-b	N2-Diff	%-P3b
(a v b)				**(b - a)**	
Fz					
1 v 2	29*	205	212	7	25
2 v 3	72*	212	285	73*	**101**
3 v 4	35	285	296	11	-
Cz					
1 v 2	28*	194	226	33	118
2 v 3	62*	226	265	39*	**63**
3 v 4	30	265	276	11	-

Amplitude. Examination of Figure 4.05 panel (b) shows an apparent decrease in amplitude between two and three items at both sites. N2 amplitude was more positive by 2.5μv (p = .011) at Cz than at Fz, a medium sized effect, F(1,11) = 9.5, p = .011, η_p^2 = .428, and reduced as items increased, F(6,66) = 3.8, p = .003, η_p^2 = .259, which was a moderate sized effect. Amplitude changes across item levels were equivalent for both sites. Planned comparisons showed that at Fz there amplitude became more negative by 2μv from two to three items (p = .018), then remained stable across subsequent item levels. Similarly, at Cz there was a significant drop of 2.2μv between two and three items (p = .006), which remained stable across subsequent item levels. Thus, Hypothesis 7 was supported. N2 amplitude became sharply more negative at three items, coincident with the large decrease in P3b amplitude.

N2 Summary: N2 latency increased and amplitude decreased substantially at three items. No other changes were significant. The latency increments to three items showed high (Fz) and medium (Cz) sensitivity to the corresponding P3b latency increases. At Cz, the nonsignificant increase of 33 ms from one item to two was also sensitive to the corresponding significant P3b latency increase.

4.3.3.4 Interim Summary and Discussion

The data clearly illustrate a number of major changes in morphology of the LPC across item levels one to three that are consistent with the onset of effortful, time consuming WM processing operations within the subitizing slope, which is within the RT derived subitizing span of this sample as a whole.

The P3b. Examination of the P3b component was organised around four hypotheses that together reflected the information detailed in Chapter 2 regarding the types of behaviour the P3b might exhibit during subitizing were it reflecting effortful, time consuming processing. The latency data are consistent with the first three hypotheses. *Within the subitizing slope* P3b latency—and thus, time to achieve context updating in WM—was longer with the addition of *each* extra item. Existence of the subitizing slope has demanded attention in theories of subitizing and led to speculation regarding the list search explanation. The significant P3b latency increase at each step in the slope provides direct evidence of time consuming WM activity during subitizing. Based on the relationship that can exist between stimulus evaluation time and RT (detailed in section 2.1.5.1.1; see Kutas, McCarthy, & Donchin, 1977; Verleger, 1997) these latency increases allowed the list scanning explanation of the RT subitizing slope to be assessed by relating the magnitude of the (non-significant) RT increments comprising the subitizing slope to that of the significant P3b latency increases. The relationship between P3b latency increments and RT increments across the subitizing slope was consistent with the prediction made in Hypothesis 3. Similar to the non-linear RT profile, the large latency increase between two items and three was more than twice as much as that between one item and two, and the critical increment to three items was highly sensitive to the corresponding RT increment that defines the non-linear slope. In fact, the delay in response execution at this increment was accounted for entirely by the increase in context updating time reflected by the corresponding P3b latency increase at Cz. With the increment from one to two items serving as a measure of scan time for a single list item, the increment to three items had 39 ms (averaged across Fz and Cz) of processing time unaccounted for by list search[2]. These findings are not consistent with the list search explanation of the subitizing slope, which would assume context updating requirements to increase linearly across the slope. That they don't suggests that the delay in response execution at three items relative to two may be determined ultimately by time consuming—serial—processing in WM that is *additional* to any proposed serial list search. As explained in section 2.1.3, the simple architecture of the item enumeration task precludes recruitment of 'extraneous' information (e.g., automatically activated semantic associates) into WM as each additional unit of information (e.g., dot) adds to the task's demands purely by requiring quantification processes to operate over it. Thus, any additional processing evident beyond that which could be seen to serve list searching could only entail the operation of time consuming quantification processes. These findings thus implicate serial quantification activity by WM in determining the non-linear subitizing slope of this group, a position at odds with preattentive or parallel processing accounts of subitizing, and consistent with the serial counting explanation.

---

[2] Using the same logic on the non-significant RT increments leads to 41 ms unaccounted for.

A key claim of the serial counting explanation is that the subitizing limit is determined by a proliferation of binding and intermediate storage operations in WM, which eventually exceeds the activation available to complete chunking in a single processing episode—which in this thesis has been defined as a single episode of context updating in WM, with the episode's duration defined by P3b latency. The P3b behaviour assessed for Hypothesis 1 illustrates that context updating takes longer across the subitizing span of this sample, and the P3b sensitivity assessed for Hypothesis 3 is consistent with the idea that this type of serial WM processing may occur within the RT derived subitizing span. According to the serial counting explanation of subitizing limits, P3b latency should cease to increase at the *last* level of the subitizing span, a proposition that was assessed as Hypothesis 2, but that was not clearly supported by the data. P3b latency increased significantly until three items, the final item level of the sample's averaged subitizing span. Trend analysis then showed that the non-significant increment to four items may also have been meaningful. Both these results would be expected from the weighted contributions of waveforms from low- versus high-span subjects if there was a tight relationship between latency and subitizing span. Alternatively, they could reflect change up to four items that is largely independent of span. Thus, before a possible relationship can be confirmed between the *capacity* of the LPC, that is, the item level at which P3b latency ceases to increase, and the RT derived subitizing capacity, the span-group waves need to be examined. Overall the P3b latency behaviour is consistent with the suggestion that at three items and still within the subitizing span of the group serial WM activity subserving quantification is occurring, and that load on available WM resources may be associated with the limited subitizing span.

In line with Hypothesis 4, the increase in P3b latency across the subitizing range was accompanied by a substantial amplitude decrease at three items, a pattern consistent with the body of literature showing that P3b amplitude decreases with increased task difficulty defined operationally by prolongation of RT (detailed in section 2.1.5.1.2; see Kok, 2001), Along with the latency findings, this amplitude reduction at three items further supports the idea that serial WM resources are recruited within the RT derived subitizing span of this group.

The P2. The grand-average waves in Figure 4.02 show that at both sites the P2 component was only vestigal at one item, became slightly more defined at two items at Cz, and then fully developed at three items at both sites. At Fz there were no significant latency changes, while at Cz development of the P2 at three items was indicated by a decrease in latency, whereafter its latency did not alter, nor were there any significant amplitude changes at any item level. These findings are not consistent with an easy interpretation of the P2s functional significance as a discrete component (i.e., not as part of the P2-N2 complex). Interpreted independently, P2 like components have been shown to index selective attention to visual features, or stimulus salience, with increases in latency and amplitude taken to indicate preferential processing. The lack of latency or amplitude movements here do not provide support for the idea that in this experiment the P2 reflects the operation of a visual *selection* and input mechanism for WM, such as proposed by Potts (2004; 2001), or Luck and Hillyard (1994). Indeed, the item enumeration task does not require selection of items based on feature characteristics that are task relevant, as do paradigms generating the selection based P2. All items in an array are relevant, and all need to be ‘noticed’ for accurate enumeration to occur. As such, the P2 might be reflecting task relevant selection on the basis of activation of individuated item representations fed forward from preattentive visual processing mechanisms in extrastriate cortex (Poiese, Spalek, & Di Lollo, 2008), a proposition in line with Pylyshyn's FINST theory (2001). The early stage of processing in which the component is active is consistent with the operation of a visual input mechanism to WM, a possible function of FINST that was argued for in section 1.2.5.

There is some support for interpreting the marked change in LPC morphology at three items with the emergence of the P2-N2 complex as reflecting an increase in WM load early in the processing stream, as Missonnier et al. (2003) suggest and as outlined in section 2.1.5.4. Given the time frame over which the P2 component is active—it begins to rise around 100 ms post stimulus, around the time that initial visual processing in the extrastriate cortex and the subsequent feedforward sweep would be completed (Poiese, et al., 2008), and peaks at ~190 ms—it is conceivable that WM 'load' at this point in the processing stream is the loading of items into WM, that is, their activation. This could reflect an initial dynamic storage operation, prior to the activated representations being passed on to subsequent integrative processing operations in WM reflected firstly by the relatively unitary P3b, when there is only a load of one or two items, and then by the N2 and the P3b when item load increases to three and beyond. If this is the case, the clear emergence of the N2 following the P2 at a load of just three items, and the subsequent prolongation of the P3b peak, suggests that under the conditions of the item enumeration task the capacity of any input process reflected by the P2, and subsequent *initial* integrative processes leading to context updating, is just two units of information prior to load related demands recruiting additional, time consuming integrative processes reflected by the N2.

These data provide solid support for the view (detailed in section 2.1.5) that devolution of the LPC *under conditions of processing load* reflects the temporal separation of two previously summating positive components, with timing and amplitude of the first component, the P2 in this case, remaining stable while latency of the subsequent P3b peak increases and its amplitude decreases, due to the intercession of processing operations reflected by negative going activity in the 200-400 ms window post-stimulus. The stable latency and amplitude of the P2 suggests that whatever process or processes generate the component do not appear to take longer, or recruit more processing resources, when item load increases to three items and beyond, and the lack of P2 latency sensitivity to the P3b latency increment indicates that the prolongation of context updating at the final step in the subitizing slope is not the result of load related increases in the timing of this early activity. Instead, this delay in context updating must be determined by activity which follows the P2 in the processing stream.

The N2. Emergence of the N2 component at three items defines the P2 peak as a distinct component in this data, interceding between it and the delayed P3b peak. N2 latency increased as a function of item level, and the increases were sensitive to the corresponding P3b latency increases, notably at the critical step to three items when context updating was delayed substantially. This pattern of latency increases is important, indicating that the N2 reflects a process or set of processes that take longer to operate over three items than over two, and that occur prior to, and can account for, the prolongation of context updating for three items relative to two reflected by the increased P3b latency. In section 2.1.5.3 findings were presented suggesting that N2 latency might reflect the duration of evaluative processes as a function of the processing load required to perform them, and that resolving more difficult task rules, or increasing WM load can increase N2 latency, possibly because of the increased difficulty in integrating these representations into the current context. In this experiment the latency behaviour of the N2 in the subitizing slope is generally consistent with the predictions made in Hypotheses 6 and 7 regarding its possible role as an index of serial processing in WM, and thus WM load.

Additional support for this claim is provided by the amplitude data. N2 amplitude became more negative at the critical step to three items, and as is evident from the grand-average waves, remained negative for a considerable period after the N2 peak. Based on the literature relating the onset of slow negative waves in the 200 to 400 ms range to effortful processing by WM (Coulson, King, & Kutas, 1998; Gunter, Friederici, & Hahne, 1999; Hahne & Friederici, 2002; Hoen & Dominey, 2000; King & Kutas, 1995), it is plausible to claim that the N2 reflects serial binding activity in WM. In terms of the rationale developed in Chapters 1 and 2 the only binding operations required to perform the item enumeration task involve quantification operations over the number of individual units of information, and the binding of the final quantification unit with its number name. Thus, within the RT derived subitizing span of this group any WM activity that might be indexed by the N2 that can’t be explained through serial search through an ordered list of number names must reflect the operation of time consuming quantification processes over the units of information. By this rationale, the appearance of the N2 at a load of three units of information—the final level of the subitizing slope, and within the RT derived subitizing span of this group—can be seen as signifying the activity of effortful, time consuming processing in WM which subserves quantification.

Summary. At a load of just two items (i.e., with the addition of just one extra item) the P3b peak, and thus context updating in WM, was delayed significantly. As processing load increased to three items—the last level of the slope and *still within the RT derived subitizing span of this group*—there was a major change in the wave’s morphology, with further prolongation of the P3b peak and a sharp reduction in its amplitude, accompanied by the appearance of components earlier in the LPC, the P2 and the N2, with latency of the N2 also increasing substantially, as its amplitude became sharply more negative. Both the P3b and the N2 latencies were sensitive to the corresponding (non-linear) increase in RT, and N2 latency was also highly sensitive to the P3b increase. However, timing of the P2 peak had no relationship to either the P3b latency increase (and thus increase in context updating time in WM), or with RT.

Implications. Overall, the group data provides support for the serial counting explanation of subitizing, not parallel based or preattentive accounts. It was argued that the P2 might reflect an input process in WM that operates in a dedicated time-frame—even at one and two items when the component has not been differentiated from the LPC by appearance of the obvious N2 at three items. Following this, the N2 and delayed P3b reflect load related time consuming activity in WM that completes quantification operations, and enumeration. In this interpretation serial WM activity would determine the subitizing slope, and constraints on capacity to engage in this activity during the course of a processing episode may be implicated in determining the limited subitizing span.

The serial counting explanation offers a perspective on the type of WM capacity that may be reflected by the subitizing limitation—that is, a limitation on the activation available to keep representations active and to perform binding operations on them in the course of a single processing episode—chunking capacity. In the current data chunking capacity may be reflected in the systematic ways both the P3b and the N2 change as a function of increasing item (information) load, which has been interpreted as possibly reflecting serial WM involvement in enumerating three items, within the RT derived subitizing span of this group. The close relationship between RT and the P3b peak latency at three items suggests that the binding required to complete the episode of enumeration occurs within the duration of the LPC, whereas at four items the P3b peak is minimally prolonged, however RT increments substantially, making it likely that all required processing is not completed within the LPC duration.

At first glance this concept seems to add little to the RT information. However, it reinterprets the classic WM explanation of subitizing as being determined by STM constraints (Klahr, 1973; also see Cowan, 2001), that is storage *per se*, in terms of a capacity limit that is not based solely on the number of units of information that can be kept active, but also on the complexity of the task relevant binding operations which must be implemented over those units. It should be noted here that it is not the intention in this thesis to *measure* binding capacity (certainly with any precision beyond the integer measure of the RT span), but to use the concept as an inferential tool. For now, the group data exhibit five properties of the LPC which may define an operational measure of chunking capacity. The first five appear at three items, the final item level of the group's span (i.e., while RT defined subitizing is still occurring): 1) a large increase in P3b peak latency; 2) a large decrease in P3b amplitude; 3) the appearance of pronounced P2-N2 deflections; 4) high sensitivity to RT of P3b latency; and 5) high sensitivity of N2 latency to P3b latency. The predominant feature of this morphology is the occurrence of the major negative deflection which was operationalised as the N2 *peak* in this experiment. However it is obvious from the grand-average waves in Figure 4.03 that the deflection has a long duration beyond the N2 peak prior to the rise to the reduced amplitude P3b. Given that the P2 becomes defined by the sharp negative deflection of the N2, and the P3b that follows shows reduced amplitude, it is reasonable to argue that the dominant feature of the index of chunking capacity is this lengthy negative deflection, and that it represents evidence of the coming online of additional load related WM resources.

4.3.3.5 Span Groups

In the rationale to this chapter it was suggested that if differences in subitizing span were related to differences in load related processing capacity, then any LPC indicators of this capacity should differ between low- and hi-span participants as some function of item level. To test for group differences in the devolution of the LPC, within each window the peak latency and amplitude values were assessed by separate Item (4) by Site (2) repeated measures ANOVAs with Span Group as a between subjects factor. Span-Group had no significant effects on the P2, or on N2 or P3b latency, and no effect on Site. However, there were some interesting Span-Group related effects on both the N2 and the P3b amplitude, which will be discussed below in reference to the average waves for the two groups.

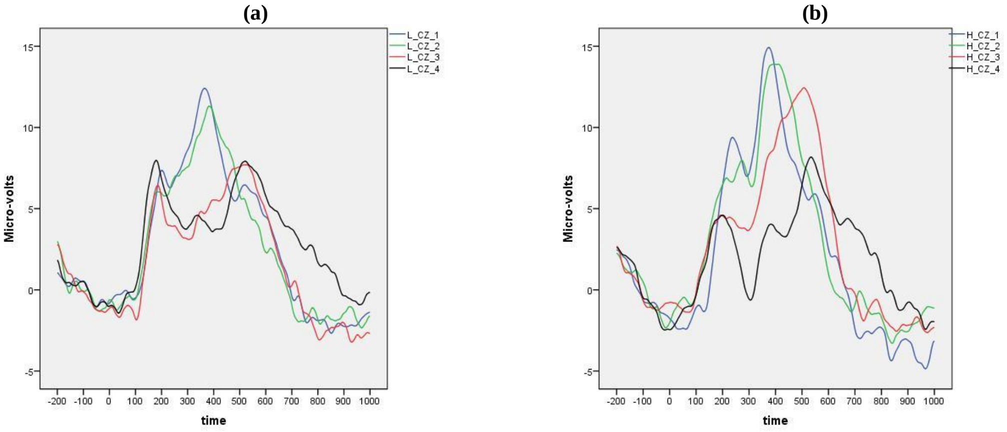


Figure 4.06. Average waves at Cz for item levels 1 to 4 for low-span (panel a) and high-span (panel b) groups. Blue trace = 1 item; Green = 2; Red = 3; Black = 4 items.

The grand-average waves for the low-span (span = 3; n = 8) and the high-span (span = 4; n = 4) groups are presented in Figure 4.06. The similarity between the groups in morphology across the four item levels is immediately obvious, an indication of a robust effect considering half the number of observations contribute to the high-span group's waves. To begin with it is clear that at Cz both groups show the same pattern of apparent prolongation of the LPC for two items relative to one item. This suggests that the significant increase in P3b latency from one to two items seen in the full group analyses was not driven by any differences in subitizing capacity, but reflects some sort of time consuming processing that occurs within the subitizing capacity and that is common to both span groups. ANOVA results confirm this. There was no significant effect for Span-Group on P3b latency—rather, both low- and high-span groups showed a significant increase from one to two items (24 ms, p = .03; 39 ms, p = .015, respectively).

At three items the addition of an extra unit of information is associated with observable amplitude reduction for both span-groups, with divergence from the waves for two items clearly beginning around 200 ms post-stimulus. For the low-span group there does not appear to be any further reduction at four items. In contrast, there is a further marked reduction in the high-span group. This pattern is captured by significant, moderate Item by Span-Group interactions for amplitude of both the N2, F(3,30) = 3.1, p = .04, η_p^2 = .234, and the P3b, F(3,30) = 3.1, p = .04, η_p^2 = .239. For the N2, amplitude decreases (by 2.7μv; p = .003) from two to three items for the low-span group and does not change significantly thereafter, while for the high-span group the numeric amplitude decrease of 0.9μv from two to three items did not reach significance, while the decrease of 2.4μv (p = .032) at the step to four items did. This interaction is described by a moderately strong quadratic contrast, F(1,10) = 5.8, p =

.036, $\eta_p^2 = .368$, which is illustrated in Figure 4.07, panel (a), however, the true nature of this interaction may not be made clear by the significance testing of the pairwise comparisons, which suggests that it occurs by *discrete* amplitude decreases at three items for the low-span group, and four items for the high-span group. When trend analyses are examined for the high-span group's amplitude behaviour from item two through to item four a different picture emerges. Amplitude for the high-span group is characterised by a very strong *linear* trend, $F(1,7) = 11.1$, $p = .045$, $\eta_p^2 = .787$, which suggests that the decrease between two and three items may reflect a real amplitude drop. Thus, instead of describing a discrete effect, this interaction may be describing a distributed effect on the N2 amplitude as a function of item load, and subitizing span, with change in the N2 amplitude occurring over two item levels for the high-span group, and over only one, overlapping level (i.e., three items) for the low-span group.

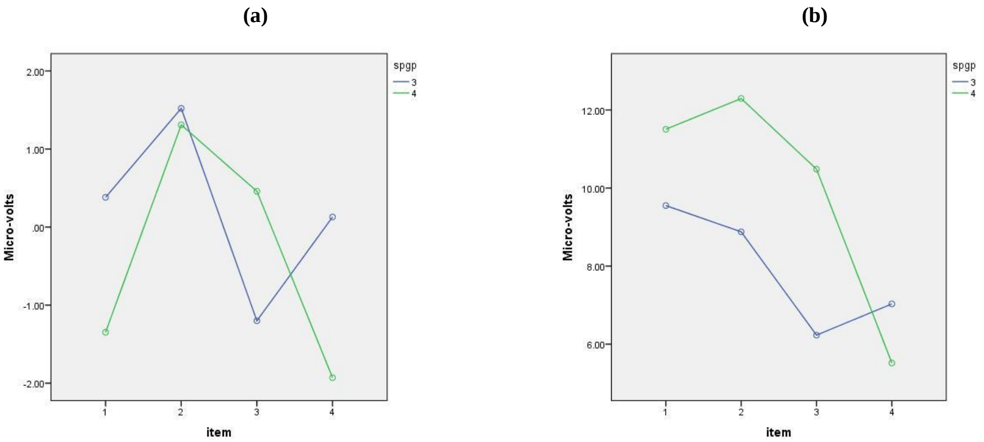


Figure 4.07. Span-group by Item interaction plots for item levels 1 to 4 for N2 amplitude (panel a) and P3b amplitude (panel b). Blue trace = low-span (3 items); Green trace = high-span (4 items).

This same pattern can descibe the interaction effect for the P3b amplitude, which decreases (by 2.6μv; $p = .02$) from two to three items for the low-span group and ceases to decrease thereafter, while for the high-span group the numeric decrease of 1.8μv between two and three items does not reach significance, while the decrease of 4.9μv ($p = .005$) at the step to four items does. This interaction is also described by a strong quadratic contrast, $F(1,10) = 8.1$, $p = .018$, $\eta_p^2 = .446$, which is illustrated in Figure 4.07, panel (b). As for the N2 amplitude, trend analyses show that the apparent steady amplitude decreases from item level two to four for the high-span group which are illustrated in the interaction plot may reflect real decreases, with the pattern described by a very strong linear trend, $F(1,7) = 21.6$, $p = .019$, $\eta_p^2 = .878$. Thus, as was surmised for the N2 interaction, instead of describing a discrete effect, this interaction may be describing a distributed effect on the P3b amplitude as a function of item load, and subitizing span, with change in the P3b amplitude occurring over two item levels for the high-span group, and over only one, overlapping level (i.e., three items) for the low-span group.

These interactions provide some interesting information regarding both the LPC profile associated with subitizing being compromised, and the possible functional significance of these amplitude reductions which begin around 200 ms post-stimulus. Firstly, they support the arguments used with the whole sample’s data that pointed to amplitude reduction in the N2 and P3b windows as being the most robust indicator of whatever extra processing demands additional units of information recruit during enumeration, and within the subitizing span. Additionally, the strong relationship between negative going activity in the LPC beginning around 200 ms post-stimulus, extending into the P3b window, and effortful cognitive processing activity suggests that changes in this amplitude related to differences in subitizing span may implicate differences in capacity for this processing activity in determining span differences.

If this is the case then interpreting the interactions as reflecting distributed rather than discrete changes in the underlying cognitive processes reflected by the LPC is consistent with proposals made in the serial counting explanation regarding chunking capacity. In section 1.1.2 chunking capacity was defined as involving the momentary activation of representations and the swift binding or integration of these representations during a processing episode. An important feature proposed of chunking capacity is that the process of integration within a processing episode does not have to proceed in parallel (see sections 1.1.2 and 2.1.5.1.1). The fact that the duration of the LPC can be prolonged by task difficulty while processing required for task completion can still be finalised within this duration (as shown by response occurring close to the time of the P3b peak) indicates that a processing episode has some capacity to absorb an amount of time consuming—serial—processing prior to chunking capacity being exceeded. For the high-span group, gradated amplitude reduction across two item levels versus a single step reduction for the low-spans could reflect a better capacity to absorb the additional binding operations demanded by a third, and then a fourth unit of information.

Finally, while there were no effects of Span-group on P3b latency, to properly address the possible relationship between P3b latency and RT derived subitizing span (Hypothesis 2), separately for each group pairwise comparisons and polynomial contrasts were examined. The high-span group had the largest overall latency increase from one to four items (151 ms v. 117 ms), and largest increment from two to three items (85 ms v. 58 ms). However, both groups' increases only approached significance at three items (low: 58 ms, p = .077; high: 85 ms, p = .085), and didn't reach it at four items (low: 35 ms, n.s.; high: 27 ms, n.s.). In contrast though, the steady numeric increases within each group are described by a strong linear contrast for the low-spans $F(1,3) = 11.1$, $p = .013$, $\eta_p^2 = .613$; and an extremely strong linear contrast for the high-spans $F(1,11) = 387.6$, $p < .001$, $\eta_p^2 = .992$. Given the limited *n* the large effect sizes for both groups give some justification for interpreting the steady latency increases as meaningfully related to increases in item load. However, it is clear that the pattern of these increases is the same across groups (see Figure 4.08) and there is no evidence of latency differences which would support an indicative relationship between latency and span.

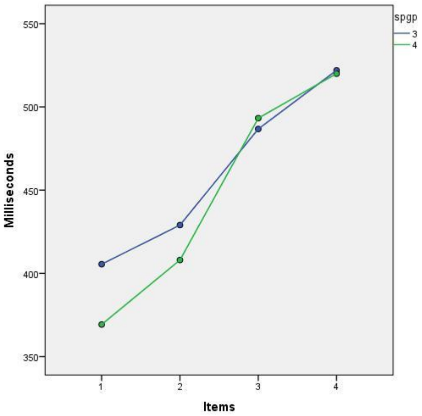


Figure 4.08. Span-group by Item interaction plot for item levels 1 to 4 for P3b latency. Blue trace = low-span (3 items); Green trace = high-span (4 items).

4.3.4 Conclusions

This experiment has shown that within the RT derived subitizing span the LPC changes morphology in a systematic way in response to the manipulation of the item enumeration task—incrementing the number of units of information that must be accurately quantified, and enumerated. In terms of the details presented in Chapter 2, the behaviour in this experiment of the LPC components as item level increased across the subitizing slope is consistent with what would be expected from a paradigm that increases processing demands on WM. P3b latency increase, P3b amplitude decrease, and the appearance of slow anterior negative deflections in the 200-400 ms range have all been suggested as electrophysiological markers of time consuming, effortful WM activity under processing load. In this experiment P3b latency increases, showing that it takes longer to accomplish the necessary operations to enumerate an array of two items than for a single item, and longer again for an array of three items than for one of two items. This lengthy increase when information load rises from two to three units is preceded by a substantial negative deflection at both the frontal and central sites, and is accompanied by a substantial decrease in amplitude of the delayed P3b.

Further evidence relating this pattern of modulations in the LPC to serial WM activity was found in the timing relationships across the subitizing slope between the sample's RT profile and latencies of the P3b, the P2, and the N2. P3b latency increased in close relationship to the RT increases across the slope, suggesting an association between increases in context updating time and increases in time taken to select a response (access the correct number word). P2 peak latency was not associated with either P3b latency movements, or RT. It did not increase in latency across the slope at all. In contrast, the duration of the N2 deflection *did* show a relationship with both the P3b and RT increments to three items, with a latency increase that could account for the subsequent P3b increase, and the associated delay in RT. Thus, the processes reflected by the P2 do not take longer once item load increases, but context updating does, and this prolongation of the P3b peak can be accounted for by the intercession of other processes between the P2 and the P3b, which are reflected by the N2 deflection. These findings demonstrate devolution of the LPC at the final level of the subitizing slope in a way that is consistent with the pattern detailed in section 2.1.5.1.2, which could reflect an increase in WM processing load, and processing time, that is related to increases in task difficulty.

Importantly, the significant increases in context updating time across the slope were unequal, and matched closely the corresponding non-linear RT increments. With the P3b latency increment between one and two items as an estimate of the time cost to scan from slot one to slot two of a list, that is, scan a single increment through the list, the P3b latency increment to three items included ~40 ms of processing time that is unaccounted for by serial list search. In terms of the architecture of the item enumeration task it has been argued (in section 1.2.3) that any serial processing across the slope that cannot be accounted for by list search has to reflect the operation of time consuming quantification operations in WM. Because intercession of the N2 defines the devolution of the LPC, and can account for the increase in P3b latency, given its time course and appearance at fronto-central sites it is reasonable to suppose that in this experiment the N2 component reflects serial processing in WM that subserves quantification, and thus (in terms of the argument made in section 1.1.2.1) may be an indicator of capacity limits of any parallel processing mechanisms involved in subitizing.

Thus, the sample’s data indicates that changes in the LPC morphology across item levels one to three—the subitizing slope—are informative about the determinants of the non-linear RT slope. Specifically, they suggest that the defining feature of the non-linear subitizing slope—a numerically larger RT step between two items and three items than between one item and two—may be determined by serial processing in WM that subserves quantification, and thus that this type of WM activity occurs during subitizing, as it is defined by RT. From this it follows that the subitizing limit may be related to limitations in capacity to engage in serial binding and maintenance operations subserving quantification across a single episode of context updating, what the serial counting explanation refers to as chunking capacity.

However, this latter proposition received only limited support from the P3b latency data. While pairwise comparisons showed that there were no significant differences in P3b latency between three and four items, either in the combined data or that for the span groups, steady increases from one to four items were described by strong to very strong linear trends. This indicates that the P3b latency increases may be meaningful at the steps to both three *and* four items, for both low- and high-span individuals. The lack of correspondence between RT span and the final level at which P3b latency increases indicates that the conception of chunking capacity in terms of a tight relationship between subitizing span and the point at which duration of the LPC ceases to increase may be too idealised.

While these findings show that P3b latency might not be a clear indicator of the exact subitizing span, the interactions from the span-groups analyses between subitizing span and P3b and N2 amplitude suggest that the pattern of increasing negativie going activity across item levels two, three, and four across these two windows may be more informative. Data from the low-span group showed a single major amplitude reduction across these windows, at three items. In contrast, the high-span group showed a gradated reduction distributed across the steps to three, *and* to four items. Besides giving a clearer indication of the RT subitizing span than the latency data, given the robust association of amplitude decrease in these windows and increases in processing load these findings provide some support for a key notion of the serial counting explanation—that is, that effortful cognitive activity is involved during subitizing, and that this activity can ‘gradually’ tax effortful processing resources prior to immediate processing capacity being exceded.

Overall, the findings from this experiment do not support the list search explanation of the subitizing slope, nor associated explanations of subitizing that conceive the subitizing limit as determined by a limited capacity to process information in parallel, whether this be by preattentive mechanisms, or in WM. They provide evidence indicating that serial WM activity that includes quantification operations may determine the subitizing slope, and that limited capacity to perform such operations within the time course of an episode of context updating in WM may have some association with the subitizing limit, and thus are broadly consistent with the serial counting explanation of subitizing. Within this conception, the experiment has identified a possible ERP index of serial WM involvement during subitizing that is informative about the RT derived subitizing span, and which might be useful in further investigations of WM involvement in subitizing, and in determining what type of capacity limitations the subitizing span might reflect.

CHAPTER 5

Experiment 2: ERP Activity During Item Enumeration

5.1 Rationale, Aims, and Hypotheses

This experiment is a replication of Experiment 1 using the Neuroscan equipment descibed in Chapter 3. This provides greater scalp coverage, and so allows investigation of the proposed anterior distribution of the negative going activity defining devolution of the LPC across the subitizing slope, which was proposed theoretically in Chapter 2, and which was evident in the frontal and central sites examined in Experiment 1. The first analyses presented here are organised around a series of hypotheses generated from the findings in the first experiment, and are conducted on data combined from the entire sample. Following these analyses, differences in LPC behaviour between groups with subitizing spans of three, and of four items are assessed.

The sample as a whole in Experiment 1 showed evidence of subitizing, as defined by RT. As has been common in studies of item enumeration (see section 1.2), the subitizing slope was non-linear, with the RT increment from two to three items being numerically larger than that between one and two items. This feature of the RT profile of the group was used to interpret the ERP behaviour within the LPC in terms of relationships that were informative about the type of processing activity that might be implicated in generating the non-linear subitizing slope, and in perhaps determining the capacity limits reflected by the RT derived subitizing span.

The key ERP findings related to the RT defined subitizing behaviour of the sample in the subitizing slope were, firstly, that devolution of the unitary LPC into three distinct components—the P2, the N2, and the P3b—occurred. The only change in P2 latency was at three items when there was a decrease which may have been driven by the marked appearance of a substantial following N2. This decrease in P2 latency was not associated with any increased duration of context updating, or RT, and there were no changes at subsequent item levels in P2 latency or amplitude. However P3b latency did increase, at the steps from one to two items *and* from two to three items, reflecting an increase in context updating time with the addition of one, and of two extra items. N2 latency also increased but only at three items (and not thereafter). Secondly, with the addition of the third item the increase in P3b latency (context updating time) was sensitive to the numeric increase in RT, and the increase in N2 latency was sensitive to the increase in P3b latency, thus suggesting that the increase in context updating time was related to the delay in response, and that the prolongation of context updating could be explained by the coming online of earlier, time consuming activity reflected by the N2. Additionally, there were significant decreases in P3b amplitude as load increased to three items and beyond, and in N2 amplitude at the step to three items, and, more generally, increasing negativity across the N2 window up to four items. So, in terms of the detailed rationale developed in section 4.1, there were meaningful changes in morphology of the LPC across the subitizing slope, and so within the subitizing range of all participants.

Together, these findings described behaviour of the LPC that was interpreted as being consistent with the advent of serial WM activity in response to increasing item load within the subitizing slope, thus supporting the inference that time consuming activity in WM was involved during subitizing and weakening any appeal to preattentive or parallel mechanisms as determining subitizing. Furthermore, the magnitide of the P3b latency increase between two and three items was larger than that between one and two items, similar to the pattern of RT increments across the subitizing slope, a finding which suggested that the time consuming WM activity recruited to enumerate three items was beyond that required for list search alone. The sensitivity of P3b latency to the RT increment at the step to three items further suggested that this WM activity influenced the numeric delay in RT, thus supporting the serial counting explanation of WM activity during subitizing over the list search explanation in explaining the subitizing slope.

Finally, the critical changes in morphology of the LPC that have just been listed occurred at three items, the final level of the subitizing slope, and also the final level at which all subjects were still subitizing. However, at four items, when only four participants were still subitizing, there was not the continuation of these same changes in the waveforms. N2 latency did not increase significantly, as it did at three items, and while P3b latency showed a trend to increase, the increase was small compared to that between two and three items. If these timing behaviours at three items were indicators of the subitizing span of the majority of the sample then the lack of continuation of this change at four items could have been because of the relatively small weighted contribution of the high-span subjects to the analyses. However, examination of the data by span-group suggested that this was not clearly the case. The pattern of latency behaviour was very similar for both groups which indicated that the proposed relationship between duration of the LPC, chunking capacity, and subitizing span did not adequately predict span, and therefore did not hold true. However, there was evidence suggesting that process(es) reflected by amplitude reductions in the N2 and P3b window may operate differently for low-and high-spans with high-spans able to absorb processing requirements over two item levels rather than only one for the low-span group. Thus, the relationship between amplitude reduction pattern and subitizing span may be a more robust indicator of chunking capacity. In the waveforms this manifested as the process of amplitude reduction being distributed across two item levels

The main findings from Experiment 1 can be summarised around the ten hypotheses that follow. Despite the latency findings which didn’t support the role of P3b latency in chunking capacity, the same questions that were asked in Experiment 1 are asked here. Hypothesis 5, which addresses the issue of chunking capacity, makes predictions that are qualified in terms of the RT profile of the current sample, which differs from that in Experiment 1. In the current sample there is a 1:1 ratio of subjects with spans of three items versus four items, as opposed to a 2:1 ratio in the first experiment. Thus, based on the equal contribution to the sample’s combined data of those with capacity of three items, and those with a capacity of four items—whether any differences between them are discrete, or distributed as was found in the span-group data from Experiment 1—it might be expected that in the combined sample the proposed ERP indicators of capacity would show an altered profile to the first experiment, and that capacity related changes reflecting some transition to serial processing load might be evident at the steps to both three and to four items. This issue is elaborated on in sections 5.3.1, and 5.3.2, and addressed specifically by the span group analyses which follow the analyses on the entire sample.

Hypothesis 1: Latency of the most positive point in the P2 window should decrease as the P2 component separates from the unitary positive going LPC, and the following values are pulled negative by the N2. It is likely that there will be no further latency changes as additional items are added to the array.

Hypothesis 2: P2 amplitude should not change across the subitizing slope, or with the addition of further items beyond the slope.

Hypothesis 3: If the LPC is sensitive to increasing processing demands of the enumeration task within the subitizing slope then we should expect latency of the P3b peak to increase with item load, as context updating takes longer.

Hypothesis 4: If the serial counting explanation of the subitizing slope is to be supported over the list search explanation, as shown in Experiment 1, then context updating time, reflected by P3b latency, should increase at both steps of the subitizing slope—that is, between one and two items, and between two and three items. Additionally, given a non-linear subitizing slope in this sample, the P3b latency increment between two and three items should be larger than that between one and two items, as was indicated in Experiment 1.

Hypothesis 5: If chunking capacity is related to the RT derived subitizing span, then P3b latency should increase up to the final item level of the subitizing span. Given the 1:1 span ratio in the current sample a significant increase in P3b latency between two and three items, and also between three and four items would support this relationship.

Hypothesis 6: If prolonged context updating in WM is implicated in delays in response execution within the subitizing slope, then any P3b latency increase at the step to three items should be sensitive to the corresponding RT increment.

Hypothesis 7: If the anterior N2 reflects load related serial WM activity that precedes and prolongs the context updating during subitizing, as suggested from Experiment 1, then N2 latency should increase at the final step in the subitizing slope, that is, the step to three items. Given the RT profile of the group, N2 latency might also increase at the step to four items.

Hypothesis 8: If the anterior N2 reflects the duration of some load related processing activity in WM that is involved in context updating during subitizing, then any increases in N2 latency should be sensitive to any corresponding P3b latency increases (and so, to RT).

Hypothesis 9: If increased processing demands are involved in prolonged P3b latency, there should also be an associated drop in P3b amplitude as item load increases. This would be a further indicator of load related WM operations.

Hypothesis 10: If the N2 represents load related serial WM activity, then N2 amplitude should become more negative as WM load increases sufficiently to require serial activity.

5.2 Method

5.2.1 Participants

Thirteen adults from Armidale participated. All had normal or corrected to normal vision, and reported being right handed. The data from four participants was not included in analyses due to excessive artifact leaving insufficient trials at all item levels, and one subject file was corrupted due to computer malfunction. The final sample comprised eight adults (3 M, mean age = 38.25 +/- 12.1, range = 20 – 62).

5.2.2 Task

The item enumeration task used for this experiment was identical to that used in Experiment 1, with one exception. Once the trial was initiated and the fixation cross was displayed, the following blank screen prior to the array being drawn to screen was only displayed for 35 ms, instead of the 235 ms in Experiment 1. As in Experiment 1, there were 420 experimental trials in total, 42 at each item level. The experimental trials were preceded by 30 practice trials, 3 at each item level.

5.2.3 Procedure

The procedure for this experiment required attendance only at a single testing session, lasting approximately two and a half hours. This session proceeded as did session two in Experiment 1, with the addition of the 30 practice trials prior to the EEG recording. Participants were again instructed explicitly to enumerate the arrays and *respond as fast as possible, while maintaining accuracy*. Thus, both speed and accuracy were stressed, with the admonition to use the practice trials to freely develop enumeration strategies within these constraints.

5.2.4 EEG Recording

For this experiment the electrophysiological signals were recorded using the NuAmps equipment, detailed in section 3.1.2. The EEG and EOG were recorded using a 40 channel QuickCap fitted with 8 mm Ag/AgCl electrodes from 36 scalp sites: FP1 FP2 Fz F3 F4 F7 F8 FT7 FT8 FT9 FT10 FCz FC3 FC4 Cz C3 C4 CPz CP3 CP4 Pz P3 P4 Oz O1 O2 PO1 PO2 T3 T4 T5 T6 TP7 TP8 A1 A2 (according to the 10-20 system). The vertical EOG was recorded from electrodes placed at the infra- and supra- orbital ridges of the left eye, and the horizontal EOG was recorded with electrodes placed at the outer canthi of each eye. All channels were referenced to A2 and had an analogue bandpass of DC to 100Hz. Gain for all channels was fixed at 19 by the NuAmps hardware, corresponding to a voltage resolution of 0.063µv at the 22 bit A/D resolution of the NuAmps. The continuous EEG/EOG was sampled at 500 Hz and recorded to a hard disk along with stimulus event marking information.

5.2.5 EEG Data Reduction

For each subject the continuous EEG was re-referenced offline to an average of A1 and A2 before being digitally band-pass filtered at 0.1 Hz to 25 Hz, 24db roll-off, using a finite impulse response, zero phase shift filter. Epochs including a 200 ms prestimulus baseline and extending 1000 ms post stimulus onset were extracted, baseline corrected, and averaged per condition. Epochs where voltage in any channel (including the EOG channels) exceeded +/- 75 μv were excluded from the average, as were trials with an incorrect response, or where RT was below 300 ms or excessive (see data screening criteria, below). The baseline corrected epochs were then averaged per condition. Grand mean waveforms were then generated for the sample, their morphology was inspected relative to the waveforms from Experiment 1, and component windows derived. These are detailed below in section 5.3.2.

5.3 Results and Discussion

As in Experiment 1, the focus of this research is on identifying the ERP changes that might occur across the subitizing slope—that is, across the first three items—as well as across four, and five items where, as information load increases most people would be expected to make the transition from subitizing to counting, as indicated by significant RT costs. So, while the task in this experiment included arrays of from one to ten items, only the data from levels one to seven are analysed. Seven was chosen as the highest level because individual subitizing spans can vary up to five or even six items.

5.3.1 Behavioural Data

Across the 2351 experimental enumeration trials from the 8 participants there were only 12 errors, a rate of 0.5%, almost identical to Experiment 1 (0.6%). This confirmed that subjects concentrated on enumerating accurately as instructed. These trials were excluded from all analyses. Error rates will not be addressed further in these analyses.

RT data was screened for each subject as in Experiment 1. This criteria led to a further 23 trials with RT under 300 ms (1%) and 95 outlier trials (> 3SDs; 4%) being excluded. A further 27 trials (1.1%) were rejected due to timing aberrations (sample drop, see section 3.5.3), additionally, 205 trials (8.7%) with a correct response that were rejected from EEG analyses due to artifact were also identified and excluded from the RT analyses. Thus 15.3% of trials (362) were excluded for all reasons (compared to 8.8% in Experiment 1), leaving 1989 trials in total for the RT and EEG analyses, with an average of 284 trials at each item level.

Mean RTs per condition are presented in Table 5.01. The RTs for one, two, and three items are consistent with those from other enumeration studies (see Gallistel & Gelman, 1991; Mandler & Shebo, 1982, Experiment 5; Svenson & Sjöberg, 1983; Trick & Pylyshyn, 1993, 1994). Along with the very low error rate this illustrates that, as in Experiment 1, both the speed and accuracy instructions were adhered to by this sample.

Table 5.01. Mean RTs, standard deviations, and number of trials included for analyses at each item level.

Items	RT(ms)	SD(ms)	N (trials)
1	465	111	278
2	473	112	281
3	506	136	284
4	581	159	292
5	788	279	283
6	1095	391	290
7	1457	487	281

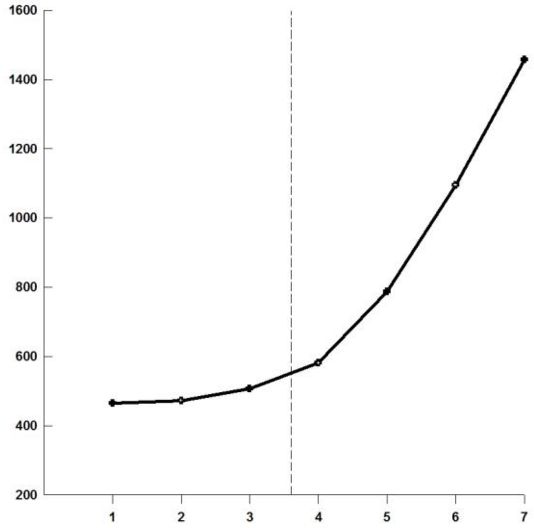


Figure 5.01. Mean RTs for each item level. The vertical dashed line in this and all subsequent plots is the RT derived subitizing span for the group.

As in Experiment 1 this sample showed a distinction between subitizing and counting. This is illustrated clearly in the means plot presented in Figure 5.01, which shows the classic ‘elbow’ in the function beginning noticeably at four items. The occurrence of subitizing in this sample was confirmed by calculating the subitizing span for each subject (by difference contrasts across item levels, as in Experiment 1). All subjects showed evidence of subitizing. Four subjects had spans of three items, and four subjects had spans of four items, with an average span for the group of 3.5 items, higher than the 3.33 item span of the group in Experiment 1. Over the three item range of the traditional subitizing slope the average time cost per item was 20.5 ms, less than the 32.5 ms per item in Experiment 1. As in Experiment 1 the increments between successive item levels within the slope were not equivalent. Thus, this sample also produced a non-linear subitizing slope. The consistency of this pattern with the wider subitizing literature again confirms the validity of the data produced from the button press response, relative to that obtained using a voice trigger. Additionally, the RT increments between item levels both within and beyond the slope differed from those in Experiment 1 in an interesting pattern which may reflect the different subitizing span profile between the two samples. The numeric difference in RT between successive item levels for Experiment 1, and for Experiment 2, and the differences in these increments between the two experiments are presented in Table 5.02.

Table 5.02: Average RT differences for increments between successive item levels for Experiment 1 (E-1), and Experiment 2 (E-2), and the difference between Experiment 2 and Experiment 1 at each increment.

Increment	Dif(ms) E-2 v E-1	Dif(ms) E-2 minus E-1
1-2	8 v 12	-4
2-3	33 v 53	-20
3-4	75 v 130	-55
4-5	207 v 326	-119
5-6	307 v 314	-7
6-7	362 v 347	+15

Examination of Table 5.02 shows that for the current experiment the increment from one to two items was almost identical to Experiment 1. However, at each subsequent increment up until the step between five and six items the values for Experiment 2 were increasingly *less* than those for Experiment 1, and it wasn’t until this step to six items that the cost per extra item reached a magnitude equivalent to the first experiment. This increment of ~300 ms in both experiments is consistent with the classic time cost associated with counting (e.g., Trick & Pylyshyn, 1993)—a cost that was established one item sooner, at the step from four to five items, for the group in

Experiment 1. This pattern may be explained by the composition of individual spans in the two samples. In the first experiment two thirds of the sample had a span of three items, with the remaining third all having a span of four items. In the current experiment *half* of the sample had a span of three items, while the remaining half had a span of four items. Thus, for the group in the current experiment the transition from subitizing to counting occurred across the increments from three to four items, and from four to five items, with half the sample counting at the step to four items. For the sample from the first experiment the transition occurred across the increments from three to four, and from four to five items, with the majority of the sample having moved into counting at the step to four items. The key point from this RT data is that, relative to the first experiment, the transition from subitizing to counting was more evenly weighted across item levels three and four, and counting was not *fully* established (i.e., with regular increments around 300 ms) until one item level later.

5.3.2 EEG Data

As in Experiment 1, due to the number of conditions and the distinction between subitizing and counting in this sample, the EEG grand-average waveforms are graphed over two overlapping item ranges—from one to four items, and from four to seven items. The grand-average waves for three midline electrodes—Fz, Cz, and Pz—are presented in Figure 5.02.

Given the association between RT and the behaviour of the LPC which was suggested by the results of Experiment 1 it might be expected that the different RT pattern in the current experiment would be reflected in the behaviour of the ERP components for this group. Specifically, if this is the case then in terms of the findings from Experiment 1 we might expect later occurrence of the notable transition in the waveforms that occurred between two and three items in Experiment 1—the clear emergence of the pronounced N2, and the preceding P2, along with the major increase in P3b latency, and reduction in its amplitude—might be less pronounced for this sample, and extend across the step to four items as well.

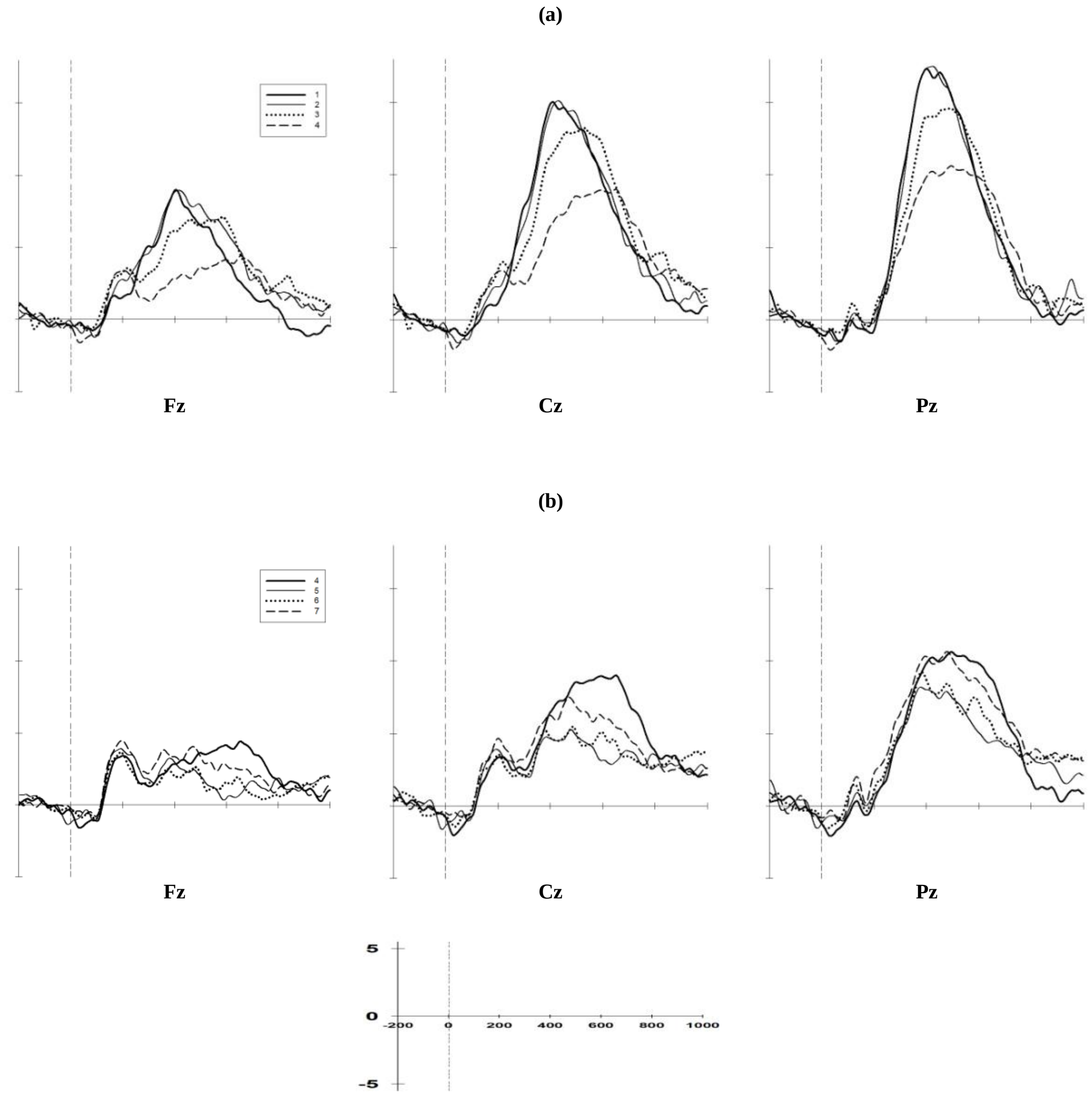


Figure 5.02. Grand average waves at Fz, Cz, and Pz for 1-4 items (panel a) and 4-7 items (panel b).

The waveforms in Figure 5.02 exhibit similar properties to those from Experiment 1. Panel (a) shows apparent latency increase and amplitude decrease of the P3b as item level increases from one to four at the three sites, and the associated appearance of the P2-N2 complex in the frontal and central electrodes, but not parietally. However, emergence of the P2 and the N2 is not nearly as pronounced at three items as it was in the first experiment. Consistent with the RT profile described above, the N2 appears to become fully established at four items, though less markedly than in Experiment 1. The N2 appears to remain similar from four to seven items (see panel b). In contrast, there are obvious amplitude changes in the following P3b, most notably a large decrease between four and five items, and what appears to be an increase between six and seven items. This amplitude pattern is clearly evident parietally as well. After examination of the individual subject’s waves the component windows were slightly adjusted from Experiment 1. Component windows are summarised in Table 5.03.

Table 5.03. Components and component windows.

Component	Start(ms)	End(ms)	Type	Distribution
P2	90	250	+ve peak, latency / amplitude	1st peak: Fronto-central
N2	180	420	-ve peak, latency / amplitude	Fronto-central
P3b	350	750	+ve peak, latency / amplitude	2nd peak: Fronto-central

For each window separate repeated measures analyses of variance (ANOVAs) are reported for the latency values and for the amplitude values across seven levels of the within subject factor *Item*, on a subset of 18 sites input as six levels of the within-subjects factor anterior-posterior (*Antpos*): Frontal (*F*; F3 Fz F4), fronto-central (*FC*; FC3 FCz FC4), central (*C*; C3 Cz C4), centro-parietal (*CP*; CP3 CPz CP4), parietal (*P*; P3 Pz P4), and occipital (*O*; O1 Oz O2), and three levels of the within-subjects factor *Laterality:* Midline, left, and right, each comprising the six electrodes along the sagittal, and left- and right-parasagital planes respectively. For all ANOVAs, when the sphericity assumption was violated Greenhouse-Geisser correction was used to assess significance. The uncorrected degrees of freedom are reported, along with the epsilon value. Partial eta-squared (η_p^2) is reported as a measure of effect size. For main effects, and two way interactions, planned comparisons (simple contrasts) are reported on the relevant composite means. For any higher order interactions simple contrasts on the individual electrode means are examined. Polynomial trend analyses are also reported where informative. Only effects involving Item are reported.

5.3.2.1 P2

Latency. Mean P2 peak latencies for the anterior/posterior and laterality composites across all item levels are presented in Figure 5.03. There was a small Antpos by Item interaction, F(30,210) 1.7, p=.013, $\eta_p^2 = .2$. Between one and two items peak latency decreased by 34 ms (p=.038) at the central sites. No other differences were significant.

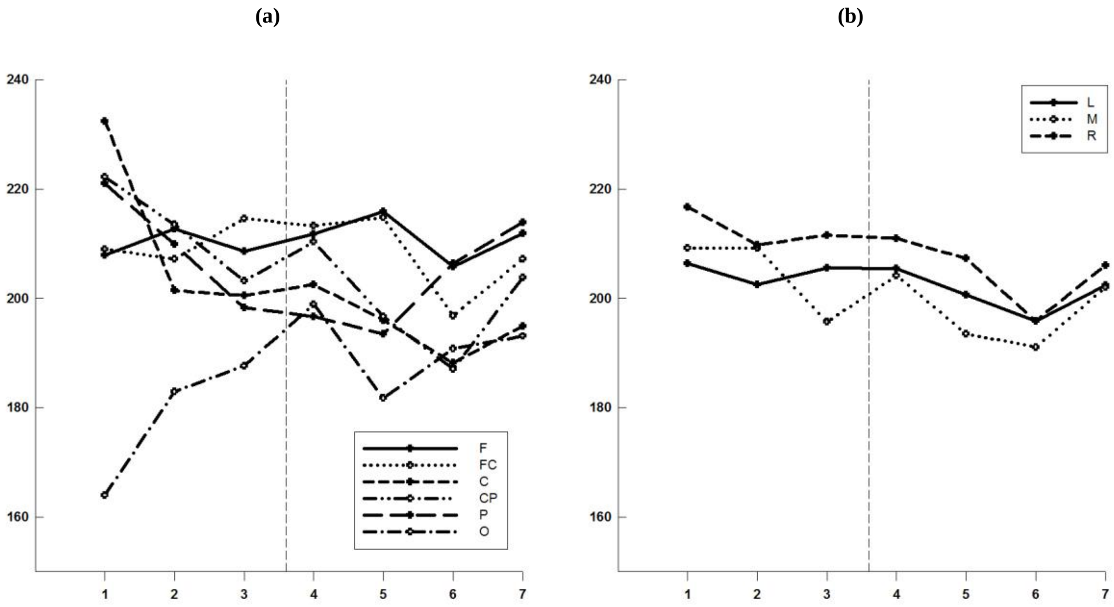


Figure 5.03. P2 mean latencies (ms): Anterior/Posterior (panel a), and Laterality (panel b), for all item levels.

Amplitude. Mean P2 peak amplitudes for the anterior/posterior and laterality composites across all item levels are presented in Figure 5.04. There were no significant effects for P2 amplitude involving Item.

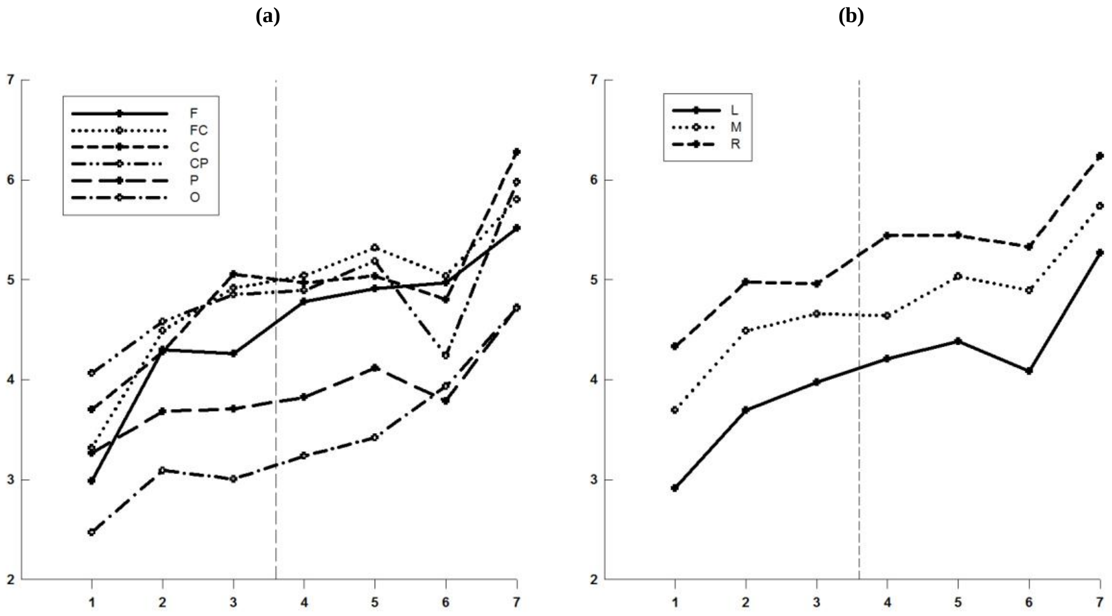


Figure 5.04. P2 mean amplitudes (μv): Anterior/Posterior (panel a), and Laterality (panel b), for all item levels.

5.3.2.2 P3b and N2 Latency

5.3.2.2.1 P3b

Mean P3b peak latencies for the anterior/posterior and laterality composites across all item levels are presented in Figure 5.05. There was a medium size effect for item, F(6,42) = 3.1, p = .014, η_p^2 = .305, embedded in a small Antpos by Item interaction, F(30,210) = 1.53, p = .045, η_p^2 = .18. All composites except Frontal sites, which approached significance, showed significant increases between two and three items (F: 25 ms, p=.083; FC: 51 ms, p = .039; C: 35 ms, p = .036; CP: 40 ms, p = .046; P: 33 ms, p = .042; O: 20 ms, p = .03), while at the step from one item to two only Frontal sites tended to show increased latency (66 ms, p = .07). Between three and four items there were a number of substantial increases but none were significant. However, at the step to five items at all but Parietal and Occipital sites there were substantial *decreases*.

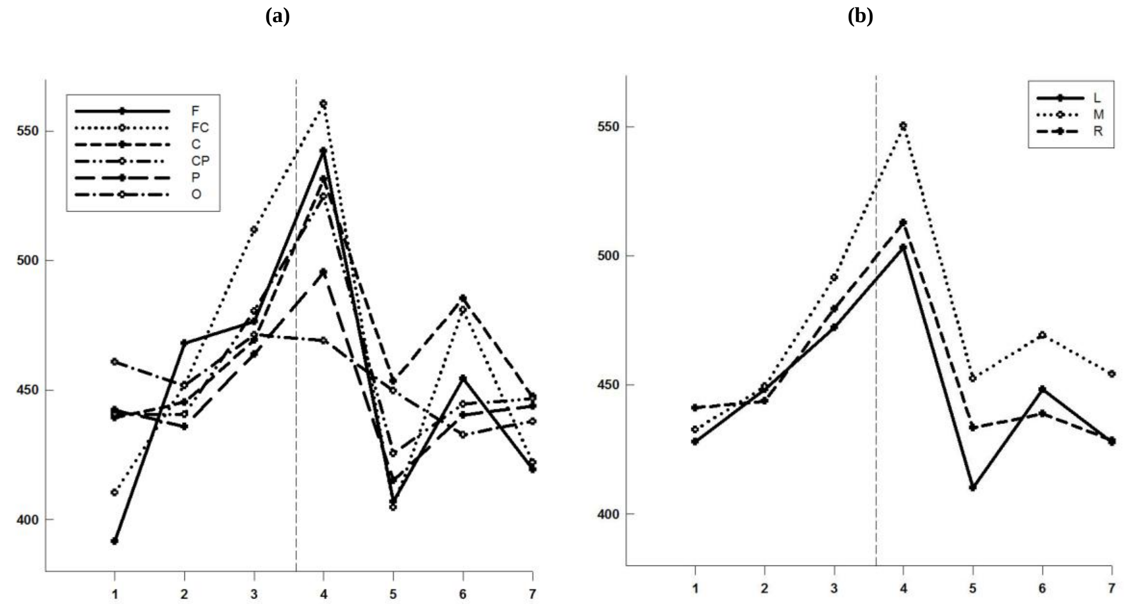


Figure 5.05. P3b mean latencies (ms): Anterior/Posterior (panel a; F=Frontal, FC=Fronto-central, C=Central, CP=Centro-parietal, P=Parietal, O=Occipital), and Laterality (panel b; L=Left, M=Midline, R=Right), for all item levels. In this and all subsequent means plots the thin dashed vertical line marks the subitizing span of the sample.

The latency decrease between four and five items was not seen in Experiment 1, where latency remained stable from four items on. However, in the current data at this step the four most anterior regions all showed a significant, large latency decrease, maximum at fronto-central sites (F: -116 ms, p=.033; FC: -143 ms, p = .005; C: -96 ms, p = .05; CP: -99 ms, p = .031), while the numeric latency decreases at parietal (-68 ms) and occipital (-19 ms) sites did not reach significance. Examination of the grand-average waves in Figure 5.02 shows that at five items at frontal and central sites there is a noticable change in morphology from ~ 400 ms through to ~700 ms post-stimulus,where the waveforms stay considerably more negative relative to four items, and there is no clearly defined late positive peak. This pattern is repeated at six items. While such a pattern may be problematic regarding the concept of chunking capacity derived from the first experiment, it could reflect differences between this sample and that from Experiment 1 in terms of the ‘group and quantify’ strategy proposed as a mechanism used in counting (see section 1.2). For now the focus is on activity over the first four items.

Examination of Figure 5.05 panel (a) shows what appears to be a considerable latency increase at frontal and fronto-central sites between one and two items. An increase between one and two items at the frontal or fronto-central sites would be consistent with the findings from Experiment 1. However, although the numeric increase at frontal sites of 65 ms was large, it only approached significance (p = .07), and the 32 ms numeric increase at fronto-central sites was not significant (p = .174). Secondly, at the step between three and four items there is an apparent increase that seems far less pronounced at parietel and occipital sites. Between three and four items there were large numeric latency increases at all composites except occipital but none reached statistical significance (F: 51 ms; FC: 60 ms; C: 61 ms; CP: 44 ms; P: 32 ms; O: 0 ms). In general these values were equivalent to (or larger, in the case of the central sites) than the significant increases found between two and three items (F: 25 ms, p = .083; FC: 51 ms, p = .039; C: 35 ms p = .036; CP: 40 ms, p = .046; P: 33 ms, p = .042; O: 20 ms, p = .03). That these relatively large numeric latency increases from three to four items, particularly anteriorly, did not reach significance may be explained by a sharp increase in variance of the timing of the peak between three and four items. The standard deviations for the F, FC, and C composites increase considerably from three (80, 80, and 81 ms, respectively) to four items (149,136, and 116 ms, respectively) indicating greater variance in the peak timing at four items, as might be expected with effectively two prototypical RT span waveform profiles—one for a three item capacity, the other for a four item capacity—contributing to the average.

Thus, the ANOVA findings suggest that latencies for one item are the same as those for two items, and those for four items are the same as for three items. However, examination of Figure 5.05 panel (a) shows an apparent pattern of sequential numeric latency increases from one to four items, particularly at each of the anterior composites, that is very similar to that observed in the plots for Fz and Cz in Experiment 1. Polynomial trend analyses were conducted *across items one to four* for each Antpos composite. The pattern of increase was described by significant, very strong linear contrasts for each composite except occipital, maximum at frontal, $F(1,7) = 11.7$, $p = .011$, $\eta_p^2 = .626$, and fronto-central sites, $F(1,7)$ 13.4, $p = .007$, $\eta_p^2 = .664$, but still strong through to the parietal sites (central: $F(1,7) = 7.4$, $p = .029$, $\eta_p^2 = .515$; centro-parietal: $F(1,7)$ 9.7, $p = .017$, $\eta_p^2 = .58$; parietal: $F(1,7)$ 5.7, $p = .049$, $\eta_p^2 = .447$). There was no significant trend occipitaly. These results are consistent with the strong linear trends found in Experiment 1.

Overall, these findings provide some support for Hypothesis 3—as in the first experiment, across one to four items context updating time showed a very strong tendency to increase as more items required processing, even though the only significant increase between successive item levels was between two and three items. This was the same level where the major morphology changes in the LPC occurred in Experiment 1. However, Hypothesis 4 is not supported. In contrast to Experiment 1, P3b latency did not increase significantly between one item and two at any of the composites, although there were considerable *numeric* latency increases at frontal and fronto-central sites. In Experiment 1, P3b latency increased significantly at Cz from one to two items. At the subsequent step to three items, at both Fz and Cz there was a further significant latency increase, which was larger than that between one and two items. The latency increase between one and two items was used as an estimate of the time taken to scan a single slot in an ordered list of number names. At the step to three items the increase in context updating time beyond that accounted for by this estimate of list scanning time was the key evidence supporting the inference that the non-linear subitizing slope was determined by increases in context updating time that reflected serial processing in WM beyond that required for list search alone, and so provided support for the serial counting explanation of subitizing. Clearly, this logic can not be supported by the current P3b latency data.

Finally, Hypothesis 5 also was not supported. Based on the RT profile of this group it was suggested that context updating time—reflected by P3b latency—might increase significantly between item levels two and three, *and* three and four, reflecting the higher subitizing capacity of this group relative to the sample in Experiment 1, which might suggest a greater chunking capacity. While the trend analyses suggest a steady latency increase up to four items, pairwise comparisons showed that following the significant increases between two and three items, the numeric latency increases between three and four items were not statistically significant. This pattern of significant linear trends to four items, but no significant difference in the pairwise comparisons between three and four items is the same as that found in Experiment 1. Given the RT profile of this group this weakens support for the proposed relationship between chunking capacity in terms of a tight relationship between P3b latency ceiling and RT derived subitizing span in the same way the span-group findings from the first experiment did.

Summary. The P3b latency findings in Experiment 1 were critical in the arguments supporting the serial counting explanation of the subitizing slope. The current findings do not immediately lend clear support to the serial counting explanation of the RT slope. Across items one to four the orthogonal contrasts of the trend analyses showed that context updating time in WM—reflected by P3b latency—increased steadily with the addition of each extra item, an effect that was strongest at anterior sites. This finding is consistent with that from the first experiment. Similar to Experiment 1, there was also evidence suggesting that context updating time in WM reliably increased between two and three items, still within the subitizing span of this group, a finding which supports the inference that time consuming WM activity occurs during subitizing. However, unlike Experiment 1, there was no clear support for there being nonequal increments in P3b latency between successive item levels over the subitizing slope. This finding does not allow the RT increment that defines the non-linear subitizing slope of this group—that between two and three items—to be interpretted as reflecting an increase in processing time in WM beyond that which would be expected from list search activity alone, as in Experiment 1, and so does not replicate the support found there for the serial counting explanation of the subitizing slope over the list scanning explanation.

In terms of chunking capacity, the overall greater subitizing capacity of this group relative to that from the first experiment was not reflected in any *significant* latency increases between three and four items as might be expected if the number of units of information that can be processed within the duration of the LPC is a deterrmining factor in the subitizing span. However, the very strong linear trends provide some support to interpreting the consistent numeric increases in latency as reflecting a progressive increase in context updating time up until four items. If this is the case then a relationship between these increments and the RT increments over the steps to three, and to four items would provide some support for a close relationship between P3b latency increases, chunking capacity, and the subitizing span.

Latency and RT. As in the first experiment, this sample's subitizing slope was defined by a larger numeric increment between two and three items than occurred between one and two items. However, the clear relationship in Experiment 1 between the magnitude of RT increases and increases in context updating time was not evident in the current P3b latency data, which did not show significant, non-equal increments across the subitizing slope. Nevertheless, a key property of the index of chunking capacity derived in Experiment 1 was P3b *sensitivity* to RT. The RT profile of this group differed from that in Experiment 1. As noted in section 5.3.2, half of this group had an RT span of four items or above, as opposed to a third with a span above three in the first experiment. Thus, if the pattern of P3b sensitivity to RT that was evident in Experiment 1 is truly reflective of a capacity limit associated with the RT derived subitizing span, then for this group the higher overall subitizing capacity—closer to four than to three—should be reflected in higher P3b sensitivity to the RT increment between three and four items, than that found in the first experiment.

To assess the relationship of P3b latency to RT, sensitivity values were computed for each of the six Antpos composites, following the procedure used in Experiment 1: *(difference P3b latency / difference RT) * 100*. The reader is reminded that values reflect three relatively broad degrees of sensitivity: values between 0-32% indicate low sensitivity; those between 33-66% indicate medium sensitivity; while values between 67% and ≥100 indicate high sensitivity. Sensitivity values were computed for each of the four steps (1 to 2 items, 2 to 3 items, 3 to 4 items, and 4 to five items). Values for the first step varied from positive to negative across the regions, ranging from ~30 to 100% at central sites, to > 180% at many others, the latter showing no meaningful relationship to the RT step. Values for the step to five items showed no sensitivity to the RT change, as would be expected from the significant latency *reduction* at this step. Thus, only the values for the steps from two to three, and three to four items are addressed. These values are plotted in Figure 5.06, panel (a).

Examination of Figure 5.06, panel (a), indicates initial support for Hypothesis 6, illustrating a relationship between P3b latency increases and the corresponding RT differences that is strongest at the four most anterior composites. P3b latency was highly sensitive to the increase in RT between two and three items at all regions except Fronto-central and Occipital (which showed moderate sensitivity), with a maximum at Central and Centro-parietal sites. The sensitivity of the average P3b peak over all electrodes was also very high at this step. Importantly, the P3b peak was also sensitive to RT at the step to four items, with a maximum at Frontal, Fronto-central, and Central sites, decreasing posteriorly. This more anterior distribution, relative to the previous step, is reflected in the decreased sensitivity of the *average* P3b peak at this step. Thus, as was predicted because of the distribution of this group's RT span, the significant increase in P3b latency at the step to three items was sensitive to the corresponding RT increase, *as was the non-significant numeric latency increase between three and four items*.

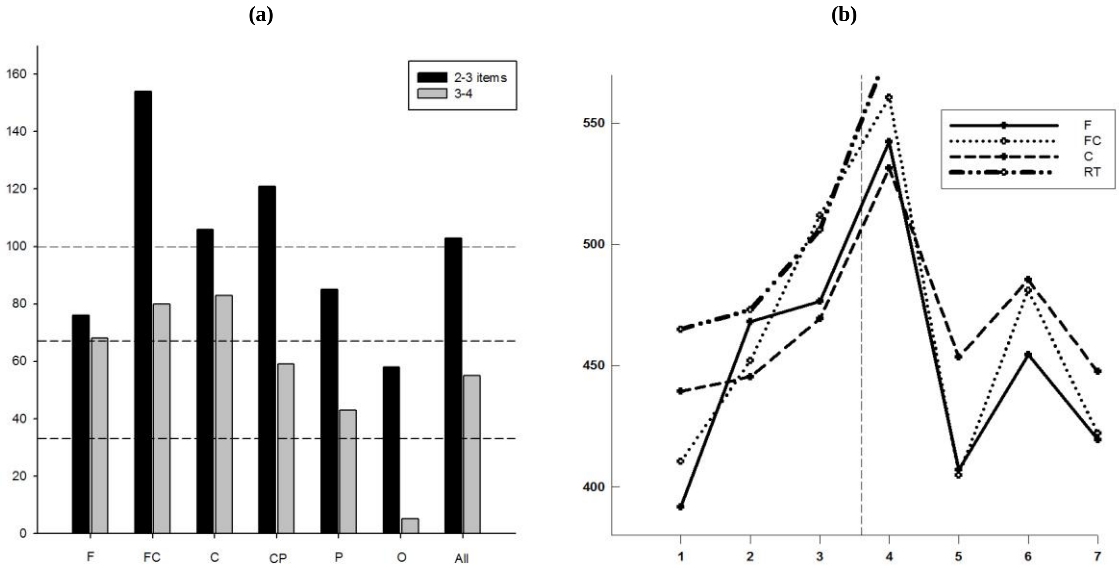


Figure 5.06. P3b to RT sensitivity values (%; panel a) for the six anterior/posterior composites, and for the mean latency over all 18 electrodes (All), for latency steps between 2-3, and 3-4 items. Percentage scores that were negative, or that exceeded 160% were recorded as a score of 5. Thus, the bars at 5% indicate NO sensitivity. Panel b plots latency (ms) of the F, FC, and C composites against RT.

These findings indicate a close relationship between delay of the P3b peak and increased RT at the steps to three, *and* to four items. In conjunction with the significant latency increase between two and three items, the high P3b sensitivity to RT at this step is consistent with the findings from Experiment 1. Context updating in WM takes longer for three items than for two, still within the subitizing span of all of this group, and this increase is associated with the corresponding numeric increase in RT, thus implicating time consuming WM activity in generating this segment of the non-linear subitizing slope. Given this finding it is reasonable to argue that the somewhat reduced sensitivity of the numeric P3b latency increase to the corresponding RT increment at the subsequent step to four items—still within the subitizing span of *half* of the group— may indicate that time consuming WM activity also contributes to the longer RT at this step. The sensitivity data then support interpreting the significant linear trend for P3b latency to increase over one to four items as reflecting a meaningful increase in context updating time at the step to four items. Thus, firstly, the numeric RT increase between three and four items may be determined by prolonged stimulus evaluation time involved in achieving context updating in WM and so may reflect a meaningful change in the duration of the processes which culminate in the P3b peak, and which may determine RT. Secondly, if this is the case then there may be some support, in this sample, for the relationship between duration of the LPC and the RT profile that defines the subitizing range. The relationship between P3b latency change between two, three, and four items at the anterior sites and RT across items two to four is well illustrated in Figure 5.06, panel b.

Summary. Together these data generally support the conclusion from Experiment 1 that P3b latency is sensitive to increasing processing requirements that also influence RT over the subitizing span. However, there is no support for the associated claim that the observed unequal increments in the RT subitizing slope may be determined by unequal increases in context updating time beyond those that would be expected from the requirement to scan an extra list entry with each extra item. Thus, there is little support for the serial counting explanation of the slope. Finally, at the step to four items Centro-parietal, Parietal, and Occipital sites show less sensitivity relative to the more anterior sites, which supports the premise that fronto-central changes in LPC morphology may be indicative of load related serial processing activity. If this is the case, then latency changes in the anterior N2 should follow a similar pattern to that in Experiment 1, increasing between two and three items, and the increase should be sensitive to the corresponding P3b step. Furthermore, given the P3b sensitivity to RT at the step to four items, and the RT span profile of the group, if this is the case then it might be expected that the N2 should also show a load related latency increase at four items.

5.3.2.2.2 N2

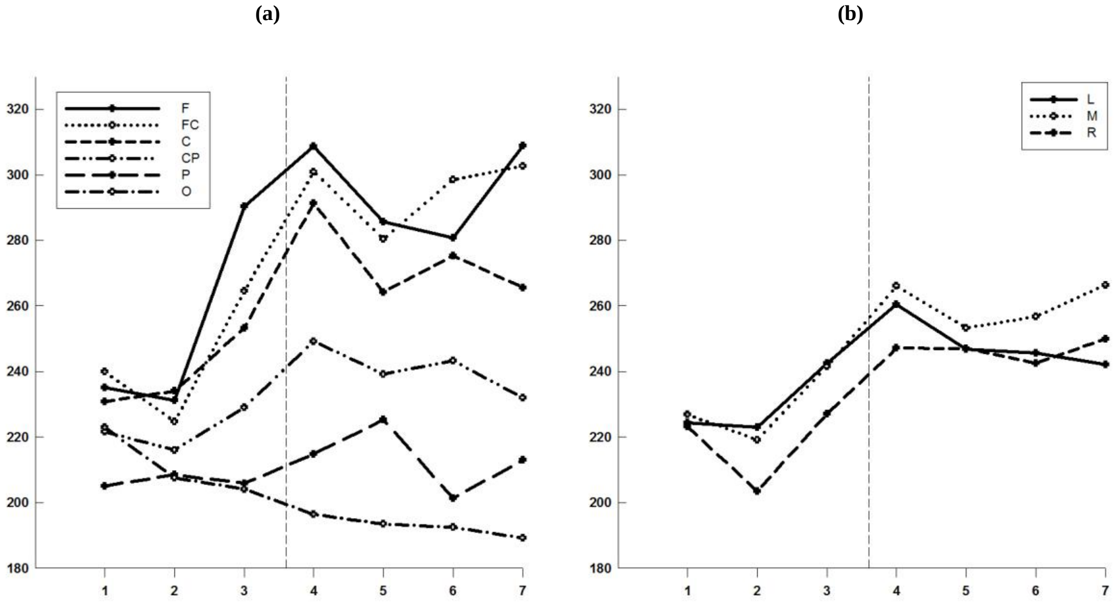


Figure 5.07. N2 mean latencies (ms): Anterior/Posterior (panel a), and Laterality (panel b), for all item levels.

Latency. Mean N2 peak latencies for the anterior/posterior and laterality composites across all item levels are presented in Figure 5.07. There was a significant medium size Antpos by Item interaction, F(30,210) = 4.0, p = .012, $\eta_p^2 = .366$, $\varepsilon = .124$, indicating that load related latency increases in the N2 window are predominant at the anterior regions. Between two and three items there were significant increases at frontal (59 ms, p=.049), fronto-central (36 ms, p=.043), sites, while the 21 ms numeric increase at central sites did not reach significance. Between three items and four there were significant increases at fronto-central (44 ms, p=.025), and central (38 ms, p=.045) sites. The numeric increase at frontal sites (29 ms) did not reach significance.

Overall, these results are consistent with Hypothesis 7—N2 peak latency increased significantly at the final step in the subitizing slope, and changed significantly at anterior, but not posterior sites. The Frontal / Fronto-central increase between two items and three was consistent with Experiment 1, where latency increased at both Fz and Cz. However, unlike Experiment 1, in the current data there was no evidence of any latency increase between one and two items in this window. Additionally, in Experiment 1, where the predominant RT subitizing span in the sample was three items (8/12 subjects), there were no significant N2 latency change between three and four items—that is, beyond the RT span of two thirds of the sample. However, in the current sample, where half had a span of four items or more, there *were* significant increases in N2 latency between three and four items at the Fronto-central and Central sites. In light of the relationship demonstrated in Experiment 1 between the timing of the N2 peak and the P3b latency increase at three items, the last item in the group's span, the additional significant N2 latency increase at four items in the current sample may signal further serial binding activity within the duration of the LPC by those subjects with a span of four. If this is the case, and it is this serial processing that is determining the predominantly anterior P3b latency increases that were shown to be sensitive to the RT increments between two and three *and* three and four items (even though the simple contrast tests did not indicate a significant latency increase between three and four items) then if the anterior N2 is an indicator of this serial processing it should exhibit sensitivity to the P3b increments at those two steps, each of which were shown to be sensitive to the corresponding RT increments.

Latency and RT. To assess the relationship of N2 latency to P3b latency, N2 sensitivity measures were computed for the steps to three, to four items, for the six

Antpos composites, following the procedure used in Experiment 1: (*difference N2 latency* / *difference P3b latency) * 100*. The relevant values are plotted in Figure 5.08.

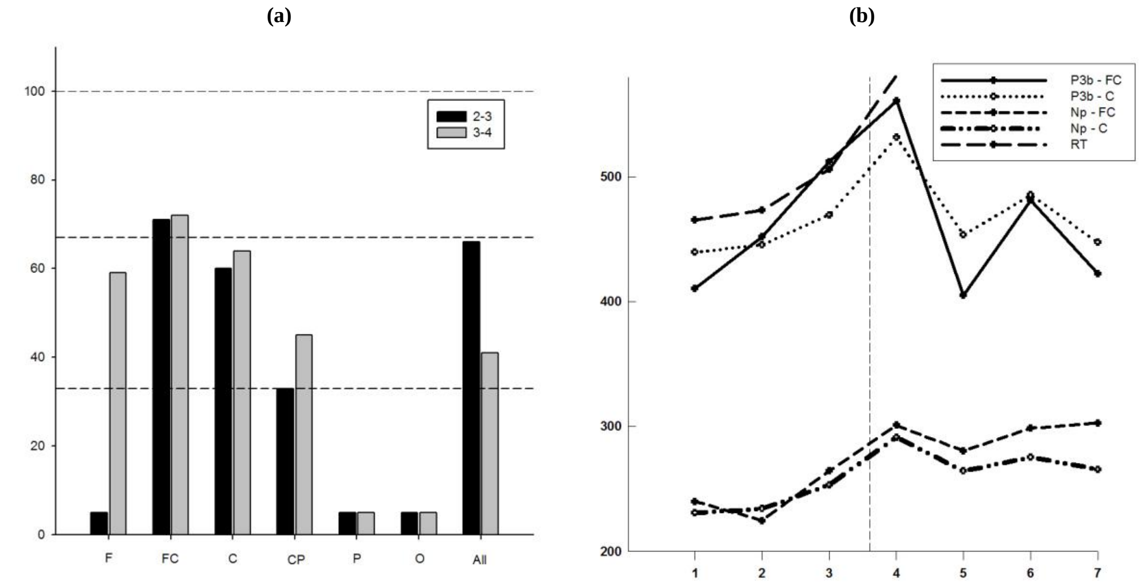


Figure 5.08. N2 to P3b sensitivity values (%; panel a) for the six anterior/posterior composites, and for the mean latency over all electrodes (All), for latency steps between 2-3, and 3-4 items. The bars at 5% indicate NO sensitivity. Percentage scores that were negative, or that exceeded 160 were recorded as a score of 5. Panel b (ms) plots latency of the FC and C composites for the P3b and the N2 against RT.

Examination of Figure 5.08, panel a, provides support for Hypothesis 8. At Fronto-central sites N2 latency was highly sensitive to the increase in P3b latency between two and three items, with a posterior going decrease from medium to low sensitivity at Central, and Centro-parietal sites respectively. Parietal and Occipital sites showed no sensitivity, driven by some large *negative* values. The lack of any relationship at frontal sites was also because of very high scores, positive in this case—N2 latency increased substantially more than P3b latency at these sites. The sensitivity of the average N2 peak latency over all electrodes was also very high at this step, which can be seen as an artifact of the averaging of these very large positive and negative values. It is obvious that its distribution is anterior. Importantly, the significant increase in N2 latency between three and four items was also sensitive to the numeric P3b latency increase at the step to four items, with values ranging from medium to high over Frontal to Centro-parietal sites, with a Fronto-central maximum. Note that at this step the Frontal sites become moderately sensitive. Posterior electrodes again showed no sensitivity. This more anterior N2 distribution, relative to the previous step, is reflected in the decreased sensitivity of the average N2 peak at this step. Thus, as was predicted N2 latency was sensitive to the significant P3b latency increase at the step to three items, and also to the numeric P3b increase between three and four items, at the anterior sites where the P3b showed the greatest sensitivity to the RT steps.

Additionally, in Experiment 1 an argument was made that the sensitivity of the N2 peak to the latency increase in the P3b peak between two and three items indicated that the N2 reflected the operation of serial WM processes, the duration of which contributed to prolongation of the P3b peak relative to two items. In terms of this inference, the current pattern of significant increases in N2 latency associated with strong N2 sensitivity to the P3b latencies at the steps to both three and four items would support the proposition made previously that for this sample the non-significant numeric increase in P3b latency at four items may in fact reflect a meaningful change in the duration of the processes which culminate in the P3b peak. This argument is functionally plausible in terms of the proposed role of the N2 as reflecting load related serial processing activity ‘coming online’, and its apparent capacity related development, illustrated in Experiment 1 through the the span group waves. It is also plausible physiologically that a process (or related collection of processes) increases intensity along a continuum. In this case the ‘processes’ reflected (at least in part) by the N2 are hypothesised to be capacity limited in terms of chunking capacity—that is, the amount of information that can be processed within the duration of the LPC—and thus the span profile, weighted towards four in this group, would explain the smaller RT step to four items than in Experiment 1, and the sensitivity of the P3b and the N2 at this step.

Overall, with the important exception of no significant differences between one and two items, the latency data for the P3b and the N2 are generally consistent with a pattern indicating serial WM activity within the subitizing span of this group, and thus with the conclusions from Experiment 1 regarding the operation of serial WM activity during subitizing. The data are not consistent with preattentive or parallel processing accounts. This supposition would be further strengthened by a pattern of P3b and N2 amplitude changes similar to that found in Experiment 1.

5.3.2.3 P3b and N2 Amplitude

5.3.2.3.1 P3b

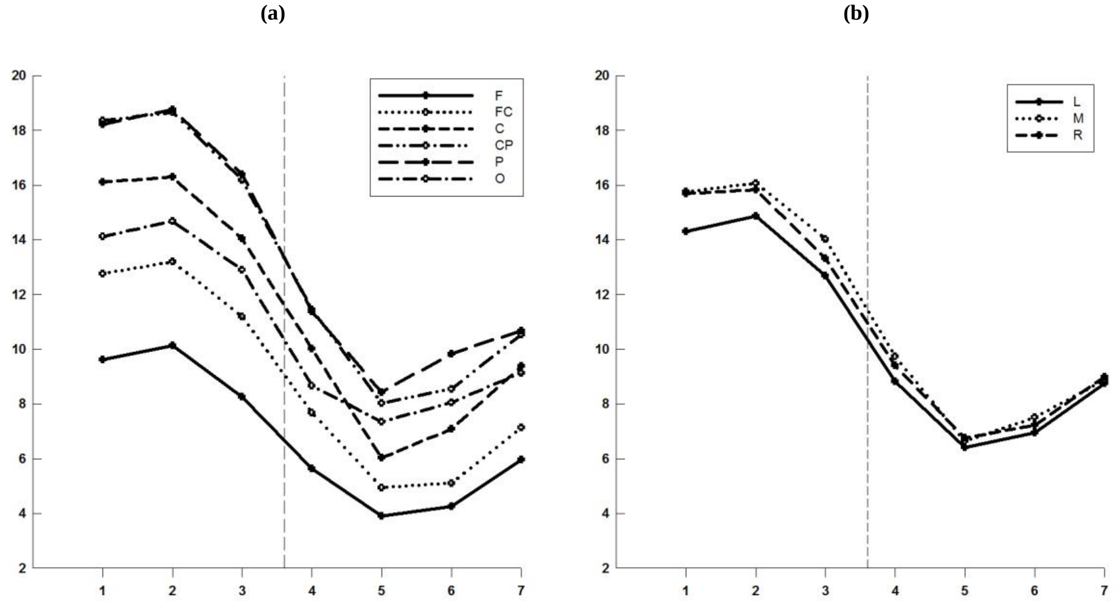


Figure 5.09. P3b mean amplitudes (ms): Anterior/Posterior (panel a), and Laterality (panel b), for all item levels.

Mean P3b peak amplitudes for the anterior/posterior and laterality composites across all item levels are presented in Figure 5.09. The results support Hypothesis 9. There was a very strong main effect for item, F(6,42) = 23.9, p = .001, η_p^2 = .773, ε = .26, showing that from three to five items amplitude decreased progressively, dropping by 2.2μv (p = .009, collapsed across sites) between two and three items, and by a further 4μv (p =.002) between three items and four, followed by another decrease of 2.7μv (p =.003) between four items and five. Overall, this pattern of decreasing P3b amplitude with increasing item load is consistent with that found in Experiment 1, where P3b amplitude decreased progressively at the steps to three, four, and five items, and thus with what would be expected with increasing load related processing in WM.

5.3.2.3.2 N2

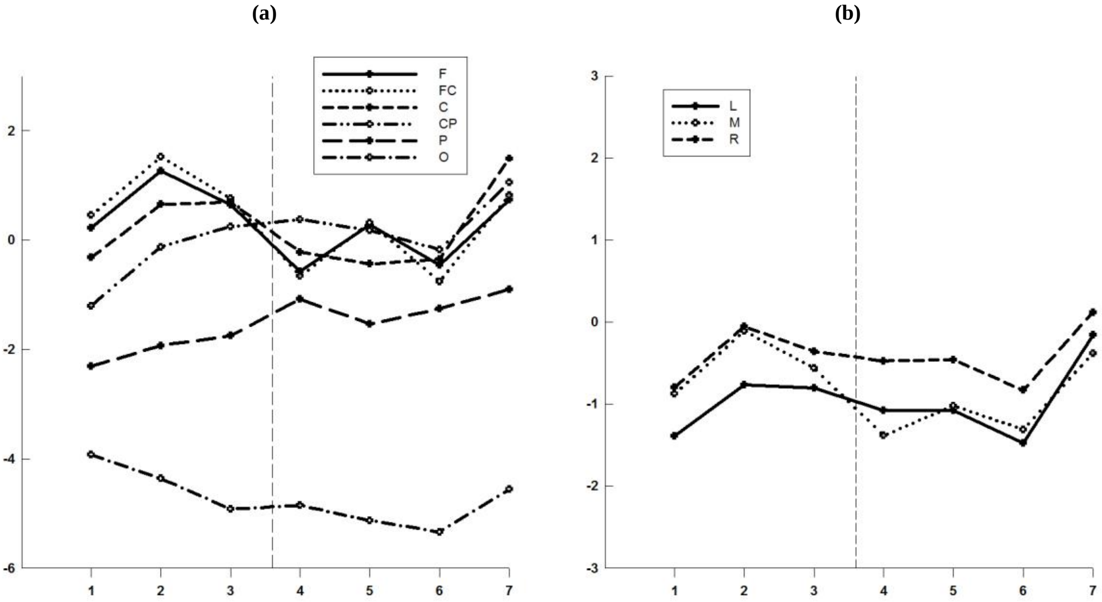


Figure 5.10. N2 mean amplitudes (ms): Anterior/Posterior (panel a), and Laterality (panel b), for all item levels.

Mean N2 peak amplitudes for the anterior/posterior and laterality composites across all item levels are presented in Figure 5.10. The only apparent increases in negativity are at the Frontal and Fronto-central sites from two to three, and three to four items. However, no effects involving item were significant, thus the results are not consistent with those from Experiment 1, and so provide no support for Hypothesis 10 . This finding is at odds with the supposition that the reduced P3b peak amplitude with increased item load is a result of the coming online of earlier, sustained negative going activity which reflects load related processing. However, while the N2 *peak* does not reflect an increase in negative amplitude, examination of the grand-average waves in Figure 5.02 shows an apparent decrease in amplitude at three items beginning around 200 ms and continuing through the N2 window. This pattern appears also for four items, and beyond, which would be consistent with that seen in Experiment 1. To assess this behaviour across the N2 window, the average amplitude for the first four items was extracted over a window from 180 ms to 420 ms. An Antpos (6) by Laterality (3) by Item (4) ANOVA showed a.significant, strong effect for Item, F(3,21) = 4.9, p=.009, η_p^2 = .414. Planned comparisons showed no change between one and two items. At the step to three, and to four items, amplitude reductions of 1.25 μv (p=.066) and 1.4 μv (p=.085), respectively, only approached significance. However, a pattern of relative amplitude reductions was described by a strong linear trend, F(1,7) = 6.0, p=.045, η_p^2 = .46, plotted in Figure 5.11. Thus, the average amplitude data also does not show significant reductions between successive conditions, although the strong linear trend provides some support for interpreting the numeric amplitude reductions between two and three, and three and four items as reflecting a sequential relative amplitude reduction in the waveforms in the period between the P2 peak, and the P3b peak.

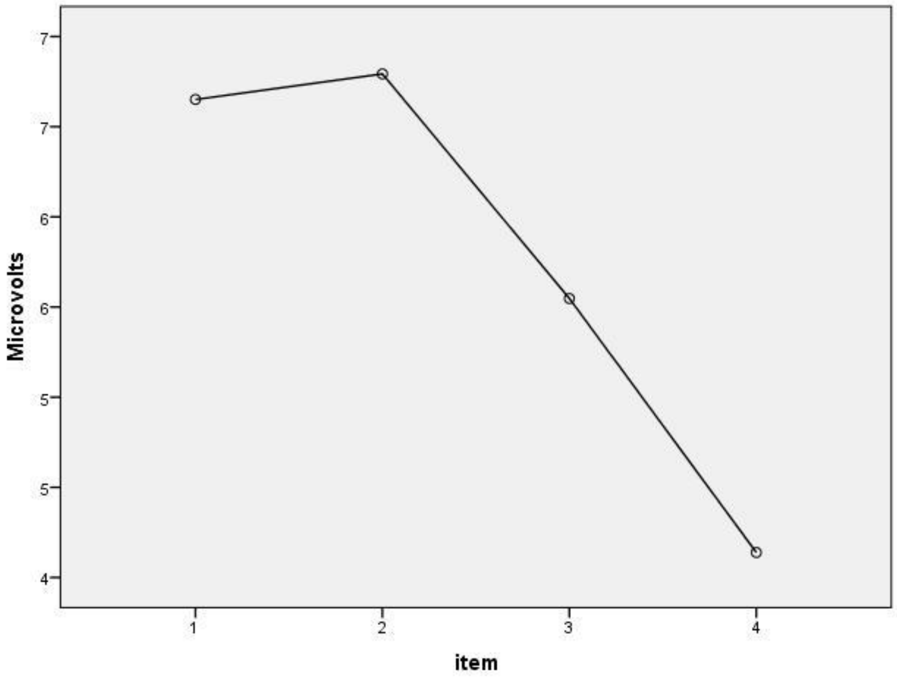


Figure 5.11. Main effect for Item for N2 mean amplitude.

5.3.2.4 Summary

The behaviour of the LPC across the subitizing slope in this experiment was broadly consistent with the findings from Experiment 1. Firstly, the LPC showed the same pattern of devolution, from a single dominant P3b peak of considerable amplitude under limited item load, to a P2-N2-P3b peak complex with reduced amplitude P3b as item load increased. The P2 showed no systematic changes in either latency or amplitude related to increasing item load across the slope, however, P3b latency *did* increase, at the step to three items, along with a reduction in P3b amplitude, and the clear appearance of a strong negative going deflection which separated and defined the two positive peaks. While there was also clear evidence of a latency increase in the N2 peak at this step, the peak was far less pronounced than in the first experiment, and did not show a dramatic amplitude increase at three items.

However, the negative deflection over the course of the N2 window and preceding the P3b peak showed some evidence of an amplitude reduction under a load of three, and of four items, though this was not as pronounced as in Experiment 1. As has been well noted already in this thesis, this overall pattern of LPC behaviour is consistent with that found in other paradigms where WM load supposedly is increased across conditions, leading to delays in response selection and ultimately, response execution. Its replication here strengthens the claim made in the first experiment that serial WM processes are recruited in the enumeration of just three items—when subitizing, as defined by RT, is still occurring.

Secondly, the increase in P3b peak latency at the step to three items was sensitive to the corresponding numeric RT increase. This supports the relationship proposed in the first experiment between the additional time taken to achieve context updating in WM, and the additional time taken to produce the response. Furthermore, at this step the increase in N2 peak latency at anterior sites was sensitive to the P3b latency increase, and this effect was strongest Fronto-centrally. Overall, these findings are consistent with the relationship proposed in the first experiment between the duration of time consuming activity in WM, beginning around 200 ms post-stimulus, and subsequent delay in context updating and response execution at the final step in the subitizing slope, as well as with the notion that the primary indicator of this activity is a predominantly anterior negative deflection in the waveforms, indexed here by the N2, and further defined by the subsequent amplitude reduction through to the P3b.

However, while these data are generally consistent with the findings from the first experiment, and thus confirm a pattern of ERP activity across the subitizing slope that is consistent with time consuming WM activity occurring during subitizing, in contrast to Experiment 1 there was no clear evidence that context updating time increased between one and two items. As such, an estimate of list scanning time for a single increment cannot be established from this data, which therefore can provide no support for the serial counting explanation of the slope over the list scanning explanation of WM activity across the subitizing slope.

Overall, these data are broadly consistent with the five properties of the LPC proposed from Experiment 1 as reflecting a chunking capacity in WM. At the step to three items, when all of the sample were still subitizing, there is clear evidence of a substantial increase in P3b latency, and a reduction in P3b amplitude. These are coincident with the clear emergence of the N2 component (though not as markedly as for the sample in the first experiment), and the preceding P2 peak, and with a strong relationship between increased latency of the N2 peak, and subsequent delay in context updating (and thus the delay in RT). At the step to four items, the final subitizing level for the remainder of the sample, there was a trend for both N2 and P3b latency to increase, and N2 relative amplitude to reduce, while P3b amplitude again decreased substantially. The N2 latency increment was sensitive to the P3b increment, which in turn was sensitive to the RT increment. Thus, while the changes may have been more minor, the pattern of change at four items was similar to that at three items. This pattern would be consistent with weighted contributions to the grand-average waves of a discrete change in morphology which was indicative of the RT subitizing span, and which included a final increase in P3b latency at the last level of the subitizing span. However, the span-group analyses from the first experiment suggested that latency increases from three to four items happened for both high- *and* low-span groups, and that it was the pattern of N2 and P3b amplitude reduction that differed between the span-groups.

5.3.2.5 Span groups

The grand-average waves for the low-span (span = 3; n = 4) and the high-span (span = 4; n = 4) groups are presented in Figure 5.11. It is immediately apparent that at three items *both span groups* exhibit a change in morphology—broadly evident as an apparent amplitude decrease and latency increase in the P3b window (see red versus blue and green traces in Figure 5.11)—that is consistent with the pattern that has so far been interpreted as reflecting a time consuming increase in processing load in WM. This broad pattern is also consistent with that seen for the high-span group in Experiment 1, where P3b amplitude reduced stepwise from two to three to four items. It is interesting that the low-span group in the current data also show this stepwise reduction, whereas the low-span group from the first experiment showed only a single step reduction from two to three items.

.

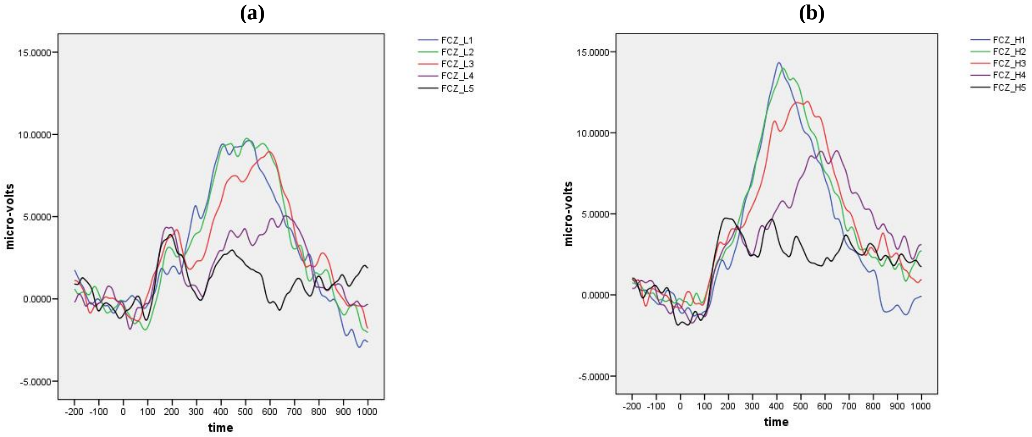


Figure 5.11. Average waves at FCz for item levels 1 to 5 for low-span (panel a) and high-span (panel b) groups.

There are also some clear differences between the group's waveforms. Firstly, at one, two, and three items the low-span group shows apparently less amplitude than the high-span group. Secondly, at one and two items the high-span group shows a prototypical unitary LPC with a sharply defined singular P3b peak, whereas the low-span group has a less defined P3b peak that plateaus for both one and two items. Thirdly, at three items the low-span group shows clear development of a P2-N2 complex, which becomes more substantial at four items, whereas for the high-span group the P2-N2 complex does not appear to become firmly established until four items. This latter observation is consistent with what would be expected if, as proposed from the analyses of the entire sample's data in this experiment, and Experiment 1, the appearance of a distinct P2-N2 deflection does reflect a load related processing capacity in WM that may be related to the subitizing limit.

To test for group differences in the devolution of the LPC, within each window the peak latency and amplitude values, and the average amplitude values, were assessed by separate Antpos (6) by Laterality (3) by Item (7) ANOVAs with Span-group as a between subjects factor. There were no significant effects involving Span-group in any window, for either latency or amplitude, and all effect sizes involving Span-group were small (less than 0.1). Average amplitude also did not differ significantly between the groups in any window, and again, all effect sizes involving Span-group were less than 0.1. There was no evidence of the strong Span-group by Item interactions seen in the first experiment. This is consistent with the visually apparent similarity in stepwise amplitude reduction for low- and high-span groups evident in Figure 5.1.1, which is in contrast to the obvious single step reduction for low-spans seen in Experiment 1.

Given that each group had only four subjects it is possible that there was insufficient power to detect any differences or interactions between the groups, however the very small Span-group effect sizes weaken this argument. If differences between the groups are not apparent statistically, it may be that similarities between them are informative. The group’s waveforms show a pattern of similarities, notably a very strong effect for P3b amplitude reduction for both span groups; for low span, F(6,18) = 7.8, p = .001, η_p^2 = .722, with a significant decrease between two and three items of 1.96 μv (p = .049), and a decrease from three to four items of 4.3 μv that approached significance (p = .084); for high span F(6,18) = 15.99, p = .001, η_p^2 = .842, with a decrease for two to three items of 2.5 μv that approached significance (p = .12), and a significant decrease from three to four items of 3.7 μv (p = .01). These *significant* differences are in line with the group’s subitizing span values, of three and four items respectively, but the very strong effect sizes suggest that the additional amplitude decreases between three and four items for the low span group, and two and three items for the high span group be interpreted as meaningful. This is supported by very strong linear contrasts across item levels one to four for both groups; for low span, F(1,3) = 6.7, p = .08, η_p^2 = .692, and for high span, F(1,3) = 20.5, p = .02, η_p^2 = .872.

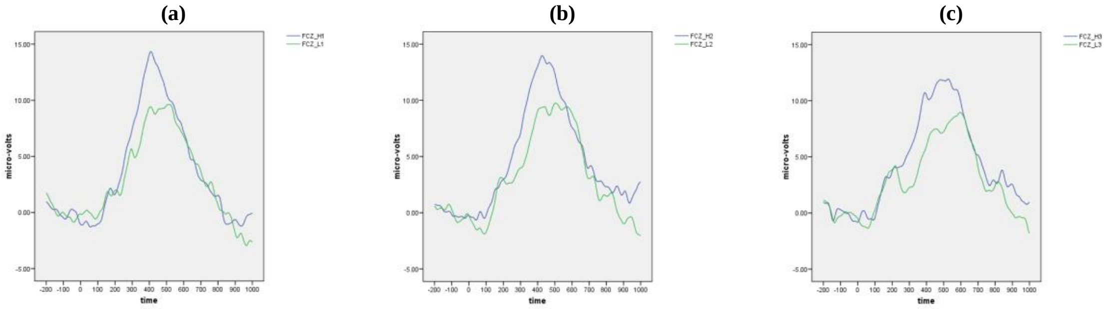
There were also similarities between the two groups in P3b latency behaviour, which, as in the first experiment, tended to increase linearly from one to four items for both groups, as shown by strong linear trends. While these trends did not reach significance at the .05 level (low-span: F(1,3) = 3.8, p = .147, η_p^2 = .558; high-span: F(1,3) = 6.2, p = .08, η_p^2 = .675), the strong effect sizes from relatively few participants suggest the steady increases across the four items might reflect meaningful increases in context updating time. Critically, there is the suggestion that latency might increase from three to four items for both groups, as was found in Experiment 1, and so this data supports the conclusion suggested there, that there is no support for a tight relationship between duration of the LPC and subitizing span.

Another *apparent* feature of the waveforms (because there is no statistical indication of between-group differences in behaviour of the P2-N2) is the clear development of a (small) P2-N2 complex occurring at the final item level of each group's RT derived subitizing span—that is, at three items for the low-span group, and four items for the high-span group. This pattern would be consistent with the observation made in section 4.3.3.4 regarding the item level at which the P2-N2 complex appeared as being an indicator of chunking capacity. However, without any apparent change in morphology being captured with statistical tools it is difficult to make any strong inferences regarding the appearance of these patterns in the waveforms. What we do know is that this pattern in the waveforms beginning around 200 ms post-stimulus is (visually) consistent with the ERP profile that has been associated with time consuming processing in WM occurring during subitizing..

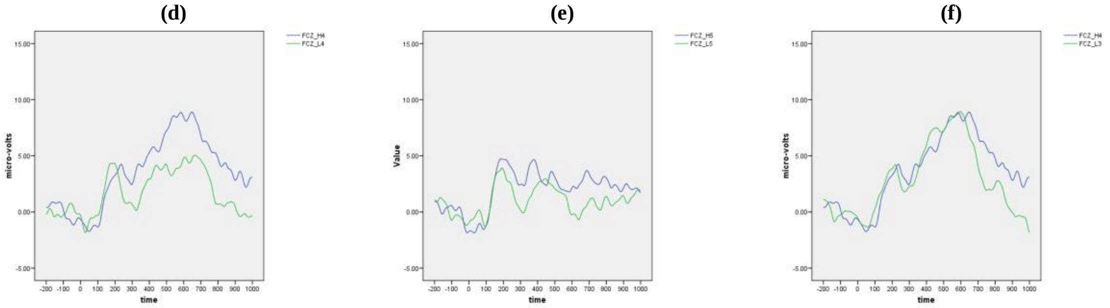
In their guidelines for publishing ERP research, Picton et al. (2000, p. 139) suggest that because of "… the ambiguities inherent in current methods for ERP quantification, the nature of an experimental effect can often be understood most effectively by visual inspection of the appropriate waveforms." While there were no significant between-groups effects in the current data, the typicality of the within-group waveforms (even with so few subjects) in terms of possible indicators identified so far in this thesis of load related WM processing warrants further examination. The following sections examine the waveforms in more detail, making a series of observations that are consistent with negative going activity in the 200-400 ms window reflecting serial WM activity occurring within the subitizing span, and thus with the serial counting explanation of subitizing.

The key point to emerge from what follows is consistent with the conclusions drawn from the span-group analyses in Experiment 1. That is, it may be that we are not looking at an ‘all or nothing’ transition from one process—subitizing—to another—counting—that manifests as a totally *discrete* transition in the waveforms one item later for the high-span group, but rather that for both span groups the developing morphology of the waveforms over the first four item levels reflects a load related processing continuum (see Balakrishnan & Ashby, 1991 for a discussion regarding the continuum view). This is indicated broadly by the development of an amplitude decrease from ~200 ms post-stimulus, which in the average waves for both span groups clearly begins at three items, and continues to decrease through to five items. Note that in terms of RT at three items both groups are still subitizing. For the low-span group this is the final level before their span is exceeded. However, at three items the high-span group are still *two items* within their subitizing span, and yet still show this initial sign of load related WM activity.

Figure 5.12 overlays the waves for the high- and low-span groups for each item level. Visually, the amplitude reduction for the low-span group is accompanied by a distinct P2-N2 deflection at three items. This P2-N2 also is evident for the high-span group, but not until one item later, at four items. This difference was not revealed by the ANOVA. It is clear in panel (c) that at three items amplitude for both groups has decreased, and that the P2-N2 is clearly evident for the low-span group, but not for the high-spans. Of particular interest here is that the low-span wave begins to diverge from the high-span wave at around 200 ms, a feature that can also be observed in the waves for two items, shown in panel (b). Panel (d) shows that at four items there is further amplitude reduction for both groups, more notable in the low-spans, and that the P2-N2 deflection has become obvious for the high-span group. Again, the waves diverge at around 200 ms, with the low-spans’ wave going more negative. At the step to five items, in panel (e), there is relative to four items a decrease in amplitude that is substantial for the high-spans, yet only minor for the low-spans. It is obvious that the amplitude reduction which began at three items for the high-spans has continued at the step to four items, and that it continues markedly again at the step to five. The reduction between four and five items for the low-spans appears much less, suggesting that the majority of the change across item levels has already occurred for the low-span group. Finally, panel (f) is of critical interest. Here the waveforms for the final level of each group’s subitizing span—three items for the low-spans, four items for the high-spans—are overlaid. Firstly, it is obvious that the waveforms both show the small but distinct P2-N2 complex which was absent at the previous level for each group. Emergence of the distinct P2-N2 complex at the last level of the subitizing span is consistent with the findings in the analyses for the entire samples from this experiment, and from Experiment 1 regarding proposed ERP indicators of subitizing capacity. Secondly, it is clear that the waveforms are all but identical—there is no divergence beginning around 200 ms as there is when waveforms of the same item number are compared in the previous panels—an observation borne out by an almost perfect correlation over the first 650 ms, $r(323) = .965$, $p < .001$ (correlations at the other fronto-central sites are of the same order).



1 Item **2 Items** **3 Items**



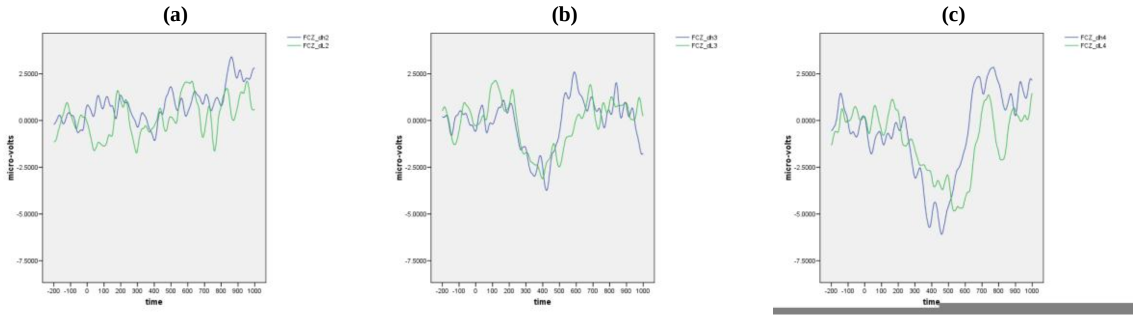
4 Items **5 Items** **Low-3 items / High-4 items**

Figure 5.12. Average waves at FCz for item levels 1 to 5 (panels a, b, c, d, and e, respectively) for low- versus high-span groups (green and blue traces respectively). Panel (f) overlays the waveforms for the final level of the subitizing span for each group; four items for the high- span group, and three items for the low-span group.

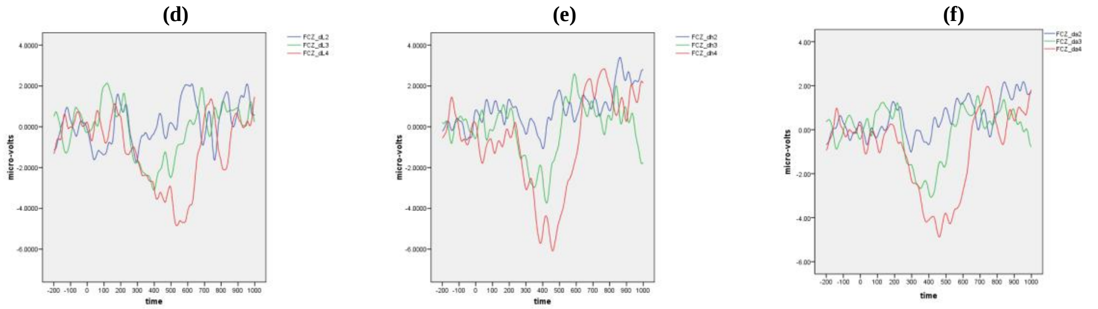
The speculation that emerges from these observations is that of a process that is *gradually* being loaded as additional units of information demand processing — perhaps beginning at two items for the low-spans, and three items for the high-spans—and which may extend over one extra unit of information for the high-span group. This supposition is consistent with that made from the span-group analyses from Experiment 1, although the pattern of data supporting the supposition does not reflect such a clear cut interaction between item number and span-group. Rather, in this sample both span-groups show a similarly gradated amplitude reduction over a couple of item levels, with that for the high-spans perhaps starting one item later. This could imply that the high-span group are better able than the low-span group to absorb the extra processing demanded by the addition of a third, and then a fourth item, within a single processing episode. In terms of the arguments put forward so far in this thesis relating chunking capacity to available activation in WM, it is plausible to suggest that the low- and high-span groups might differ in their capacity for WM activation. Because the group's waveforms start to diverge around 200 ms post-stimulus, that is, in the N2 window, and because negative going ERP activity in this window and beyond into the P3b window has been associated with WM activity, it is reasonable to assume that any differences between the group's waves that would be informative about WM activity would be evident in this window.

5.3.2.5.1 Difference Waves

Difference waves can be informative in psychophysiological research and subtraction techniques are widely used to isolate the degree of a similar activity which varies across conditions (Coles & Rugg, 1995; Luck et al., 2009; Näätänen, 1990; Rugg & Coles, 1995). If WM is fundamentally involved in subitizing, and each additional item adds extra load on these WM processes as the serial counting explanation suggests, then difference waves computed *within-subject* between successive item levels and averaged within span-group should reflect differences in WM activity related to any increasing WM load. The within-group difference waves presented in Figure 5.13 have been computed by, for each subject, subtracting the activity for one item from that for two items, two from three, and so on, and then averaging the waves per condition, and then for each span-group.



2 Items - 1 Item **3 Items - 2 Items** **4 Items - 3 Items**



Low-span **High-span** **All Subjects**

Figure 5.13. Within-group difference waves at FCz for each step to four items (panels a to c; low-span = green trace; high-span = blue trace). Panels (d) and (e) plot the steps to 2, 3, and 4 items for low- and high- span groups, respectively. Panel (f) plots the steps to 2, 3, and 4 items for the combined sample.

Examination of the within-group difference waves in panels (a), (b), and (c) in Figure 5.13 shows clearly that between successive item levels the ERP activity of both span-groups differs in the same general way. For two items (obviously the first item level that the subtraction routine can provide a difference component for) there is little difference to speak of for either group. However, for three, and four items both groups generate substantial difference components that start around 200 ms post-stimulus and continue over the next perhaps 300 ms. This time frame encompasses the N2 window and the P3b window right up to the peak occurrence at each of these item levels. This time frame is consistent with the findings from the first experiment, and from the group data in this experiment, which showed no emergence of major indicators of processing load until three items, when differences began to emerge from around 200 ms post-stimulus.

The plots in panels (d) and (e) overlay the waves for two, three, and four items for low- and high-span groups respectively and are of considerable interest. They show a very consistent start point for the major difference component of around 200 ms for both groups at each item level. They also show what appears to be a systematic increase in the size and duration of the component across successive item levels for *both* groups, even though according to RTs the low-span group is not subitizing any more at four items. To assess any between group differences, area under the curve was extracted for these difference waves and the values for two, three, and four items were input to a 6 (Antpos) by 3 (Laterality) by 3 (Item) repeated measures ANOVA with Span-group as a between subjects factor. There were no effects involving Span-group, with the test of Item's main effect having η_p^2 of only .002, and for the Item by Span-group interaction η_p^2 was only .018. However, collapsed across groups (see panel (f) for the waves averaged across all participants) there was an extremely strong effect for Item, $F(2,12) = 9.06$, $p = .004$, $\eta_p^2 = .602$, supported by an even stronger linear contrast, $F(1,6) = 18.61$, p .005, $\eta_p^2 = .756$. Planned comparisons between successive Item levels showed that area under the curve increased significantly between two and three items ($p = .019$), but the increase between three items and four did not reach significance ($p = .098$). However, given the extremely strong effect size of the linear contrast and the relatively small sample size the idea that successive increases in the number of units of information that had to be enumerated generated successively larger difference components can not be dismissed out of hand. What is plain though is that it is at the same stage in the processing stream where high- and low-span groups differ when enumerating the same number of items, a stage beginning around 200 ms post-stimulus when it is known that integrative processing subserving context updating in WM is underway (see section 2.1.5).

The key point to emerge from examination of the span-group data is that in this sample there are no definitive statistically significant differences between the groups on any of the measures that so far have been informative about differences between item levels. It seems unlikely that this is simply a power issue. Between group effect sizes are miniscule, whereas the within subject effect sizes are large. However, in assessing the difference waves further evidence has emerged showing firstly that the processing that differs between item levels two and three, and three and four, occurs from around 200 ms on, well after any preattentive mechanisms have finished operations. Secondly, at this point in the processing stream, when the effect of increased processing load becomes apparent, the primary indicator of load related activity is prolonged, increased negativity in the ERP waveforms that continues through the entire remainder of the processing episode.

5.4 Conclusions

This experiment sought to further address the ERP profile associated with item enumeration that was found in Experiment 1. The general profile of devolution of the unitary P3 into the LPC at a load of three items was replicated in this experiment, indicating that a robust morphology is generated from the response contingent item enumeration task. The P2 component exhibited the same general behaviour as in Experiment 1, becoming fully established at three items and showing no change in characteristics with the addition of any more items, nor did it show any relationship to increases in either RT or duration of context updating as item load increased. At a load of three items emergence of the clearly defined P2 was accompanied by the distinct N2, and an increase in the latency of the subsequent P3b peak. Confirming the arguments put forward in Chapter 2 regarding the topographical distribution of the anterior negativities associated with effortful processing in WM, these effects were predominant fronto-centrally.

In the data for the entire sample P3b peak latency did not increase significantly between one item and two, but only at the step from two to three items. Thus, the argument suggesting that the non-linear subitizing slope is determined by gradually increasing load on WM resources during subitizing could not be supported as it was in Experiment 1. However, there was a clear pattern of numeric latency increases from one to four items, described by significant, very strong linear contrasts for each composite except Occipital, maximum at Frontal, and Fronto-central sites. The latency increments of the P3b peak were sensitive to the RT increments from two to three items, and from three to four items. This sensitivity was maximal at fronto-central and central sites. At these same two steps the latency increases of the N2 peak were sensitive to those of the P3b, maximal fronto-centrally, and they were both significant.

Taken together the latency data supported the relationship found in the first experiment between the timing of the activity indexed by the N2 peak and subsequent delays in context updating indexed by the P3b peak. As such, these latency findings illustrate clearly a number of major changes in morphology of the LPC across item levels one to three that are consistent with the onset of time consuming WM operations within the subitizing slope. Therefore they provide evidence against any preattentive mechanism(s) determining the characteristics of subitizing. However, the lack of any significant difference in context updating time between one item and two means that they provide no direct support for the serial counting explanation of the non-linear subitizing slope. Thus, parallel processing mechanisms could be involved in subitizing and the time consuming activity evident within the subitizing slope could be explained by the list search explanation.

Moving beyond the subitizing slope and addressing the relationship between LPC behaviour and subitizing span, the latency data do not provide definitive support for the idea that chunking capacity within the duration of the LPC determines the subitizing limit. Trend analyses suggest that P3b latency might increase sequentially up until four items for those with a span of three items, as well as those with a span of four items. Consistent with the findings from the first experiment, this finding does not support the idea that there is a tight relationship between the duration of the LPC, and the RT derived subitizing span. Additionally, the dramatic *decrease* in P3b latency at five items is problematic for the definition of chunking capacity per se. There is a possibility that age of the sample could have influenced this effect though for reasons stated in the earlier discussion this would appear unlikely. Another possibility is that this sample employed different strategies in the counting range then those in the first experiment.

Overall, the latency data provided very useful information for drawing inferences about what types of processing might be occurring during subitizing *within the subitizing slope*. However, in terms of ERP indicators of subitizing capacity P3b latency behaviour was not immediately informative. It seems that duration of an episode of context updating in WM may not be tightly related to differences in span (at least between spans of 3 and 4), but that differences in how much work (processing) can be done in the same duration might be a better indicator of 'chunking capacity', and this activity probably is reflected by amplitude changes (reductions) from around 200 ms on.

The amplitude data across the first four items in this experiment was also generally consistent with the findings from Experiment 1, although the movement in the N2 window was not as pronounced. Nevertheless, average amplitude in the N2 window showed a strong trend to reduce at both three and four items, and P3b amplitude reliably reduced at three and at four items. Looking just within the subitizing slope this pattern across the first three items is consistent with the interpretation that the primary indicator of the onset of processing load is the amplitude decrease beginning around 200 ms post-stimulus with the N2 and extending through the P3b window. This supposition was strongly supported by the span-groups' difference wave analysis, which showed evidence at three items of the appearance of a negative going difference component beginning around 200 ms post-stimulus that extended through both the N2 and P3b windows.

Moving beyond the subitizing slope the significant amplitude decrease across the P3b window at four items for the high, but not for the low-span group might have been seen to indicate a discrete relationship between subitizing capacity and this amplitude behaviour. However, polynomial contrasts suggest that amplitude reduces steadily to four items for both high- and low-span subjects, a pattern which might suggest that the processing response to increasing item load is distributed across a number of item levels. Importantly, this pattern of change is not reflected by any interaction between span size and item level as it was in the first experiment where the distributed response was only evident in the high-span group. It was suggested that this interaction might have reflected a greater capacity for the high-span group to absorb any processing requirements demanded by a fourth unit of information. However, in the current data this capacity is evident for both span-groups and so provides no clue as to how to differentiate people with different subitizing spans on the basis of ERP amplitude changes.

However, the findings from this experiment do need to be interpreted with some caveats in mind. Attrition because of artifact reduced the overall N substantially, and this reduction in power was even more noted in the span-group analyses where each group only had four participants. Inferences made on statistical outcomes from such small sample sizes leave open the clear possibility of false positives and so the conclusions suggested in the previous paragraphs should be treated with some caution. Alternatively though, it could be argued that these results are strong *given* the low subject numbers—and the robust morphology and strong, relatively consistent effect sizes might support this position. Nevertheless, the veracity of these interpretations will be illustrated by any continued robustness in the experiments to come.

Taking these caveats into account, there are key findings to emerge from these first experiments examining the ERP profile associated with subitizing during response contingent item enumeration. Firstly, the latency and amplitude data are equivocal regarding ERP indicators of subitizing span. There does not appear to be a definitive relationship between duration of the LPC and span. The relationship between *amount* of processing able to be accomplished in the duration of the LPC and span, proposed from the Item by Span-group interaction for both N2 and P3b amplitude in the first experiment, is also discounted by the current results. However, while the ERPs have not provided concise information regarding transition from subitizing to counting, at three items when all subjects were still subitizing there were reliable, major changes in morphology of LPC which appear informative about possible changes in processing behaviour as a function of processing (item) load. The most robust of these changes have been increases in P3b latency and amplitude reductions in the N2 and P3b windows at the step to three items. On the basis of these findings, the experiments which follow will each examine questions addressing WM involvement in subitizing by focusing predominantly on P3b latency changes and the amplitude behaviour of the LPC in these windows across the subitizing slope and specifically at the step to three items.

So far in this thesis this related increase in context updating time and onset of pronounced negative going activity has been interpreted as reflecting effortful, time consuming processing in WM, on the basis of the timing relationships between the N2 and the P3b, and the body of evidence linking amplitude reductions in these windows with increased processing load in WM. These interpretations have been based on theoretical and logical considerations, with the primary supposition (articulated in the serial counting explanation) being that subitizing recruits quantification operations that demand WM resources. However, two other possible determinants exist. Firstly, it is possible that pattern matching and associative memory operations underlie subitizing, as suggested by Mandler and Shebo (1982), in which case it could be the operation of these mechanisms which are determining the changes in the LPC (see section 1.2.7). Secondly, as noted in section 2.1.5, in a visual search task increasing perceptual load can lead to decreases in P3b amplitude (e.g., Lorist, Snel, Kok, & Mulder, 1996). Thus, it is possible that the LPC morphology identified so far may be a response simply to increased item load, independent of any task related processing requirements. Experiment 3 assesses these alternative explanations to ascertain whether this negative going activity in the N2 and P3b windows reflects WM activity involved with quantification operations or is determined by configural operations, or simply by an increase in item load.

CHAPTER 6

Experiment 3: Configuration and Item Load

6.1 Rationale, Aims, and Hypotheses

The robust finding from the first two experiments has been that at a load of just three items, still within the subitizing span of all participants, there is clear evidence of emergence of negative going activity in the LPC from approximately 200 ms post-stimulus. This activity has a predominantly fronto-central distribution, and encompasses the initial deflection to the N2 peak as well as ongoing sustained negativity through to the reduced amplitude P3b peak. Additionally, emergence of this negativity has been accompanied by an increase in P3b peak latency, relative to two items, which was also preceded by an associated increase in latency of the distinct N2 peak. On the basis of these timing relationships and the task demands proposed by the serial counting explanation, it has so far been suggested that the clear onset of this negative going activity with the addition of a third unit of information might reflect the operation of *time consuming quantification operations in WM* (that are not obviously recruited when just two units of information must be enumerated). This suggestion has been supported by the argument put forward in section 2.1.5.2 relating this general pattern of fronto-central amplitude reduction to effortful, time consuming processing in WM. However, the first two experiments did not include any empirical control for another possible process which may underlie subitizing—that is, configural processing—and thus for the possibility the negative going activity in question might reflect processes occurring during subitizing that reflect configural processing rather than quantification activity in WM.

A configurational explanation of subitizing was addressed in section 1.2.7, where it was noted that some (e.g., Mandler & Shebo, 1982; Logan & Zbrodoff, 2003) believe that subitizing reflects recognition of canonical patterns—two dots are always joined by a line; three are always configured as some sort of triangle (or a line); and four items are always configured as a quadrilateral (or a line). By Mandler and Shebo's explanation the array elements are bound into a configural representation and the number word is accessed through learned association with the properties of the configuration, and no quantification operations are required during subitizing. Logan and Zbrodoff would suggest that the operations of a pattern matching mechanism underpin subitizing, with the number word also accessed through associative processes. If pattern matching and association are the processes that subserve subitizing, then the negative going activity seen at three items in this program so far might in fact be reflecting configural binding operations and associative processes and not serial quantification operations in WM, as the serial counting explanation suggests and has been argued for so far in this thesis. Note that configural binding operations and associative processes may also occur in WM, but the flat RT function over the subitizing range that is associated with enumerating canonical patterns (see section 1.2.7) suggests that as item level increases no additional, time consuming demands are made on WM resources.

There is also another possible determinant of this negative activity. Increasing perceptual load has been shown to reduce amplitude over the P3b window (Lorist, Snel, Kok, & Mulder, 1996). In the item enumeration task increasing the number of items to be enumerated also increases the perceptual load of a trial, independent of any additional processing demands that the requirement to engage in enumeration might introduce.

Therefore, the primary aim of this experiment is to assess whether processing arrays of three items in terms of their configuration also generates the predominantly anterior negativity in the N2 and P3b windows that has so far been generated during item enumeration. A second aim is to see whether any such negativity is also accompanied by prolonged context updating, reflected by P3b peak latency, which would indicate additional time consuming operations and thus serial processing.

The data reported here addresses these issues by comparing the ERP waveforms from a pattern recognition task to those from the enumeration task. The pattern recognition task required subjects to identify swiftly, and accurately, the *name* of a dot pattern (‘dot’; ‘triangle’; ‘square’), rather than the *number* of dots in an array, as in the item enumeration task. This task can be seen as isomorphic to the enumeration task for item levels one, three, and four. In each case, the same number of units of information need to be bound into a representation that then needs to be bound to the appropriate name word (see section 1.2.3). The key property on which the paradigms might differ is the requirement to engage in enumeration activity, or not, prior to accessing the name word.

Given the proposed isomorphism of the two tasks, and the key property upon which they might differ, two initial hypotheses can be generated to test whether the negative going activity of interest might reflect simple item load, configural operations, or quantification operations which might subserve subitizing. Firstly (*Hypothesis 1*), if simple item load, or configural operations, which might include pattern matching, are involved in generating the anterior negative activity at three items, then the waveforms for triangle and for three items should both show this activity, and it should not differ between these conditions. Secondly (*Hypothesis 2*), given the proposed difference in binding activities between the two paradigms, if it is the case that quantification activity (as required by the enumeration task), but not pattern recognition activity (as required by the pattern recognition task), generates this negative activity, then the waveforms for triangle, and for three items should differ. Specifically, the waveform for three items would be expected to exhibit the anterior negative deflection beginning around 200 ms post stimulus, while the waveform for the triangle condition would not. Additionally, any such deflection should be accompanied by a subsequent reduction in P3b amplitude.

A third hypothesis (*Hypothesis 3*) can be tested to assess whether any evident negative going activity might be reflecting *serial* processes in WM, as speculated so far. If this is so then, consistent with the findings from Experiments 1 and 2, we should see appearance of any notable negative deflection in the N2 window at three items being accompanied by a within condition increase in P3b peak latency relative to the preceding item level. Finally (*Hypothesis 4*), if the processing subserving subitizing is more resource demanding than that required for pattern recognition then we might expect context updating to take longer for subitizing than for pattern association, which would be reflected by delayed P3b latency.

While the focus of these four predictions is on what ERP behaviour emerges at a load of three items, the paradigm included item levels one, three, and four. Data from all three levels is included in the analyses, and discussed where appropriate, as this approach provides a picture of the morphology of the effects of Type of task as a function of item load.

6.2 Method

6.2.1 Participants

Eleven adults (5 M, mean age = 31 +/- 11.4, range = 18-46) from Armidale participated. All had normal or corrected to normal vision, and reported being right handed. None had participated in the previous experiments, and all were naïve to the item enumeration task.

6.2.2 Tasks

Pattern recognition task. All timing, display, and response parameters of the pattern recognition task were identical to the item enumeration task used in Experiment 2. However, instead of randomly configured dot arrays the stimuli were three canonical patterns of dots: 1) a single Dot centred in the 5 * 5 grid; 2) three dots, arranged as the points of a centred equilateral Triangle; and 3) four dots, arranged as the corner points of a centred Square. Required response was identical to that for the item enumeration task—that is, participants were asked to respond with a button press (the numberpad '+' key) when they were sure of which configuration was displayed—that is, as soon as they had accessed the correct name word. As in the item enumeration task, each trial was cued by the participant when they were ready. In order to ensure automatic access to the name words associated with each configuration, prior to beginning the task proper there was a brief association regimen where initially the Dot was presented with the word "DOT"; Triangle with "TRIANGLE"; and Square with "SQUARE" Following this, three blocks of thirty experimental trials were presented. Each block contained 10 trials of each of the patterns, presented in random sequence.

Item enumeration task. In Experiment 4 (reported in the following Chapter) an item enumeration task was used that contrasted performance under two memory load conditions—having to maintain a verbal memory load, or not. The actual item enumeration task was identical to that used in Experiment 2. However, each trial began with either eight asterisks (in the 'no-load' condition) or an eight character CVCVCVCV phonetically legitimate nonsense word (in the 'load' condition) displayed for 1500 ms, after which a fixation cross appeared signaling the start of the enumeration trial as in Experiment 2. After the subject responded to the numerosity prompt which signaled the end of the enumeration trial, another prompt appeared on the screen and the subject had to articulate the nonsense word, in the load condition, or do nothing in the no-load condition. The data examined here for the Enumeration condition was extracted from item levels one, three, and four from the no-load condition.

6.2.3 Procedure

Upon arriving at the lab participants were initially given basic information on the experimental procedure, and informed consent was gained. They were then seated comfortably in the testing room approximately 1.2m from the computer monitor and informed that they would be monitored by closed circuit TV throughout the procedure. Bi-directional voice communication was then explained, followed by a briefing on the importance of minimising ocular and movement artifacts during an EEG recording, and an explanation of a trial sequence for the pattern recognition task. The self-paced nature of the paradigm was explained, with instructions to delay cueing each trial until subjects were sure their eyes were relaxed and they were settled and comfortable in the chair, and to avoid blinking and gross movements during the *critical period* – from onset of the fixation cross through to the numerosity prompt. The pattern recognition task was presented first, followed by the item enumeration task. For the pattern recognition task subjects were instructed to accurately identify the correct pattern name and respond with a button press as quickly as possible. For the item enumeration task subjects were instructed to accurately enumerate the arrays and respond as fast as possible. Thus, for both tasks both speed and accuracy were stressed. The task order was chosen so that subjects would as far as possible engage in pattern matching while seeing the dot arrays during the pattern matching task, and not engage in automaticised enumeration of the arrays, which would be a likely response if exposed to the pattern matching task after just enumerating over 400 dot arrays.

6.2.4 EEG Recording and Data Reduction

EEG recording was identical to Experiment 2, as was data reduction, with the exception that current epochs are of 750 ms post stimulus, with a 200 ms prestimulus interval. All of the components of interest manifest well within this duration.

6.3 Results

6.3.1 Behavioural Data

There were no errors at any of the three item levels for either condition. RT data was screened for each subject as in Experiments 1 and 2, leaving a total of 1015 trials, an average of ~30 trials per subject per item level (~330 trials per item level), for the enumeration condition, and 865 trials, an average of ~27 trials per subject per item level (~297 trials per item level) for the pattern condition. Mean RTs for each item level per condition are presented in Table 6.01.

Table 6.01. Mean RTs (SD) for each condition at each item level.

Items	RT(SD)	
	Pattern	Enum
1	535(103)	580(111)
3	534(106)	595(119)
4	537(130)	723(203)

An Item (3) by Type (2) repeated measures ANOVA revealed a strong Type by Item interaction, $F(2,20) = 9.2$, $p = .001$, $\eta_p^2 = .48$. The interaction is plotted in Figure 6.01. RTs for the Pattern condition did not increase at all between any of the conditions (i.e., between one and four items), consistent with previous findings that canonical configurations eliminated any subitizing slope (e.g., Mandler & Shebo, 1982; Wender & Rothkegel, 2000; Wolters et al., 1987). Interestingly, for the Enumeration condition there was also no significant increase in RT across the subitizing slope (i.e., between one and three items), an increase which might be expected given the findings in the first two experiments. However, there was a substantial significant increase of 128 ms between three and four items ($p = .002$) ,a finding consistent with a transition from subitizing to counting occurring within the group at the step from three to four items. This was confirmed by difference contrasts, reported below. The main effect for Type failed to reach significance, with the numeric increases for the Enumeration versus the Pattern condition of 45 ms and 61 ms at one and three items respectively being unreliable. However enumerating four items took 186 ms longer than identifying four items as a square ($p = .033$). Taking longer to respond when required to enumerate rather than identify canonical patterns is consistent with other findings (e.g., Mandler & Shebo), and with the idea that enumeration recruits serial WM resources beyond those required for pattern identification.

Finally, to confirm that in the enumeration task all subjects were still subitizing at a level of three items—and thus that any negative going activity post 200 ms in the enumeration waveforms would reflect its occurrence *while subitizing was still occurring*, as in the first two experiments—the subitizing span for each subject was calculated by computing difference contrasts across item levels. Subitizing span was taken as the item level prior to that at which the contrast was significant. All subjects showed evidence of still subitizing at a load of three items. Three subjects had a span of four items, and eight had a span of three items.

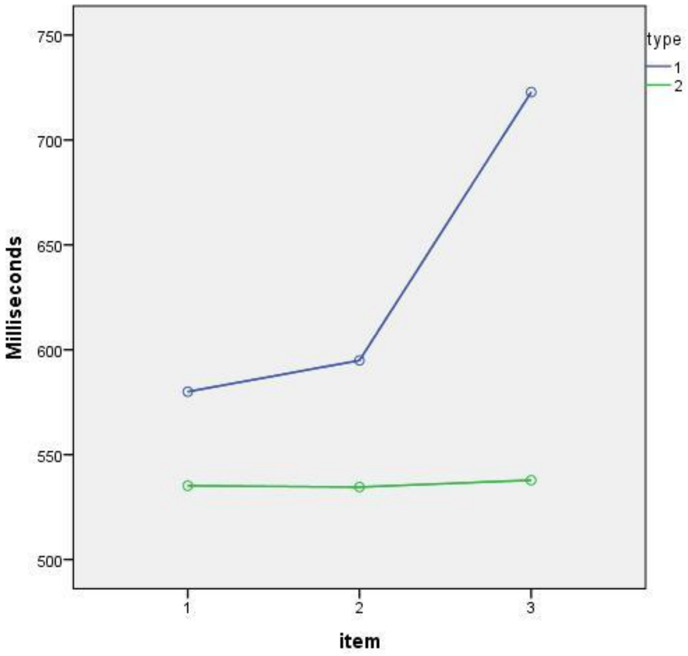


Figure 6.01. Type by Item interaction for RT. Blue trace = Enumeration condition; Green trace = Pattern condition.

In summary, the RT showed that there was no increase in time taken to identify a dot (1 item), a triangle (3 items), or a square (4 items) when the arrays needed to be processed in terms of pattern matching. When the arrays needed to be enumerated there was no evidence of any time cost between one and three items, that is, across the subitizing slope, as might be expected from the consistent appearance of such a slope when enumeration was occurring in the first two experiments. Thus, the functions across loads of one to three items were flat for both conditions. Possible reasons for this will be addressed in the discussion. However, while for the Pattern condition the function remained flat at the step to four items, when enumeration had to occur there was a substantial increase in RT at this step, signifying a transition from subitizing to counting, and indicating that time consuming processing was occurring that wasn’t required to process the items simply in terms of their configuration.

6.3.2 EEG Data

The grand-average waveforms at midline sites Fz, Cz, and Pz for the Pattern and Enumeration conditions, each by Item Load, are presented in Figure 6.02. Examination of the Enumeration waveforms in Figure 6.02 Panel (b) shows the clear emergence of negative going activity from ~200 ms post-stimulus at three items, which is also evident at four items. This pattern is most notable at Cz, and reflects an apparent sequential amplitude *decrease* from one, to three, to four items. Overall, this pattern, notably the pertinent change in morphology at three items, is consistent with the findings from the first two experiments. However, looking at Panel (a), the Pattern waveforms, it is obvious that at a load of three, or four items, the negative going activity which becomes evident is much less notable than for the Enumeration condition. In addition, at the three sites illustrated there is a suggestion of *increasing amplitude* across the N2 window from one, to three and four items, most notable at central and parietal sites, in contrast to the decreasing amplitude for Enumeration.

This observation of differential behaviour between the two conditions as item load increased is illustrated clearly by the pattern of apparent differences between the Enumeration and Pattern waveforms, notably at three, and at four items, illustrated in Figure 6.03, which displays the waveforms for Item Load by Type (1 v. Dot; 3 v. Triangle; 4 v. Square). The waveforms in panels (b) and (c) show a clear fronto-central negative deflection, beginning ~200 ms post-stimulus for the enumeration waveforms that is far more substantial than that seen in the pattern recognition waveforms. Additionally, there appears to be greater P3b peak latency, notable parietally in these plots, for the enumeration condition.

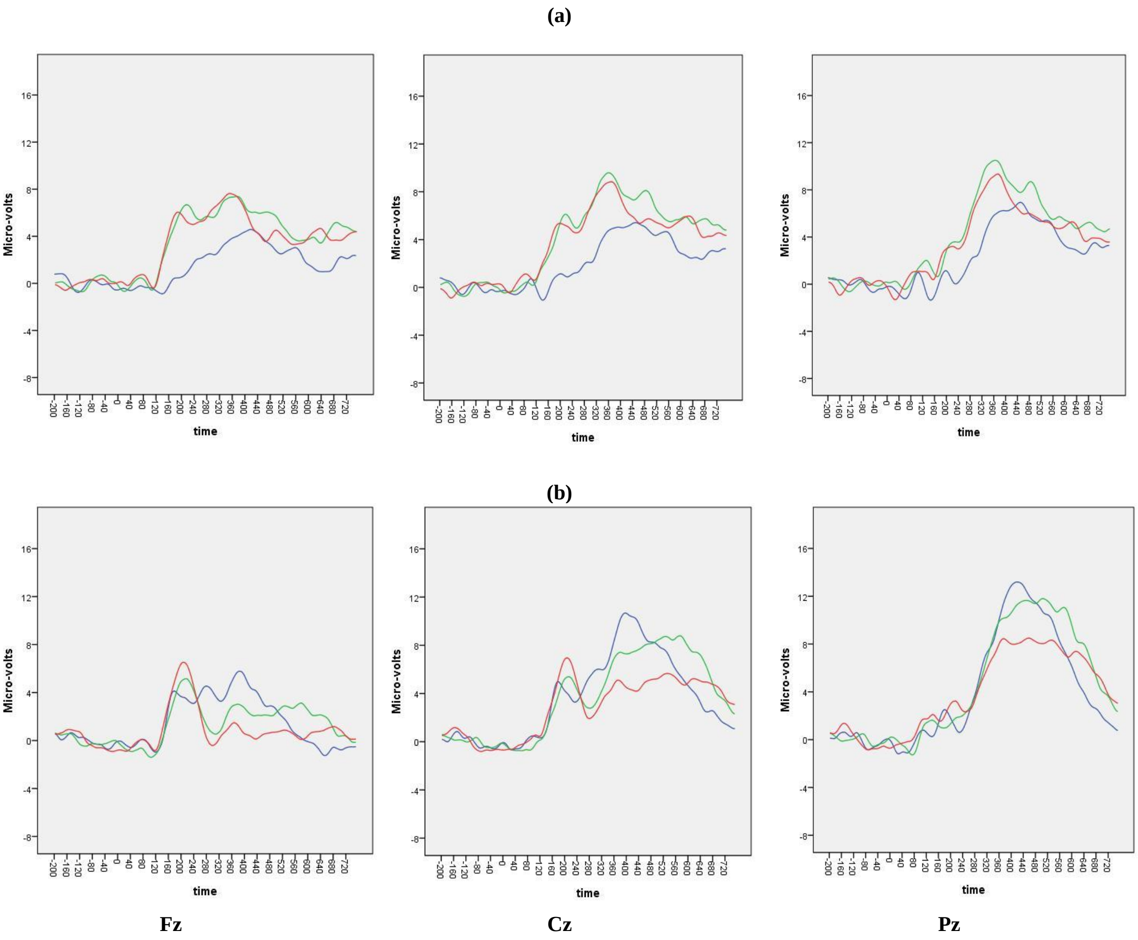


Figure 6.02. Grand average waves at Fz, Cz, and Pz for the Pattern condition by Item Load (panel a), and for the Enumeration Condition by Item Load (panel b). 1-item/Dot = Blue trace, 3-items/Triangle = Green trace, 4-items/Square = Red trace.

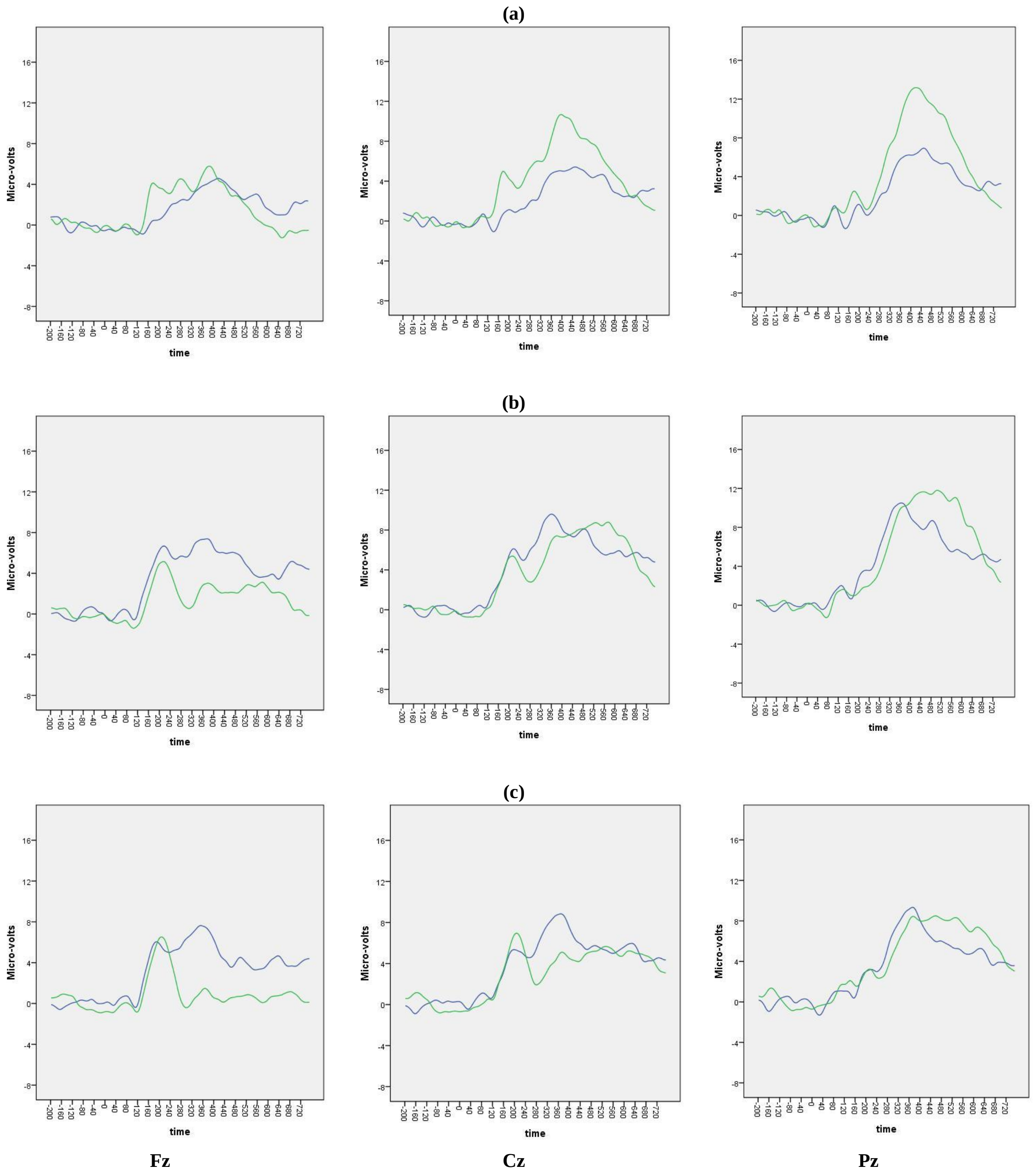


Figure 6.03. Grand average waves at Fz, Cz, and Pz for 1 item versus dot (panel a), 3-items versus 'triangle' (panel b), and 4-items versus 'square' (panel c). Blue trace = Pattern condition, Green trace = Enumeration condition.

To assess the pattern of difference between the Enumeration and Pattern conditions that is apparent in the grand-average waves, a series of 6 (Antpos) * 3

(Laterality) * 3 (Item) * 2 (Type: Enumeration; Pattern) repeated measures ANOVAs were conducted. The Antpos and Laterality composites were identical to those used in Experiment 2. Because the lengthy duration of the negative going activity seen so far from around 200 ms post-stimulus, we have used the average amplitude values across the N2 window (180 ms to 420 ms post-stimulus) rather than peak amplitude values to examine amplitude behaviour within the N2 window. Given the appearance of a clear N2 component, N2 peak latency was also examined. Behaviour of the P3b peak latency and amplitude was examined using as dependent variables the latency and amplitude of the highest positive peak in a window 300 ms to 650 ms post-stimulus. Partial eta-squared (η_p^2) is reported as a measure of effect size. Pairwise comparisons are used to assess differences between successive item levels within condition, or between conditions at the same item level.

6.3.2.1 N2 Amplitude and Latency

N2 Average Amplitude: As might be expected from observation of the grand-average waves, there was a very strong Item by Type interaction, $F(2,20) = 19.4$, $p < .001$, $\eta_p^2 = .66$. Planned comparisons showed that for a load of three items average amplitude across the window was 1.9μv *less* for the Enumeration condition than for the Pattern condition ($p = .005$, $\eta_p^2 = .57$). Additionally, for a load of four items the Enumeration condition was 2.1μv less than for the Pattern condition ($p = .016$, $\eta_p^2 = .453$). The substantially greater relative negativity across this window for the Enumeration condition at a load of three items provides strong support for Hypothesis 2, and thus for the idea that the negative going activity in this window that was observed in the first two experiments *is not* related to configural processing, but reflects other processing that the requirement to enumerate items recruits. This could be the operation of quantification processes that are recruited during item enumeration. These findings do not support Hypothesis 1. This observation is further supported by the significantly greater negativity for the Enumeration condition that is also seen at a load of four items. Mean amplitude values for this interaction are plotted in Figure 6.04.

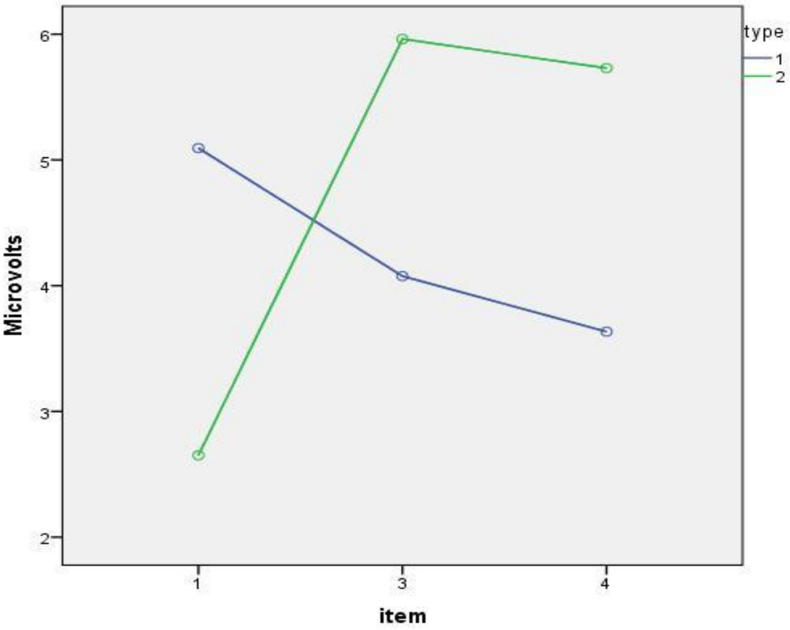


Figure 6.04. Mean amplitudes (µv) across the N2 window for Enumeration (Type 1, blue trace) v. Pattern (Type 2, green trace), by Item Load.

The moderate size Antpos by Type interaction was also significant, F(5,50) = 5.3, p = .001, η_p^2 = .348. The pairwise comparisons for this interaction showed that the Enumeration condition was significantly more negative than the Pattern condition at Frontal sites (1.8µv, p = .023, η_p^2 = .416), and almost so at Fronto-central sites (1µv, p = .063, η_p^2 = .262). These differences are consistent with the N2 topography identified in Experiment 2 and the suggestion made throughout this thesis that the negative going activity of prime interest in terms of association with possible effortful processing during subitizing has a predominant anterior presentation. The interaction is plotted in Figure 6.05. Again, these results provide support for Hypothesis 2, and speak against Hypothesis 1.

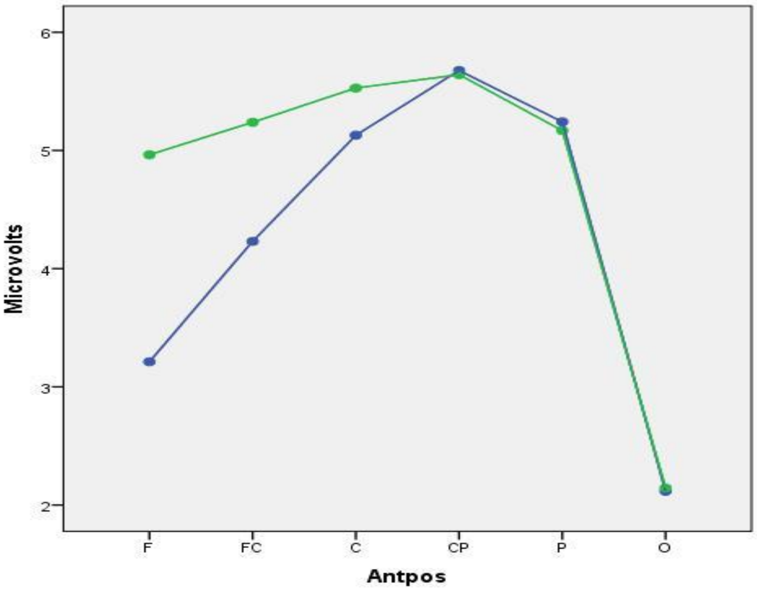


Figure 6.05. Mean amplitudes for the Antpos by Type interaction for amplitude across the N2 window. Type 1 (blue trace) is the Enumeration condition. Type 2 (green trace) is the Pattern condition.

N2 Latency: There was a moderately strong Antpos by Type interaction, F(5,50) = 5.6, $p < .001$, $\eta_p^2 = .36$, which showed that N2 latencies for the Enumeration condition were longer than those for the Pattern condition at Frontal sites (45 ms, $p = .003$, $\eta_p^2 = .615$). Numeric differences approximating 20 ms at Fronto-central ($p = .135$, $\eta_p^2 = .209$), and Central ($p = .119$, $\eta_p^2 = .205$) sites did not reach significance.

Together these data suggest that N2 amplitude tends to be more negative, and N2 peak latency tends to be longer at Frontal sites when three items of information need to be enumerated versus identified as a pattern. These effects are consistent with the anterior effects seen during enumeration in the first two experiments.

6.3.2.2 P3b Latency and Amplitude

P3b Latency: For P3b peak latency there was a strong effect for the Item by Type interaction, F(2,20) = 7.2, p=.004, $\eta_p^2 = .419$. Examination of Item level within Type showed for the Enumeration condition a significant increase from one to three items of 52 ms (p = .012). There was no further increase, with latencies for three and four items being identical. There were no significant differences between the item levels for the Pattern condition. Between conditions, however, collapsed across sites P3b latencies were longer for the Enumeration than the Pattern condition for a load of both three, and four items, by 62 ms (p = .027, $\eta_p^2 = .402$) and 76 ms (p = .037, $\eta_p^2 = .367$) respectively. This interaction is plotted in Figure 6.06. There was also a significant Antpos by Type interaction, F(5,50) = 6.7, p < .001, $\eta_p^2 = .402$, where numerically greater P3b latencies the Enumeration versus Pattern conditions of 34 ms and 46 ms for Frontal and Fronto-central sites of 34 ms (p = .175, $\eta_p^{2=}$ .176) and 46 ms (p = .104, $\eta_p^{2=}$ .243) respectively did not reach significance, but differences at Central and Centro-parietal sites of 71 ms (p = .013, $\eta_p^2 = .477$) and 60 ms (p = .024, $\eta_p^2 = .413$) respectively were significant. This interaction is plotted in Figure 6.07. Thus, there was no support for Hypothesis 3. Greater P3b latencies for the Enumeration condition were evident at the same item level where the N2 initially showed more negative going activity, but not clear at Frontal or Fronto-central sites where there was evidence of greater relative negativity for the Enumeration condition in the N2 window. Hypothesis 4, however, was supported. At a load of both three and four items context updating took longer when the items had to be enumerated than when they had to be processed in terms of configuration,

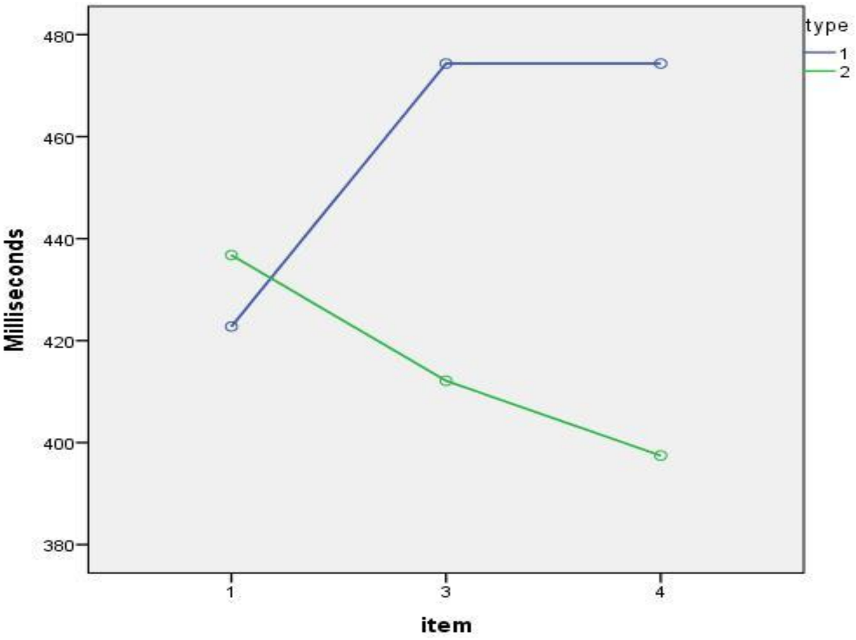


Figure 6.06. Mean latency (ms) for P3b peak for Item by Type interaction (Blue trace = Enumeration; Green trace = Pattern).

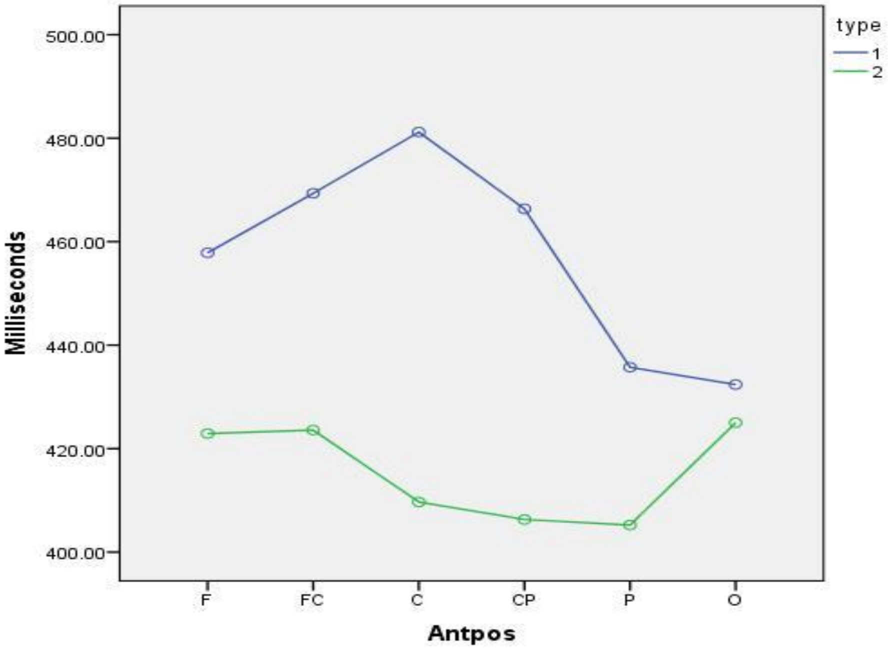


Figure 6.07. Mean latencies for P3b peak for Antpos by Type interaction. (Blue trace = Enumeration; Green trace = Pattern).

P3b Amplitude: There was a strong Item by Type interaction for P3b peak amplitude, F(2,20) = 20.9, $p < .001$, $\eta_p^2 = .677$. For a load of one item amplitude for the Enumeration condition was substantially greater than for the Pattern condition (3.2µv, p = .002, η_p^2 = ..63). Amplitudes were equivalent for three items, while for four items the Enumeration condition was considerably less than the Pattern condition (2.7µv, p = .05, $\eta_p^2 = .31$). This interaction is plotted in Figure 6.08, which illustrates clearly the strong linear contrast which also describes it F(1,10) = 30.1, $p < .001$, η_p^2 = .751.

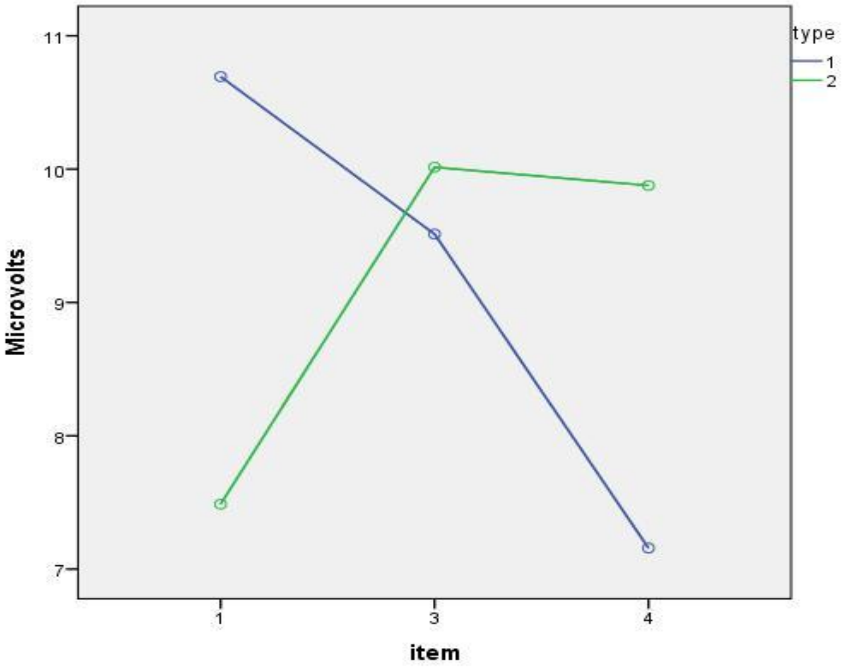


Figure 6.08. P3b amplitudes (µv) for the Item by Type interaction. Blue trace = Enumeration condition, Green trace = Pattern condition.

There was also a strong Antpos by Type interaction, F(5,50) = 9.1, $p < .001$, $\eta_p^2 = .477$, driven by significantly lower amplitude for the Enumeration condition Frontal sites (-2.6μv, p = .048, $\eta_p^2 = .337$), and Fronto-central (-1.8μv, p = .118, $\eta_p^2 = .416$) sites, while tending to be greater than the Pattern condition at Parietal (2.1μv, p = .097, $\eta_p^2 = .454$), and Occipital (2.7μv, p = .054, $\eta_p^2 = .558$) sites. This interaction is plotted in Figure 6.09, which nicely illustrates the very strong linear contrast which also describes it, F(1,10) = 12.8, p = .005, $\eta_p^2 = .562$.

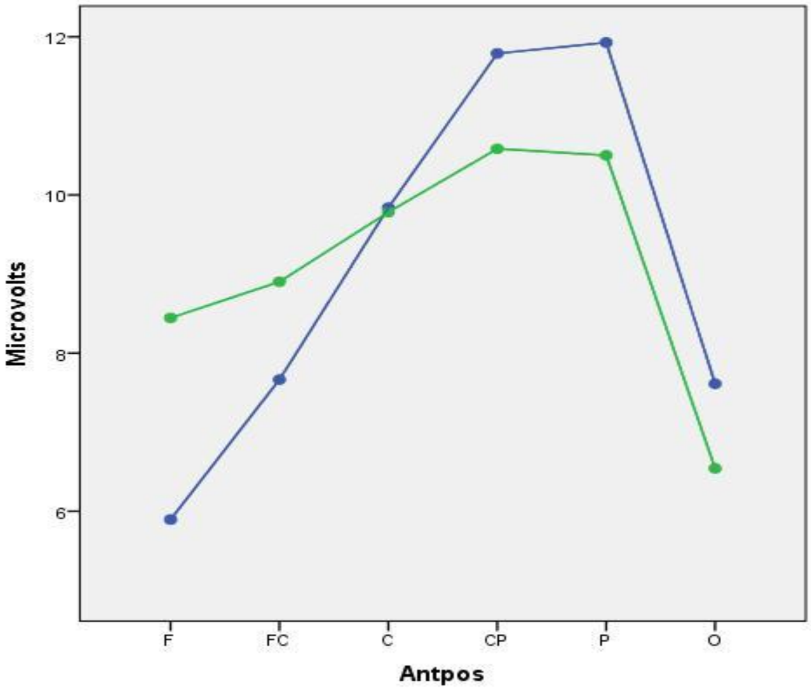


Figure 6.09. P3b amplitudes for the Antpos by Type interaction. (Blue trace = Enumeration condition, Green trace = Pattern condition).

Together, the P3b data provides further support for Hypothesis 2, suggesting that enumeration within the subitizing range recruits more effortful cognitive resources than required for pattern matching operations over the same item load, and that these operations associated with enumeration consume additional time, supporting Hypothesis 3. Additionally, they also strengthen the proposal that these effects are predominantly anterior.

6.4 Discussion

While at a load of three items there was evidence of negative going activity within the N2 window in both Pattern and Enumeration conditions these results show that this activity was far stronger when the items must be enumerated rather than processed only in terms of their configuration, and it was also accompanied by reductions in P3b amplitude that were not evident in the Pattern condition.. The effect of requiring quantification of the array elements was strong. Thus, there was no support for Hypotheses 1, and so for the explanation that the negative going activity in the N2 window might reflect processes related to configural processing. However, the findings provide clear support for Hypothesis 2—and so, for the idea that the negative going activity in the N2 window, and subsequent reductions in P3b amplitude, which manifest in the waveforms at three items reflects the operation of quantification operations that are associated with item enumeration while subitizing, as defined by RT, is still occurring.

There was only limited support for Hypothesis 3—the clear emergence of the N2 deflection at three items in the Enumeration condition was only significant at Frontal sites whereas the significant increase in P3b latency relative to the preceding item level was not clear at anterior sites The clear increase in P3b latency which accompanied emergence of the N2 deflection at three items in Experiments 1 and 2 was used as evidence to support the supposition that the negative going activity in the N2 window reflected time consuming, serial processes in WM. The data from this experiment therefore does not allow such clear inference regarding the functional significance of the N2 deflection, beyond its relationship to the operation of processes related to quantification. This finding is at odds with the robust presentation of increased P3b latency at a load of three items in both Experiment 1, and 2. It is possible that the addition of the extra task elements ‘bookending’ the item enumeration task in the memory load paradigm from which the Enumeration condition data for this experiment was collected could have disrupted clear performance of the task, as was surmised earlier in relation to the lack of an RT subitizing slope as seen in the first two experiments. Although there were no additional *explicit* processing demands in the No-load condition (just as there were none in the enumeration task administered in the previous experiments), there were additional perceptual components in the task structure which may have disrupted the clean flow of the procedure relative to the first experiments. This issue will be pursued in discussing Experiments 4, and 5, to follow.

P3b latency was considerably greater for the Enumeration condition than for the Pattern condition—that is, *between* the conditions. For a single item this was not the case. Only when the N2 deflection became obvious in the Enumeration condition, at three items, did context updating time become longer for the Enumeration condition than for the Pattern condition. The increased P3b peak latency for both three and four items for the Enumeration condition (compared to the Pattern condition) suggests that context updating took longer when the items needed to be enumerated rather than simply processed in terms of their configuration. However, as discussed above, this increase in context updating time, while coincident with the notable appearance of the N2 deflection, did not occur at the sites where the N2 was maximal.

The design of this experiment suggests two possible issues which could confound the interpretations being made. The Pattern matching procedure was always administered first, and the Enumeration condition, which followed, contained vastly more trials and was particularly taxing on participants, who reported considerable fatigue from middle of the procedure onwards (which required managing, as will be discussed in the following experiment). This raises the obvious possibility of an order effect. However, as stated earlier in this Chapter, it was not considered possible to counterbalance the task order because of the automaticity which develops in enumeration over the course of over 400 enumeration trials. All participants in this research program reported this effect. Once you have become used to enumerating the simple dot stimuli it would not be possible to expect that the simple associations required to validly perform the Pattern matching task could replace this learned behaviour. However, the issue of fatigue influencing the differences seen between the conditions cannot be entirely discounted. Future work on this Pattern Matching versus Enumeration paradigm should attempt to match the trial load between conditions.

Keeping these caveats in mind, together, these findings provide some indication that it is not simply an increase in item load per se that causes the amplitude reduction beginning in the N2 window. Only when the requirement to enumerate the items was introduced did the major N2 deflection occur, and the associated decrease in P3b amplitude, and then at the same item load where this has become apparent in the previous two experiments. While the current findings provide some empirical support for the supposition made from the data from the first two experiments, that is, that the N2 deflection reflects processing activity (perhaps in WM) which subserves quantification processes that operate during subitizing, the P3b latency data do not provide clear support for the idea that the N2 deflection reflects serial processing in WM that occurs within the subitizing slope. However, the ERP behaviour in this experiment is consistent with the research detailed in Chapter 2 showing that paradigms which incrementally increase load on WM generate load related changes in the LPC that may include the appearance of predominantly fronto-central negative going activity in a window extending from ~200 ms to 400 ms post-stimulus, as well as increased latency and decreased amplitude of the following P3b peak. The scalp distribution of the often very strong sizes of these effects supports this conclusion. Overall, this experiment has provided evidence against the possibility that one key explanation for subitizing—configural processing—generates this ERP profile. Additionally, it has shown that increasing perceptual load does not by itself generate the profile, and so has strengthened the interpretation of this behaviour of the LPC during item enumeration as reflecting WM activity that subserves quantification operations which occur during subitizing. The two experiments which follow test this supposition directly by introducing a memory load concurrent with the requirement to enumerate.

CHAPTER 7

Working Memory Load During Subitizing

7.1 Rationale, Aims, and Hypotheses

From the first two experiments it was argued that the change in morphology of the LPC evident at three items, while subitizing was still occurring, might reflect serial WM activity associated with quantification operations. Increases in context updating time and the concurrent indication of processing load provided by the robust amplitude reduction in the N2 and P3b windows were taken as being inconsistent with theories of subitizing as a parallel process *not* dependent on WM resources. These characteristics were instead interpreted in support of the serial counting explanation of subitizing. Experiment 3 went on to provide support for the proposed relationship between the negative activity across the N2 and P3b windows and some type of processing operations which are demanded by subitizing, but which are not demanded by configural processing. Given the task differences between the enumeration and configuration conditions it was claimed that this activity likely reflected quantification operations of some sort, a position consistent with the serial counting explanation. However, Experiment 3 did not provide compelling evidence linking this negative going activity evident in the enumeration condition to time consuming activity that delayed context updating in WM. Thus, to this point any interpretation of the ERP profile observed so far as being an indicator of time consuming, effortful processing in WM during subitizing lacks definitive empirical support, and is still logically based on the theoretical considerations introduced in Chapter 2. The current chapter reports two experiments that directly address the question of the functional significance of this ERP profile as an indicator of WM activity.

The logic underlying the approach taken in these experiments is simple, and consistent with the dual-task paradigms used to assess WM structure, and capacity (introduced in sections 1.1.1.1 and 1.1.1.2), and with methodologies that have been used in the few RT based studies that have examined verbal, and spatial WM involvement in subitizing using response contingent displays. In these studies subjects have been required to enumerate arrays of simple items—the primary task—while performing other, secondary tasks that might place a concurrent load on WM resources. In one study (Logie & Baddeley, 1987, Study 1) item enumeration was examined in terms of Baddeley's (1984) multi-component model of WM, focusing on the verbal component, the articulatory loop. It was found that relative to a pure counting condition, articulatory suppression (repeating '*the*' 4 times a second while enumerating item arrays) had no effect on RTs for arrays in the subitizing range, but RTs increased significantly into the counting range. It was suggested that subitizing does not rely on subvocalisation, and so does not demand resources from the limited capacity phonological loop component of WM.

In an effort to clarify and extend the findings of Logie and Baddeley (1987), Trick (2005) examined dual-task interference during a dot enumeration task where performance on a baseline condition could be compared to that under four different types of secondary load; simple or complex articulation, and simple or complex tapping. Concurrent articulation was designed to engage verbal WM resources, as Logie and Baddeley did. It was hoped the tapping task would engage spatial WM resources. While interference measures (secondary task RT – baseline RT) showed some time costs during subitizing, irrespective of the type of secondary task there was still the classic RT discontinuity around four items, signifying that subitizing over the first three items was certainly preserved. Still, across items one to three there were substantial interference effects for complex articulation (repeating an alternating two letter sequence such as FG; ~80 ms), simple tapping (~110 ms), and complex tapping (160 ms). Simple articulation (repeating a single letter) produced no appreciable interference effect.

Together these data have been interpreted as suggesting that subitizing is not affected by compromising either verbal, or spatial WM resources, and therefore does not require WM resources to procede. However, Logie and Baddeley's verbal secondary task (repeating '*the*') may not have been sufficiently taxing to have an effect. Indeed, it was not until well into the counting range at a load of eight items that they found this task affecting RTs. In Trick's work even though the secondary tasks did not eliminate subitizing, with the exception of simple articulation (which was an even easier task than that used by Logie and Baddeley) they did produce interference effects at each level of the subitizing slope. For each secondary task these effects were of the same magnitude at each level, thus having the effect of a linear transform on the RT data. Given that these tasks were designed to utilise some attentional resources, the fact that they produced interference effects at all indicates that some resource sharing may have been occurring between the processes subserving subitizing, and those required to perform the secondary tasks. However, besides these two studies mentioned there have been no studies of which the author is aware that have explicitly examined WM activity during subitizing (using either response contingent or tachistoscopic displays) using dual task methodology, and certainly none using ERPs to gather information about what processing activities might be involved as a processing episode unfolds prior to a response being made.

Each experiment reported in this chapter uses an analogue of the basic dual-task paradigm (as described in sections 1.1.1.1, and 1.1.1.2), where the enumeration task must be performed under conditions where there is, or is not, a concurrent maintenance load on WM. Experiment 4 imposes a verbal information load, while Experiment 5 imposes a spatial information load. This design allows predictions to be tested which address the functional significance of the negative going activity of the anterior N2 and the P3b, and associated P3b and N2 latency increases, as indicators of WM activity, as well as the way in which WM might be involved in subitizing. Because most people are still subitizing when required to enumerate an array of three items, and major changes in morphology have so far emerged at three items, within the subitizing span, it is within this item range that any effects of Load on subitizing and on the components of the LPC indicative of effortful use of WM resources would occur. As such, all analyses in this Chapter will focus on what changes might occur across item levels one to three.

If the behaviour of these amplitude reductions from around 200 ms post-stimulus at a load of three items seen in the previous experiments *does* reflect effortful, serial WM activity, then during subitizing the addition of a concurrent load *sufficient to tax WM resources which are actively involved in subitizing*[1] should generate changes in these components of the LPC. Such changes could become evident in one of two ways. Firstly, they may occur at a load of two, or even one item. In this case we would expect to see at these item loads (*Hypothesis* 1) longer P3b latency with widespread scalp distribution for the Load condition if memory load is utilising resources that would otherwise be involved in achieving context updating in

---

[1] Bearing in mind that the simple articulation tasks used by Logie and Baddeley (1987), and Trick (2005) may not have been sufficiently demanding on WM resources to generate an effect, any changes in processing that might be driven by holding a concurrent maintenance load in WM would be dependent on the secondary task being sufficiently taxing on WM activation resources. For this reason the memory stimuli in the current verbal task consist of four sound segments, two more than Trick's (2005) complex articulation task, which did produce interference during subitizing. The spatial memory stimuli used in Experiment 5 also are complex, consisting of conjunctions of spatial information (i.e., colour, location, and orientation).

terms of the primary task—that is, to enumerate. We might also expect to see (*Hypothesis 2*) increased P3b latency also be accompanied by increased anterior N2 latency. Given the findings so far in this thesis we would also expect these characteristics to be accompanied by (*Hypothesis* 3) generalised reduced P3b amplitude for the memory load compared to the No-load condition, accompanied by (*Hypothesis 4*) reduced anterior N2 relative amplitude. Together this behaviour would indicate that WM resources are integral to subitizing two items, or even a single item, a position that would be consistent with evidence regarding the effect of attentional disruption on subitizing that was presented in section 1.2.4., and with the serial counting explanation of subitizing. Alternatively, the effect of memory load (if there is one) might not emerge until a load of three items, a load where the data have so far suggested that effortful, time consuming WM activity may already be occurring. If a memory load effect first emerges at a load of three items then at this item level (*Hypothesis 5*) we would again expect to see reduced N2 / P3b amplitude for the memory load compared to the No-load condition, along with (*Hypothesis 6*) the possibility of longer P3b latency for the Load condition. Emergence of these indicators at a load of three items but not at two, or one, item would firstly provide some confirmation that the morphology of the waveforms for three items seen in the first two experiments does indicate WM activity related to processing load. Secondly, it would suggest that it is only at an item load of three that WM resources are critical during subitizing, and that WM resources may not be utilised in subitizing one or two items. Finally, of course, carrying this memory load may have no effect on the ERP profile identified so far. That is, (*Hypothesis 7*) the waveforms for the Load and No-load conditions will exhibit the same morphology across the three item levels. It could be that the addition of a memory load recruits other neural circuitry than that which is generating the ERP profile and so might generate new component behaviour, and not influence the ERP components that so far have varied with item load during subitizing.

7.2 Experiment 4: Verbal Load

7.2.1 Method

7.2.1.1 Participants

Eleven adults (3 M, mean age = 28 +/- 8) from Armidale participated. All had normal or corrected to normal vision, and reported being right handed.

7.2.1.2 Stimuli

The verbal memory load condition was operationalised using eight letter consonant(C)-vowel(V)-C-V-C-V-C-V phonetically legitimate nonsense words (e.g., Zopuzeha; Hipabuni; Buwigazu; Rodamiwu). The list was vetted by four independent raters who identified any nonsense words that they believed were phonetically similar to real words and thus may have provided phonetic or semantic associations which could have aided memorisation and subsequent recall. These items were removed from the nonsense word pool, from which 180 exemplars were chosen to be presented during the procedure. The No-load condition was identical in structure to the Load condition but the letter sequences were replaced by eight asterisks. Both the letter sequences and asterisks were presented in uppercase 20 pt Times New Roman font, centred across the (invisible) array grid such that four characters fell either side of the centre of the centre cell. As such, the full sequence protruded either side of the array space and so subtended a little more than one degree of horizontal visual angle. Since no EEG data was of interest during this exposure phase there was no concern for any small saccades.

7.2.1.3 Task

The *base item enumeration task* used for this experiment was identical to that used in Experiments 1, and 2, where a response contingent dot array was preceded by a 1500 ms fixation cross, and followed by a delay of 1000 ms and then a numerosity prompt. In the current experiment this base task was embedded in the memory load paradigm, such that a trial sequence now began with a 1000 ms fixation cross, followed by presentation for 1500 ms of a nonsense word in the Load condition, or an equivalent length sequence of asterisks in the No-load condition. In the Load condition subjects were asked to read the word to themselves and to memorise its sound, and maintain that memory over the remainder of the trial. There was no task requirement in the No-load condition. The base item enumeration task then followed, and once the subject had responded to the numerosity prompt, a memory prompt was presented where in the Load condition they were required to verbalise the nonsense word memory stimulus, and the experimenter recorded its accuracy (see explanation of accuracy criteria, below). The same prompt appeared in the No-load condition however no act was required of the participant and they could immediately progress to the next trial. The participant then pressed the ‘+’ key on the keyboard numberpad to initiate the next trial. There were 11 blocks of 42 stimuli per block, with each block containing one array each of ten, and eight items, and two arrays each of nine, and seven items. These six arrays were presented without memory load and were included to offset item density judgements for the largest arrays presented. The remaining 36 trials comprised six exemplars each of arrays of from one to six items, three of which were presented with concurrent memory load, and three of which were not. Presentation order was randomised.

Accuracy: Only trials where the participant could not provide the first two syllables of the memory stimulus were scored as incorrect, and removed from subsequent analyses. For all other trials the memory stimulus recalled was scored as correct. This criterion was followed because of (a) the primary aim of the task, and (b) the task structure during the retention interval which made precise recall of the memory stimulus difficult due to retroactive interference (for review see Pashler, 1994). Firstly, the aim of the task was for participants to carry an active maintenance load in WM *while enumerating one to three items.* To this end participants were given very clear instructions to try hard to continue remembering the memory stimulus *while they were actively enumerating the array.* Instructions were also given to then try very hard to then keep the memory stimulus active for subsequent recall, of course. Between each block participants were also reminded that the task of paramount interest in each memory load trial was to ensure they maintained the memory stimulus active during enumeration.

Given the primary aim of the task, for practical purposes the first 1000 ms after onset of the item array certainly was the critical time for having concurrent memory load, as enumeration up to an item load of three was always accomplished within this window. However, there was a delay of at least 5000 ms between presentation of the memory stimulus and the prompt to recall it. In this time a great deal goes on. Firstly the ‘word’ is displayed for 1500 ms, then there is the 1500 ms fixation cross preceding onset of the item array. Thus, participants had close to 3000 ms in which to rehearse the sounds of the nonsense word, rehearsal which it is hoped was still proceeding during the subsequent time as the item array was being enumerated. After responding to the array there was a further delay of 1000 ms until the numerosity prompt appeared, when the subject had to verbalise of the array’s numerosity. During this time the number had to be kept active in WM, ready for response. Finally the recall prompt was presented. Because of these delays, and the inclusion of additional interference the chances of correct recall were reduced. As such, given the repeated instructions given to subjects to focus on remembering the verbal stimulus while enumerating, it was considered that partial recall would be a sufficient indicator of compliance with the secondary task.

7.2.1.4 Procedure

Upon arriving at the lab participants were initially given basic information on the experimental procedure, and informed consent was gained. They were then seated comfortably in the testing room approximately 1.2m from the computer monitor and informed that they would be monitored by closed circuit TV throughout the procedure. Bi-directional voice communication was then explained, followed by a briefing on the importance of minimising occular and movement artifacts during an EEG recording. After being presented with the pattern matching paradigm previously explained when reporting Experiment 3, subjects were given an explanation of a trial sequence for both the Load, and No-load conditions for the memory load paradigm. The importance was stressed of *remembering* the load stimulus and responding to the memory prompt, in the load condition, as was the importance of *not needing to remember* any information in the no-load condition. The self-paced nature of the paradigm was explained, with instructions to delay cueing each trial until subjects were sure their eyes were relaxed and they were settled and comfortable in the chair, and to avoid blinking and gross movements during the *critical period* – from onset of the fixation cross following presentation of the memory stimulus (or the asterisks for the no-load condition) through to the numerosity prompt. For the no-load condition subjects were instructed to accurately enumerate the arrays and respond with a button press as fast as possible when they were sure of the numerosity of the array. For the load condition, subjects were instructed to keep the memory stimulus active in memory, to then accurately and swiftly enumerate the following dot array and respond with a button press as fast as possible when they were sure of the numerosity. Then after reporting the numerosity of the array when prompted, subjects were presented with a memory prompt where they were to verbalise the memory stimulus.

7.2.1.5 EEG Recording and Data Reduction

EEG recording and data reduction was identical to Experiments 2 and 3 with the exception, for the same reason as in Experiment 3, that epoch length was 750 ms.

7.2.2 Results and Discussion

7.2.2.1 Behavioural Data

There were only 10 errors in enumerating at any of the three item levels for either condition. RT data was screened for outliers for each subject as in Experiments 1, 2, and 3. Out of 2167 trials this led to 96 (4.4%) being removed from the analyses. Trials that were excluded from ERP analysis because of artifact (108; 4.9%) were also removed, along with the 10 trials where errors were made, leaving a total of 1953 trials, an average of ~30 trials per subject per item level per condition (~326 trials in each condition). Mean RTs for each item level (one to three) per condition are presented in Table 7.01, along with those for four items, to assess whether subitizing did in fact occur in this paradigm. A 4 (Item) * 2 (Load) repeated measures ANOVA showed that RTs do not differ between Load conditions at item levels one to three, but a significant effect for Item, $F(3,30) = 16.1$, $p < .001$, $\eta_p^2 = .617$, was driven by substantial increase at four items of 128 ms, collapsed across conditions.

Table 7.01. Average RTs (ms) and standard deviations for each item level, including four items, by Load. None of the differences between Load conditions were significant.

Items	1	2	3	4
No-load	580 (111)	565 (114)	595 (118)	722 (203)
Load	597 (113)	600 (108)	624 (82)	752 (188)
Difference	17	35	29	30

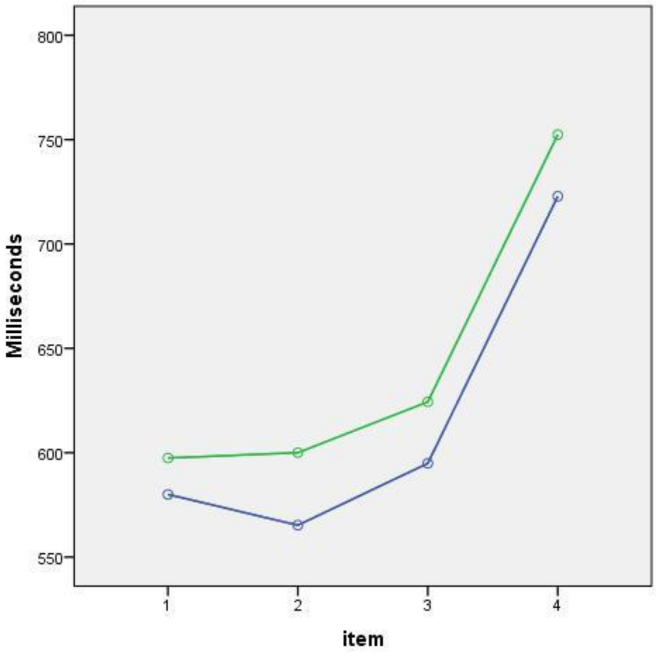


Figure 7.01. RTs for items one to four, by Load. Blue trace = No-load, Green trace = Load.

The RT data, plotted by Load in Figure 7.01, thus indicate that, defined by RT, subjects are still subitizing at an item load of three whether they are holding a concurrent memory load in WM, or not. This was supported by difference contrasts.

This finding is consistent with that of Trick (2005), where subitizing was preserved even when there were secondary task demands on WM capacity. Additionally, the lack of any appreciable interference effect would suggest that the requirements of maintaining the verbal stimulus in memory may be similar to those demanded by Trick’s simple articulation task. However, this supposition is at odds with subject reports regarding the taxing nature of holding the four syllable memory load, which simply in terms of its information load would appear to be more complex than the repetition of a single letter. Finally, it should be noted that the subitizing slope seen clearly in the first two experiments is not evident in this data. During debriefing subjects reported that the task was very demanding, with one aptly describing it as ‘busy’. Their reports were in contrast to those from the first two experiments, where subjects reported going into a ‘zone’ where they had relaxed focus on the trial progression in each block. It is possible that the increased task demands precluded subjects being able to relax into a rhythm and therefore that subjects hurried their button press responses to two, and to three items in order to reach the recall prompt as soon as possible, and so be seen to perform well on the memory task, and that this lack of focus could have carried over to performance in the No-load trials. In any case, whether active (stimulus to remember) or passive (no stimulus to remember), introduction of the secondary task into the enumeration trial structure has obviously disrupted something in the process that was evident for the enumeration paradigm alone.

7.2.2.2 EEG Data

Examination of the grand-average waves in Figure 7.02 shows clearly that the basic morphology of the waveforms for the Load condition is similar to that seen so far in this thesis for the No-load condition. Thus, carrying a memory load while subitizing does not appear to cause radical changes (that are visually evident) in the ERP profile associated with enumeration. There is a suggestion that P3b peak latency may be later for three items than for two, though this is not as clearly evident as in the first two experiments. Both conditions show a P3b amplitude reduction between two and three items, but this is not as marked as in the first two experiments, and is also barely evident at posterior sites.

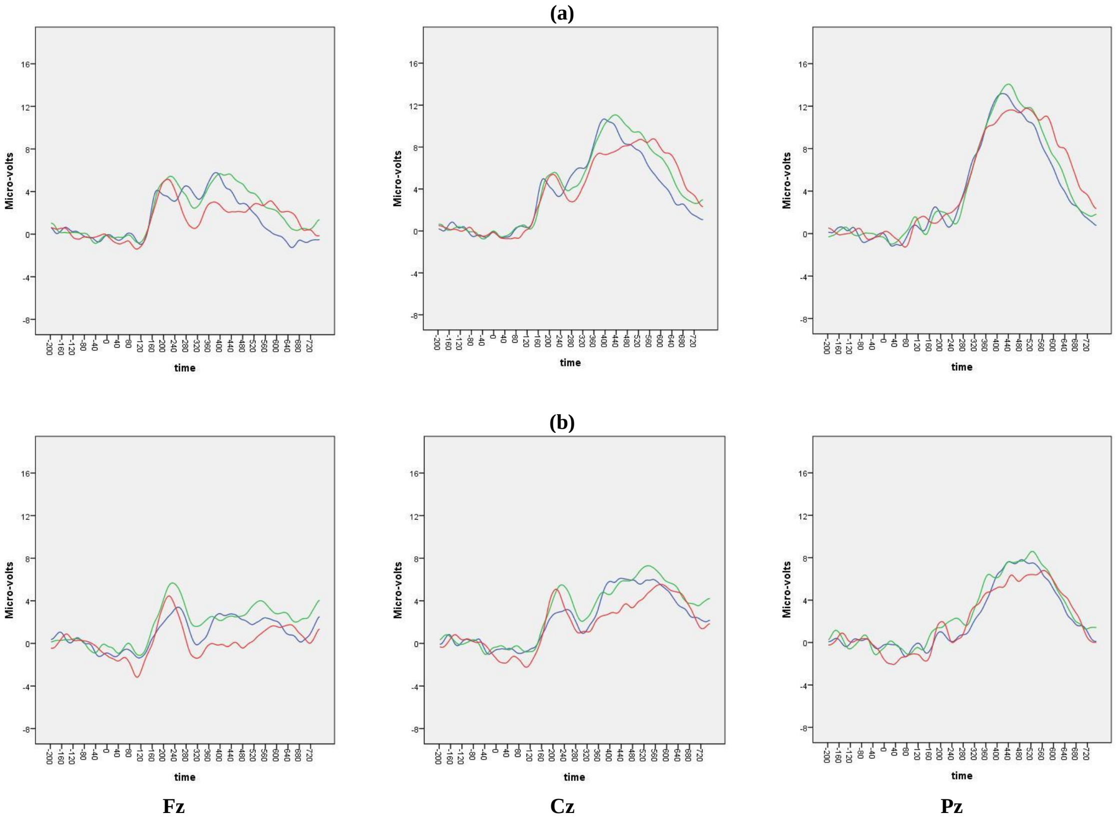


Figure 7.02. Grand average waves for one (blue trace), two (green trace), and three (red trace) items at Fz, Cz, and Pz for enumeration without concurrent memory load (panel a), and with concurrent memory load (panel b).

To test these observations regarding the P3b patterns *within conditions* separate 6 (Antpos) * 3 (Laterality) * 3 (Item) repeated measures ANOVAs were conducted on P3b peak latency and amplitude values for the No-load, and Load conditions. Consistent with these grand-average waves, the P3b window was taken between 300 ms and 650 ms post-stimulus. For the No-load condition examination latencies were significantly longer at three items than at two, $F(2,20) = 7.7$, $p = .003$, $\eta_p^2 = .435$, by 41 ms collapsed across sites ($p = .023$). There were no differences in the Load condition. For No-load a significant Antpos by Item interaction $F(10,100) = 5.2$, $p < .001$, $\eta_p^2 = .342$, suggested that the latencies increased at Frontal (56 ms, $p = .05$, $\eta_p^2 = .665$), Fronto-central (57 ms, $p = .019$, $\eta_p^2 = .593$), Central (77 ms, $p = .005$, $\eta_p^2 = .607$), and Centro-parietal (56 ms, $p = .015$, $\eta_p^2 = .464$) sites. There was no such interaction for the Load condition. This latency behaviour for the No-load condition is broadly consistent with the findings from the pure enumeration task in the first two experiments where P3b latency showed a strong tendency to increase across the subitizing slope. However, the lack of any significant effects on P3b latency in the Load condition is puzzling, and will be examined in more detail in subsequent between-Load condition analyses.

The apparent decrease in P3b amplitude evident in the grand-average waves only approached significance in both conditions, which yielded identical F values, $F(2,20) = 3$, $p = .071$, $\eta_p^2 = .233$, both driven by decreases between two and three items of 1.3μv ($p = .04$) and 1.9μv ($p = .083$) for No-load and Load conditions respectively. Thus, the substantial amplitude reductions seen at three items in the first two experiments were not evident in the current data for the No-load condition. There were no significant effects involving item within either condition.

Additionally, similar to the morphology observed so far there is evidence of P2-N2 deflections which are more apparent Fronto-centrally, and which within the No-load condition at least appear more pronounced at three items. Examination of the waveforms suggests the P2 is apparent across a similar window as has been used so far, thus all P2 analyses are conducted on a window 90 ms to 250 ms post-stimulus. However, the examination of the Load condition’s overlaid waves in Figure 7.03, which follows, indicates a constrained and regular peak occurrence in this paradigm of around 320 ms. Therefore the N2 window will be taken between 280 ms and 360 ms post-stimulus. Within condition analyses on the peak latency and amplitude values within these windows showed that there were no significant effects involving either latency or amplitude changes in the P2, nor the N2.

Together these data illustrate a number of differences in behaviour of the LPC over the subitizing slope in the No-load condition compared to the enumeration task used in the first two experiments. While increases in P3b latency were observed, there were none of the robust decreases in P3b amplitude at three items seen so far, and no indication of the robust decrease so far seen in the N2 window. Interestingly, when enumeration had to occur while maintaining a load in WM there were no P3b latency increases across the items, and no amplitude decreases for either the P3b, or the N2. Clearly, these findings support the conclusion suggested from the RT data—that is, that introduction of the secondary task into the enumeration trial structure disrupts something in the process that was evident in the base enumeration task from Experiments 1 and 2. This disruption seems more dramatic when the secondary task includes stimuli to be retained in WM during subsequent enumeration.

While deviations from the established LPC morphology are important to note, the question of key interest in the current experiment is whether the morphology of the LPC during subitizing differs when subitizing occurs while there is, or is not, a concurrent memory load. Figure 7.03 overlays the waveforms for the No-load and Load conditions for each item level at Fz, Cz, and Pz. It is immediately obvious that at a load of just one item there are substantial apparent differences in the appearance of the waves between the Load conditions, with Load appearing to be associated with pronounced N2 deflections at frontal and central sites, as well as reduced amplitude in the N2 and P3b windows, compared to the No-load condition.

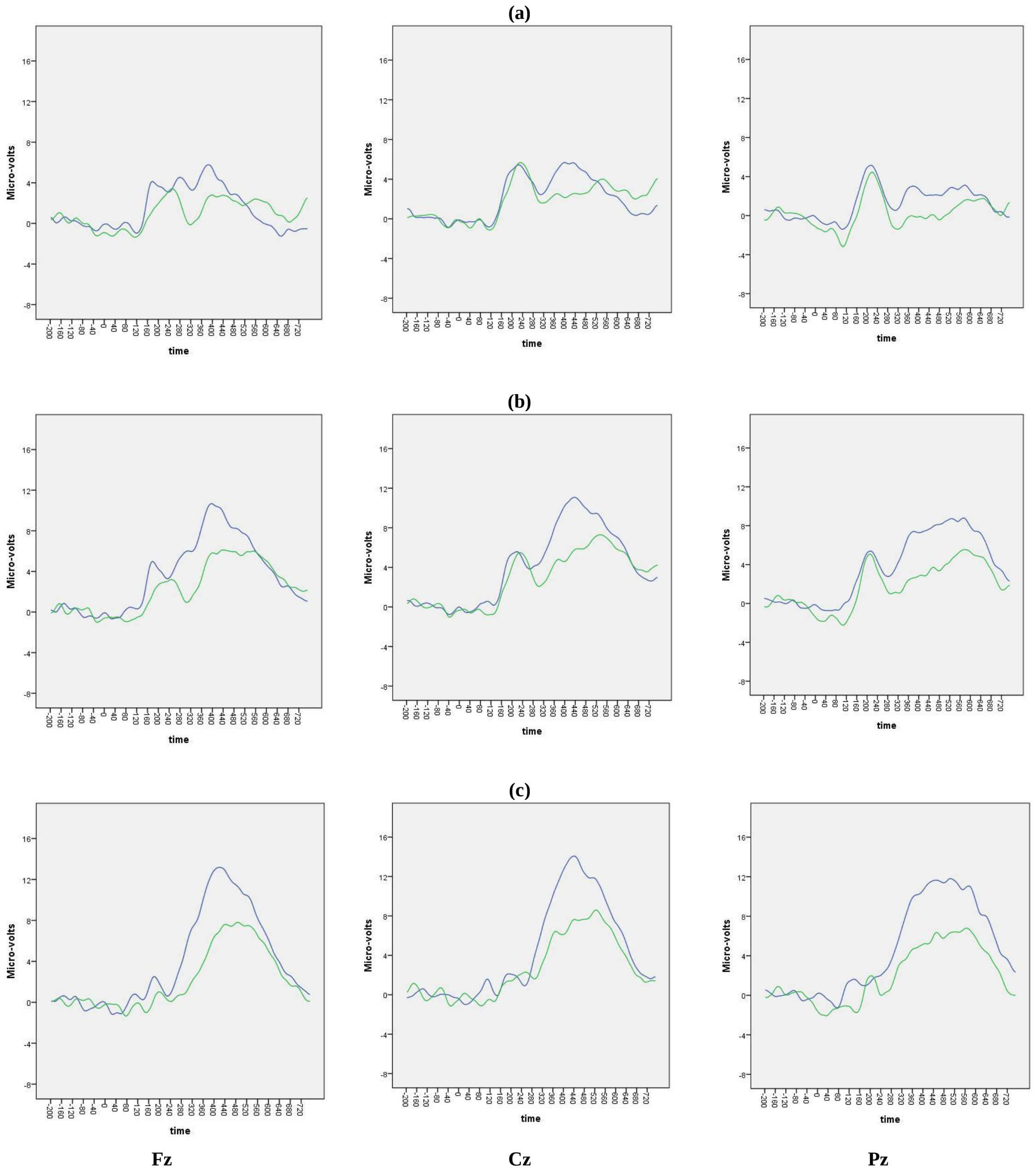


Figure 7.03. Grand average waves at Fz, Cz, and Pz for 1 item (panel a), two items (panel b), and three items (panel c) without (blue trace) and with (green trace) concurrent memory load.

To assess the pattern of any difference between the No-load and Load conditions, separate 6 (Antpos) * 3 (Laterality) * 3 (Item) * 2 (Load: No-load; Load) repeated measures ANOVAs were conducted on P2, P3b, and N2 peak latency, and amplitude values. There were no effects involving the P2, therefore only analyses for the P3b and the N2 are reported.

7.2.2.2.1 P3b and N2 Latency

P3b Latency: Effects on P3b latency were described by a moderately strong effect for Load, F(1,10) = 5.8, p = .037, η_p^2 = .366, where latencies collapsed across Item and Antpos were 40 ms greater for the Load condition. Latencies for Load conditions at each Antpos composite are plotted in Figure 7.04. This effect was embedded in a weak Antpos by Item by Load interaction, F(10,100) = 2.2, p = .021, η_p^2 = .183. This interaction was driven by substantially longer latencies for one item for the Load versus the No-load condition at Frontal (52 ms, p = .06, η_p^2 = .301) Fronto-central (62 ms, p = .019, η_p^2 = .436), Central (68 ms, p = .001, η_p^2 = .665), Centro-parietal (60 ms, p = .006, η_p^2 = .543), and Parietal (44 ms, p = .001, η_p^2 = .548) sites. For two items latencies were longer only at Central (82 ms, p = .019, η_p^2 = .441) and Parietal (48 ms, p = .027, η_p^2 = .402) sites At three items the only difference was at Parietal sites (42 ms, p = .027, η_p^2 = .401). Thus, in terms of the anterior distribution of effects seen in the previous experiments the effect of verbal memory load on increasing P3b latency is clear at anterior sites at one item, but is not significant as item load increases during subitizing, with significant effects for two and three items constrained to central and Parietal sites.

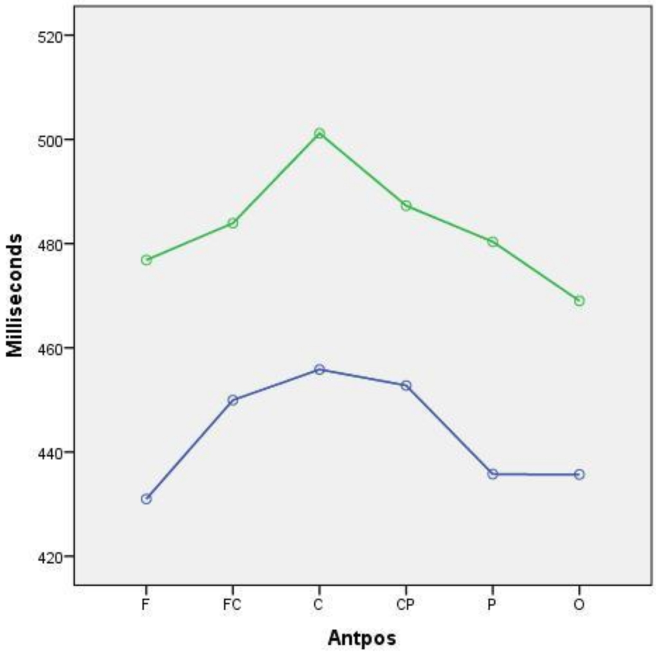


Figure 7.04. P3b latencies collapsed across Item at each Antpos composite by Load. Blue trace = No-load condition, Green trace = Load condition.

N2 Latency: N2 latency was described by a strong effect for Load, F(1,10) = 12.8, p = .005, $\eta_p^2 = .56$. Collapsed across sites latencies in the N2 window were 13 ms longer for the Load condition. No other effects were significant. The changes of interest in the N2 window in Experiment 2 had a fronto-central distribution, so in the current data a maximum effect over anterior sites would be expected if latency increases are reflecting load on the same processes generating the anterior N2 seen in Experiment 2. There is no evidence of any specific effect on latency in terms of scalp topography in this data, and so no claims can be made regarding the relationship of latency behaviour in the N2 window and any *specific* anterior component. However, there are reliable effects at anterios sites. N2 latencies collapsed across Item level are plotted at each Antpos composite by Load in Figure 7.05. It is clear that latency increases related to holding a verbal memory load while subitizing are evident in the N2 window across the scalp.

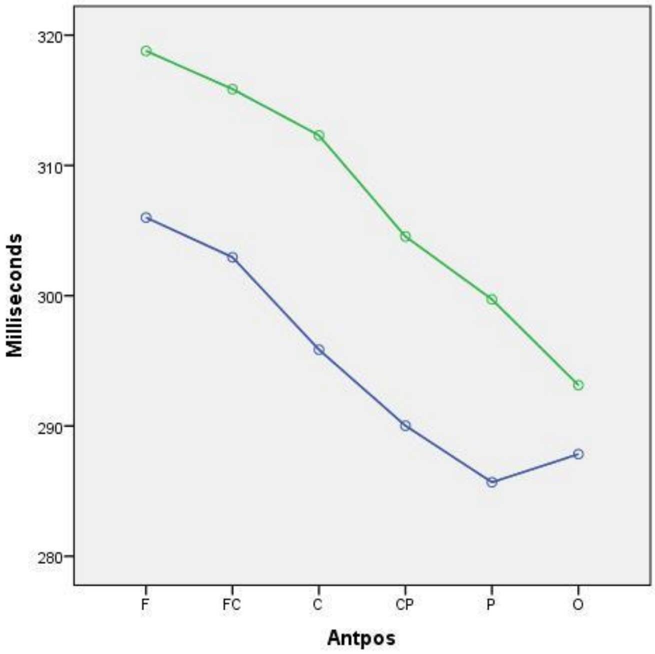


Figure 7.05. N2 latencies collapsed across Item levels at each Antpos composite, by Load. Blue trace = No-load condition, Green trace = Load condition.

The P3b latency data provides strong support for Hypothesis 1, and so for the idea that holding a load in WM utilises resources that are required for subitizing to occur. Latencies across the scalp—including anterior sites—were substantially greater in the Load condition even when only a single item had to be enumerated, indicating that something was delaying context updating. Interestingly, only Central and Parietal sites showed significantly greater latencies at a load of two items, and while numerically greater latencies still occurred at anterior sites at a load of three items, these were not reliably different, with only Parietal sites showing reliable increases. Hypothesis 2 was also supported. Even though the difference was only 13 ms N2 latencies were also reliably longer when subitizing under a memory load. While the anterior specificity of this behaviour was not evident, with similar increases occurring right across the scalp, there *were* reliable increases at all anterior sites. However, these were nothing like the magnitude seen in the first two experiments, and could in no way account for the subsequent substantial increases in P3b latency.

7.2.2.2.2 P3b and N2 Amplitude

P3b Amplitude: There was a very strong effect of Load on P3b amplitude, $F(1,10) = 20.5$, $p = .001$, $\eta_p^2 = .672$, with amplitude, collapsed across Item levels and sites for the Load condition being 3.3µv less than when no Load was held during subitizing. This effect was embedded in a moderate Antpos by Load interaction, $F(5,50) = 5.6$, $p < .001$, $\eta_p^2 = .365$.

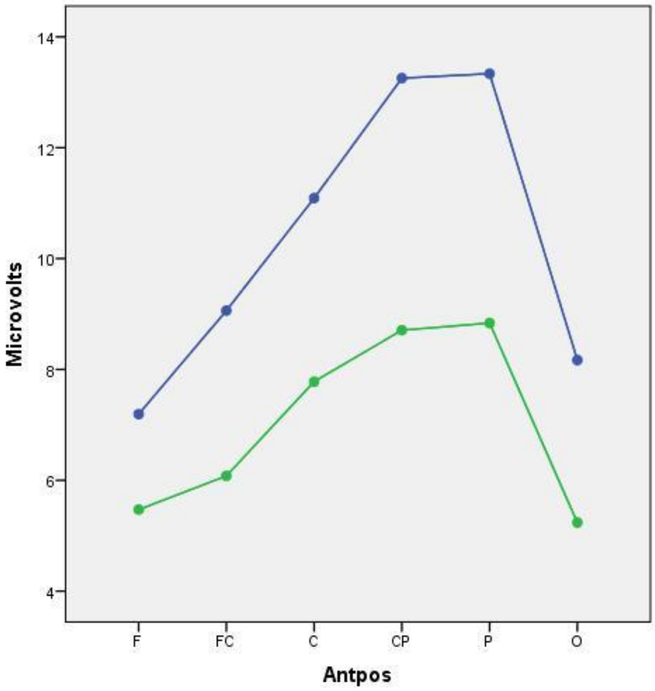


Figure 7.06. P3b amplitudes collapsed across Item levels at each Antpos composite, by Load. Blue trace = No-load condition, Green trace = Load condition.

Examination of the pairwise comparisons showed that with the exception of Frontal sites (which approached significance; 1.7µv, $p = .061$, $\eta_p^2 = .308$) amplitudes were less for the Load versus No-load condition at all other composites, (Fronto-central, 3μv, $p = .003$, $\eta_p^2 = .591$; Central, 3.3μv, $p = .003$, $\eta_p^2 = .591$, Centro-parietal, 4.5μv, $p = .001$, $\eta_p^2 = .696$, Parietal, 4.5μv, $p = .001$, $\eta_p^2 = .71$, Occipital, 2.9μv, $p = .002$, $\eta_p^2 = .647$). Thus, the effect of verbal memory load on decreasing P3b amplitude while subitizing is clear across all item levels at all sites, weakest Frontally and becoming stronger posteriorly through to Parietal sites. The Antpos by Load interaction is plotted in Figure 7.06.

N2 Amplitude: There was a strong effect for Load, $F(1,10) = 7.1$, $p = .024$, $\eta_p^2 = .416$. Collapsed across Item levels and sites N2 peak amplitude was more negative for the Load versus the No-load condition by 2.3μv. The N2 peak amplitude values collapsed across Item levels for each Antpos composite are plotted in Figure 7.07. It is clear that this effect is substantial, and widespread. No other effects were significant.

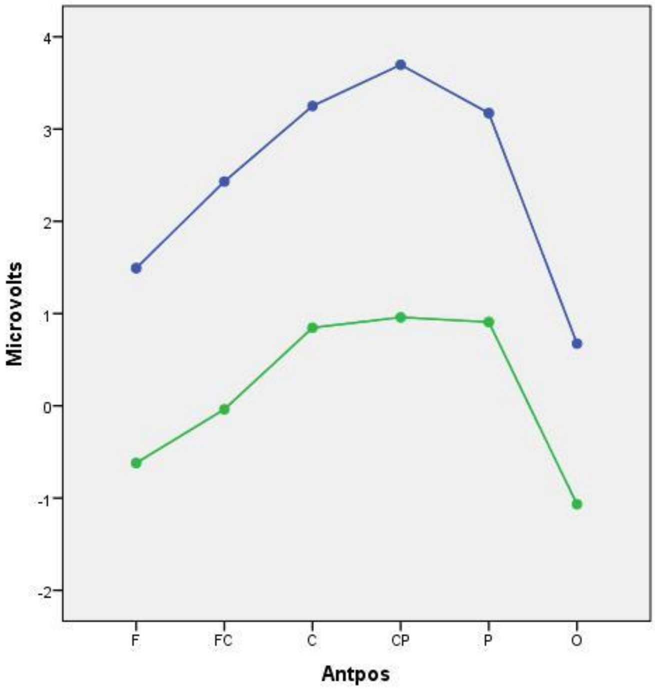


Figure 7.07. N2 peak amplitudes collapsed across Item levels at each Antpos composite, by Load. Blue trace = No-load condition, Green trace = Load condition.

Thus, similar to the findings for the P3b amplitude, the effect of verbal memory load on decreasing relative amplitude in the N2 window is clear from even one item, and there are strong effects at the anterior sites. However, given that the N2 component of interest in this thesis has an anterior distribution, thenotable effects at posterior sites raise some concerns regarding its interpretation. These posterior effects are unlikely to reflect behaviour of the same operations which generate the anterior N2, and so are not of direct relevance to establishing the functional significance of the anterior N2. However, it is possible that these posterior amplitude differences are reflecting more general behaviour in the overlapping P3b window that is related specifically to *maintenance* operations in WM (e.g., Mecklinger, 1992). In terms of the anterior N2 there were substantial numeric differences at the anterior composites that were in the predicted direction for signifying WM load related interference on the component.

Overall, the substantial amplitude decrease of the P3b in response to subitizing even a single item while holding a concurrent memory load was entirely consistent with the profile interpreted so far in this thesis as being related to increased effortful processing, and provides solid support for Hypothesis 3. The co-occurrence of decreased relative amplitude of the N2 at anterior sites is also consistent with this profile, an idea given more credence by the obvious presentation in the grand-average waves of the very distinct N2 component at Central and more anterior sites which precedes reduced amplitude throughout the remainder of the epoch. However, while the obvious N2 is clearly most evident at these more anterior sites the widespread presentation of a consistent negative deflection in the small window used in this experiment suggests only limited support for Hypothesis 4.

7.2.3 Discussion

The ERP waveforms produced by both Load conditions in this experiment conformed visually to the general morphology seen across the first three item levels in the previous experiments. However, looking at the data within conditions made it clear that for *both* load conditions introducing the secondary task to the item enumeration trial structure disrupted the clear pattern of P3b and N2 latency increase and amplitude decrease at a load of three items that has been seen so far in this thesis. For the No-load condition, while there were anterior latency increases, there was no definitive P3b amplitude reduction evident at three items. However, disruption was greater when subitizing while holding a memory load. For arrays of three items there were no clear increases in P3b or N2 latency, or decreases in their amplitude. Disruption of processing stages preceding speeded response is one way that carrying a concurrent memory load can generate interference (Pashler, 1994). While disruption of *both* conditions could indicate that modification of the enumeration task structure may be the source of this interference, there are key differences between the two conditions which suggest the presence of measurable interference effects directly related to the addition of a maintenance load on WM while subitizing. These effects are not the same at a load of three items as they are for a load of one or two items.

Looking firstly at the latency findings for an item load of one, or two, there were differences in P3b latency between the Load conditions, and thus differences in duration of whatever cognitive operations are necessary to achieve task relevant context updating. If, when enumerating the same number of items, the same processes (those which subserve subitizing) take longer to achieve context updating when a concurrent WM load is present, then it is plausible to infer that the load is making the task more difficult. In terms of a resource sharing model of WM such as described by the activation model, it follows that the processes being made more difficult must be processes which utilise, and thus share, WM resources. Importantly, at a load of just a single item (i.e., a single unit of information) context updating took longer when there was a concurrent maintenance load held active in WM. For one item this effect was evident, and strong, at all except Occipital sites The same relationship held true at a load of two items, but only at Central and Parietal sites. This distribution is not entirely consistent with the strong P3b latency increases seen anteriorly at the step to three items in the first two experiments, and so casts doubt on clearly interpreting this behaviour in terms of the anterior model supported from the first experiments.

There were also greater N2 latencies for the Load versus the No-load condition when one or two items were being enumerated, but this this effect was strong across the scalp, again not lending itself supporting the model of anterior changes in the LPC related to WM load that has so far been supported in this thesis. Additionally, while strong, the increase was only 13 ms and could not account for the substantial subsequent delays in P3b latency. In the first two experiments anterior N2 latency increase was associated with subsequent P3b latency increase most notably at the step to three items, and predominantly at Central and anterior sites. This association was interpreted as a marker possibly reflecting the operation of time consuming WM processes which were prolonged due to the additional processing demands of enumerating a third unit of information, with this prolongation contributing to the overall greater time to achieve context updating reflected by the subsequent greater P3b latency. In the current data the memory load related latency delays the N2 do not conform directly to this model and so does not clearly support this interpretation that the N2 peak reflects the timing of some WM related process or processes that are involved in subitizing.

In the earlier experiments this pattern of predominantly anterior latency increases that marked dramatic item load related changes in the LPC was accompanied by substantial amplitude reduction in the N2 and P3b windows. At first glance the current findings of lesser anterior amplitude for these components for the Load condition at both one and two items might seem to fit this profile. However, while there were significant relative decreases in N2 amplitude for the Load condition at Central, Fronto-central, and Frontal sites, significant amplitude differences occurred at the more posterior sites as well, not solely at the anterior sites where the main effects of incrementing item load were seen in the previous experiments. Experiment 2 showed clearly that the N2 amplitude reductions across the subitizing slope were evident only at the more anterior sites, where N2 and P3b latencies also increased. While P3b amplitude reductions were also evident at these same anterior sites, as is often the case with the P3b component (see Kok, 2001) substantial amplitude reductions were also seen right across the scalp, although this was not focused on. The anterior component morphology of the LPC has so far been the focus of this thesis because of the proposed association of anterior, negative going components with WM related activity from around 200 ms post-stimulus (detailed in section 2.1.5.2), and this anterior association with similarly functional behaviour of the P3b strengthened this inference. However, more posterior amplitude reductions in the P3b window have been noted as a response to increased load in dual-task paradigms such as the n-back task (Watter, Geffen, & Geffen, 2001), the flanker task with or without concurrent WM load (Pratt, Willoughby, & Swick, 2011), and a tone identification task with a concurrent Sternberg memory task (Low, Leaver, Kramer, Fabiani, & Gratton, 2009). Reductions in posterior amplitude in the P3b window have also been seen as a function of memory load in semantic memory search paradigms (e.g., Mecklinger, Kramer, & Strayer, 1992). Thus, these topographical changes could reflect the activity of additional neural populations recruited by WM to perform maintenance operations, and these processes might operate across the N2 window as well. Additionally, posterior components in the N2 window, such as the N2pc, have been associated with target detection in the presence of conflicting distractor stimuli, with the suggestion that they reflect attentional filtering or focus in the face of distraction (Luck, 1995). These sorts of processes may have been recruited in order to perform the primary task of enumerating while having attention distracted (utilised) by the secondaryvisually based memory load.

The key observation to draw from these findings is that relatively well established indicators of time consuming WM activity—P3b peak latency increases and amplitude reductions—occur when only a single item needs to be enumerated while a concurrent information load is held active in WM. These robust indicators are accompanied by a substantial change in morphology of the LPC relative to that seen when enumerating a single item without holding any memory load. Referring to the grand-average waves for a load of one item in Figure 7.03, a very distinct N2 component is obvious in the waves across the anterior scalp when a load on WM resources is imposed, whereas no such *pronounced* deflection is evident when there are no additional demands on WM. This difference between the waveforms from the No-load and Load conditions is evident again for arrays of two items. At a load of three items the N2 component becomes much more evident for the No-load condition as well, just as we have seen in Experiments 1 and 2, and both waveforms exhibit the same morphology over the N2 window, indicating devolution of the LPC. Appearance of the N2 component being conditional on WM load leads to the obvious inference that the N2 component reflects some WM related process(es) that respond to resource demands.

Overall, the addition of a verbal information load in WM while subitizing one or two items has generated changes in the ERP behaviour that are similar to those seen at a load of three items when enumeration occurs with no additional memory load, and which are indicative of load related, time consuming processing in WM. Experiment 5 examines what sort of effect the addition of a spatial information load in WM during subitizing might have on morphology of the LPC.

7.3 Experiment 5: Spatial Load

This experiment addresses the same Hypotheses as did Experiment 4, assessing whether there are Load related (*Hypothesis 1*) increases in P3b latency, (*Hypothesis 2*) any associated increases in N2 latency, (*Hypothesis 3*) reductions in P3b amplitude, and (*Hypothesis 4*) N2 amplitude, to examine whether holding a spatial WM load during subitizing influences the morphology of these ERP components in ways that suggest effortful WM activity is involved in subitizing.

7.3.1 Method

7.3.1.1 Participants

Nine adults (7 M, mean age = 29 +/- 14) from Armidale participated. All had normal or corrected to normal vision, and reported being right handed.

7.3.1.2 Stimuli

The spatial memory load condition was operationalised using configurations of four squares bordered in black with solid fill of one of four colours (green, blue, yellow, and red) which in each trial were randomly assigned to each square. Each square covered the four cells in each corner of the 5*5 grid in which the enumeration arrays were presented, such that there was one column of empty cells separating the left and right coloured squares vertically, and a row of empty cells separating the top and bottom squares horizontally. Two squares randomly chosen in each trial had diagonal lines, either ascending or descending left to right (again, orientation was randomly assigned). Thus, stimuli had properties of location of the four colours, and location and orientation of the two diagonals, for a total of eight pieces of spatial information. It was hoped that this amount of information would preclude the use of any verbal memorisation strategies, and subjects were also given specific instructions to not use verbal maintenance strategies. Example stimuli are illustrated in Figure 7.08.

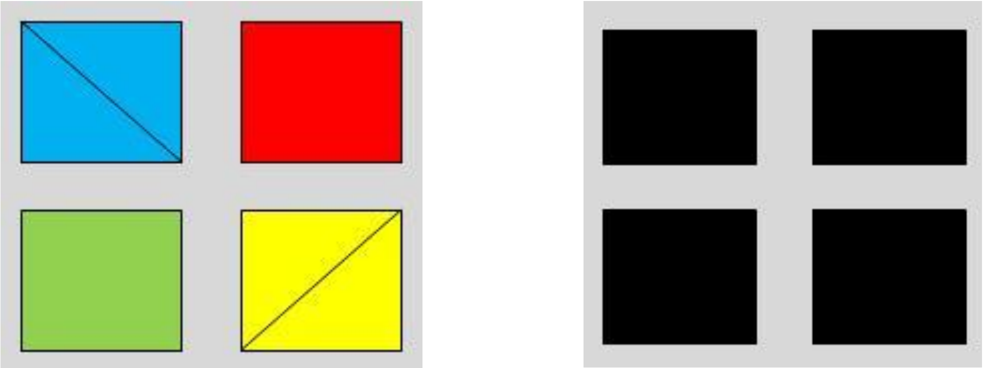


Figure 7.08. Examples of spatial Load stimuli. Left pane illustrates a Load stimulus consisting of four coloured squares with two occupied by a diagonal line. Participants were required to remember position of each colour and position and orientation of each line. Right pane depicts the No-load stimulus, signifying no memory requirement..

7.3.1.3 Task

All timing and display parameters of the trial structure were identical to those used in Experiment 4, with two exceptions. Firstly, the memory stimuli were the spatial arrays described above. Subjects were asked to observe the memory array carefully and note the colour of each square, and the location and orientation of the diagonals, and to try very hard to maintain a memory of these features over the remainder of the trial. Secondly, memory for the stimulus was assessed by requiring a same or different judgement to a probe that was or wasn't the same as the memory stimulus. See below for accuracy criteria. As for Experiment 4, all trials were initiated by the participant. However, given the difficulty of this task the number of blocks was reduced to 9 (from 11) blocks of 42 stimuli per block, with each block containing one array each of 10, and eight items, and two arrays each of nine, and seven items. These six arrays were presented without memory load and were included to offset greatest item density judgements for the largest arrays presented in the Load condition. The remaining 36 trials comprised six exemplars each of arrays of from one to six items, three of which were presented with concurrent memory load, and three which were not. Presentation order was randomised.

Accuracy: Pilot testing indicated that the incidence of interfering events in the lengthy delay between exposure to the stimulus and the recognition probe, along with the complexity of the stimulus, made accurate recognition of the memory array probe difficult. Pilot subjects admitted being unsure on nearly 50% of trials, placing accuracy at no better than chance level. Even so, it was decided not to reduce the complexity of the stimuli, so as to preclude verbal memory strategies that would be easier to implement over fewer pieces of spatial information (e.g., with only two elements to remember: "Left box yellow, line left to right. Right box blue, no line"). This decision was also based on the fundamental aim of the task which, as for Experiment 4, was for participants to carry an active maintenance load in WM *while enumerating one to three items.* Thus, for this experiment *all* trials were scored as correct and so no trials were excluded from analyses on the basis of memory array recognition inaccuracy. To help ensure that the memory stimulus was kept active during enumeration participants were given very clear instructions to try hard to continue remembering the memory stimulus *while they were actively enumerating the array*. Instructions were also given to try very hard to then keep the memory stimulus active for subsequent recall. Between each block participants were also reminded that the task of paramount interest in each memory load trial was to ensure they maintained the memory stimulus active during enumeration. Additionally, similar reminders were provided one or two times *during* each block in a further effort to ensure task compliance.

7.3.1.4 Procedure

The procedure was identical to Experiment 4. Again, the importance was stressed of *remembering* the load stimulus in the Load condition, as was the importance of *not needing to remember* any information in the No-load condition. For the No-load condition subjects were instructed to accurately enumerate the arrays and respond with a button press as fast as possible. For the load condition, subjects were instructed to keep the memory stimulus active in memory while accurately and swiftly enumerating the following dot array, respond by a button press, and then to make a same or different judgement when presented with a probe that was or wasn’t the same as the memory stimulus. For both conditions both speed and accuracy of enumeration were stressed.

7.3.1.5 EEG Recording and Data Reduction

EEG recording and data reduction was identical to Experiment 4.

7.3.2 Results and Discussion

7.3.2.1 Behavioural Data

RT data , presented in Table 7.02 and Figure 7.09, was screened for each subject as in Experiments 1, 2, 3, and 4. RT outliers (72 trials) that were excluded from the analysis comprised 4.9% of the total 1458 trials for the six item conditions assessed in this experiment. A further 96 trials (6.6%) were removed because of artifact, leaving a total of 1289 trials, an average of ~24 trials per subject per item level per condition (~215 trials in each condition). Mean RTs for each item level per condition are presented in Table 7.02. A 4 (Item) * 2 (Load) repeated measures ANOVA showed that as in Experiment 4 RTs do not differ between Load conditions at any item level, nor between item levels within either Load condition, as would be expected with numerically faster RTs for some of the Item levels in the Load condition.

Table 7.02. Average RTs and standard deviations for each item level, by Load. None of the differences between Load conditions were significant.

Items	1	2	3	4
No-load	566 (134)	576 (134)	615 (137)	796 (272)
Load	572 (122)	576 (147)	622 (188)	756 (264)
Difference	17	-5	-6	-40

The RT data thus show that, defined by RT, subjects are still subitizing at an item load of three whether they are holding a concurrent memory load in WM, or not. Difference contrasts confirmed this for both conditions, as in Experiment 4. This finding is consistent with that of Trick (2005), where subitizing was preserved even when there were secondary task demands on WM capacity. However, this supposition is at odds with subject reports regarding the taxing nature of holding the multiple spatial elements in memory.

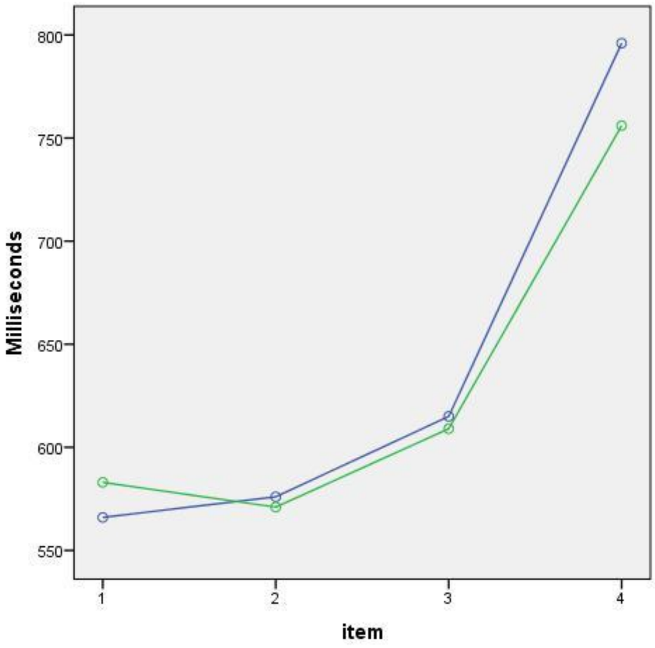


Figure 7.09. RTs for items one to four, by Load. Blue trace = No-load, Green trace = Load.

Finally, it should be noted that, as in Experiment 4 the subitizing slope seen clearly in the first two experiments is not evident in this data, nor is there any interference effect. These observations together cast some doubt on the accuracy of the RT data in terms of the button press response. Again as for Experiment 4, during debriefing subjects reported that the task was very demanding. Their reports were in contrast to those from the first two experiments, where subjects reported going into a ‘zone’ where they had relaxed focus on the trial progression in each block. Without being able to relax into a rhythm it is possible that subjects hurried their button press responses to two, and to three items in order to reach the recall prompt as soon as possible, and so be seen to perform well on the memory task, and that this lack of focus could have carried over to performance in the No-load trials. In any case, whether active (stimulus to remember) or passive (no stimulus to remember), introduction of the secondary task into the enumeration trial structure has obviously disrupted something in the process that was evident for enumeration alone.

7.3.2.2 EEG Data

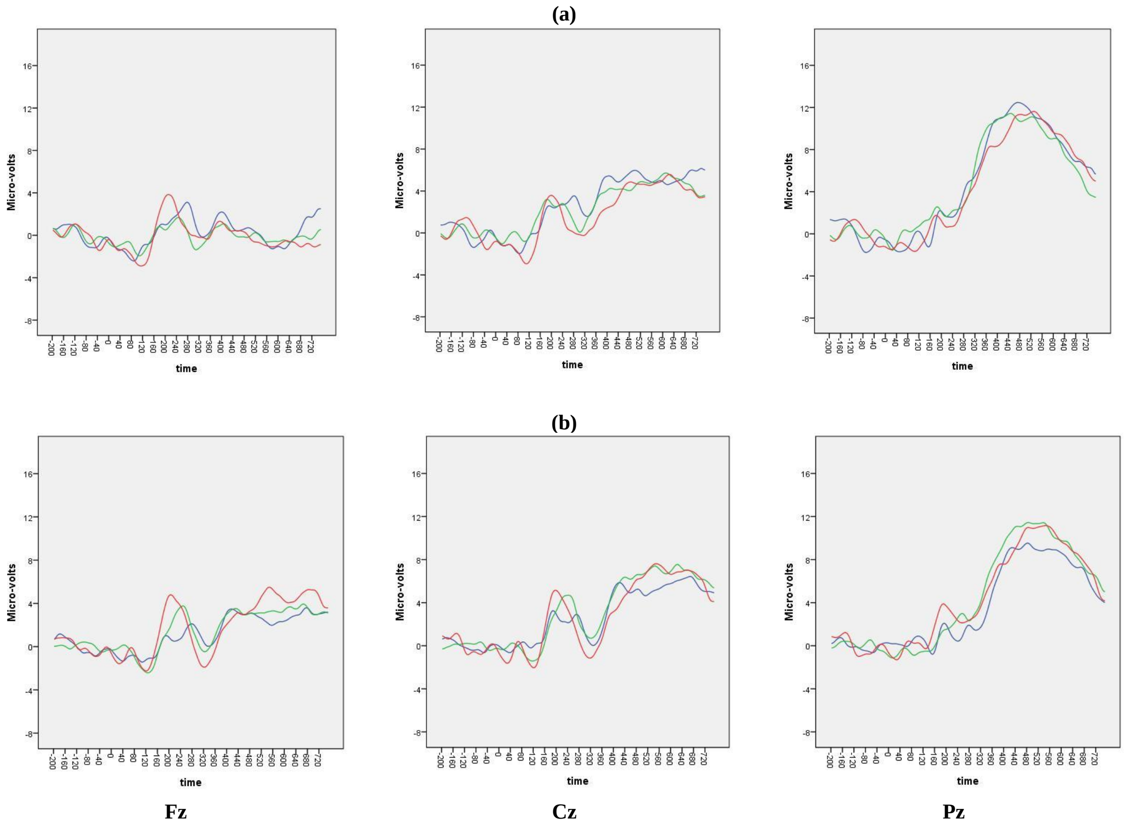


Figure 7.10. Grand average waves for one (blue trace), two (green trace), and three (red trace) items at Fz, Cz, and Pz for enumeration without concurrent memory load (panel a), and with concurrent memory load (panel b).

Examination of the grand-average waves in Figure 7.10 suggests that the morphology of the LPC across the subitizing slope does not clearly conform to the robust profile observed in the first two experiments. Firstly, in both the No-load and

Load conditions there is no indication of any obvious P3b peak latency increases, or of clear amplitude reductions, which were observed for three items in the ‘pure’ enumeration paradigm used in these previous experiments. These observations were confirmed by separate 6 (Antpos) * 3 (Laterality) * 3 (Item) repeated measures ANOVAs on P3b peak latency and amplitude values for the No-load, and Load conditions (using component window identical to Experiment 4; 300-650 ms). There were no significant effects for Item on P3b latency or amplitude for either memory load condition. Thus, as was found in Experiment 4, alteration of the trial structure to include presenting the additional secondary task components appears to disrupt the processing which underlies context updating during enumeration even on trials where no memory load is presented. Additionally, the disruption is seemingly greater here, where the secondary task utilises spatial processing resources, than that associated with the verbal stimuli used in Experiment 4.

Secondly, the waveforms for both conditions exhibit marked P2-N2 deflections and except at Parietal sites, less distinct P3b components than seen in all of the previous experiments. Interestingly, the N2 deflections exhibit almost identical latencies as those in Experiment 4, justifying the use of the same component window (280-360 ms). The P2 also falls in the same latency window as the verbal paradigm, justifying choice of the same component window of 90-250 ms. Within conditions, neither the P2 or N2 show any no significant effects involving either latency or amplitude changes.

However, as for Experiment 4 the question of key interest here is whether the morphology of the LPC across the subitizing slope differs when subitizing occurs while a concurrent memory load is held active—in this case one involving the use of spatial, not verbal, processing resources. Figure 7.11 overlays the waveforms for the

No-load and Load conditons for each item level at Fz, Cz, and Pz. It is immediately obvious that there is remarkable similarity between the conditions, notably regarding presentation of the N2 deflection, and to a lesser extent the P2. Again, there were no significant effects in the P2 window so only results for the P3b and the N2 components are reported.

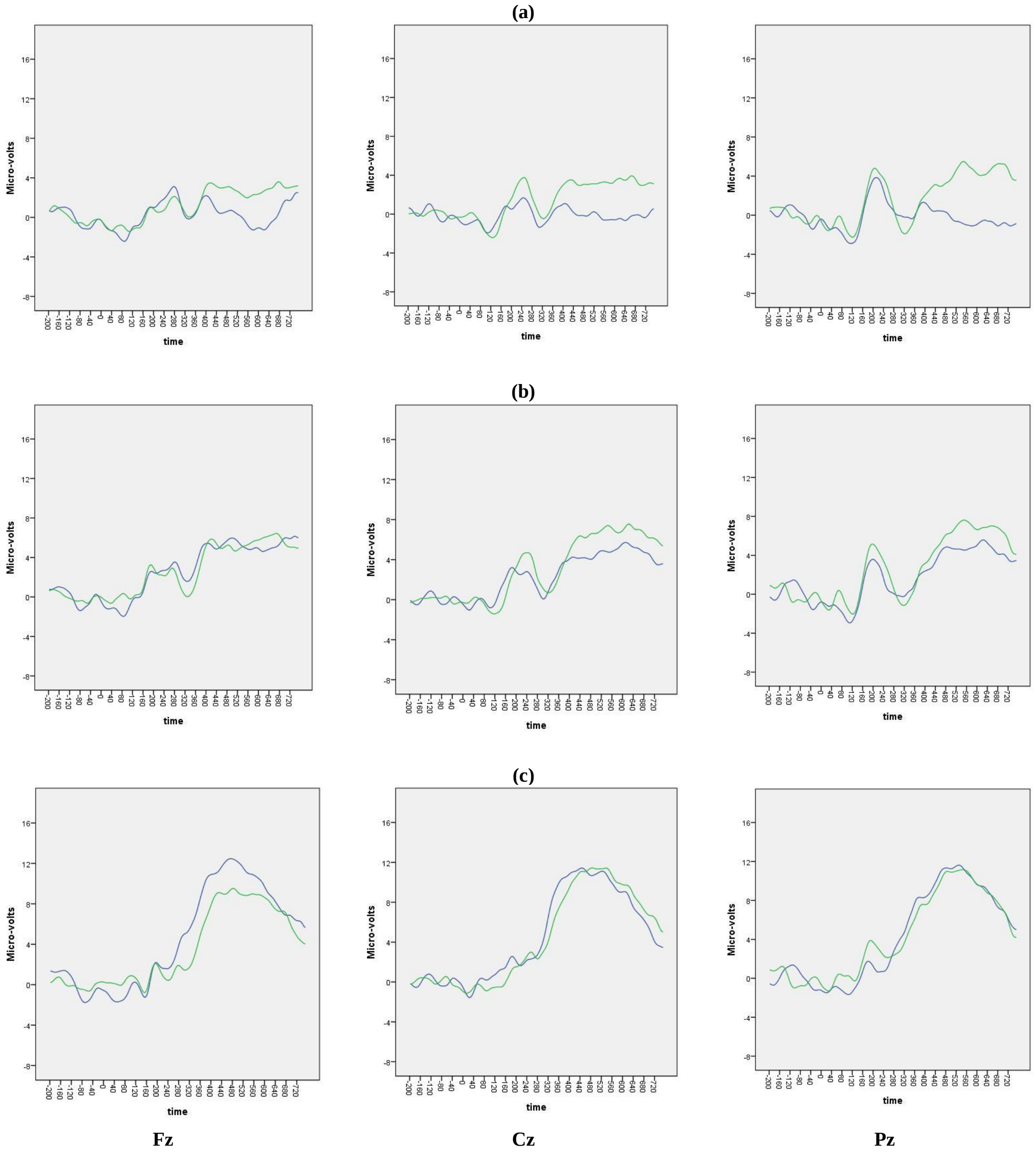


Figure 7.11. Grand average waves at Fz, Cz, and Pz for 1 item (panel a), two items (panel b), and three items (panel c) without (blue trace) and with (green trace) concurrent memory load.

To assess the pattern of any difference between the No-load and Load conditions, separate 6 (Antpos) * 3 (Laterality) * 3 (Item) * 2 (Load: No-load; Load) repeated measures ANOVAs were conducted on P3b, and N2 peak latency, and amplitude values. There were no effects involving the P2, therefore only analyses for the P3b and the N2 are reported.

7.3.2.2.1 P3b and N2 Latency

P3b Latency: There was a significant moderately strong effect for Load, F(1,8) = 5.8, p = .042, $\eta_p^2 = .422$, with longer latencies for the Load condition of 35 ms collapsed across Item levels and sites. There was also a weak Antpos by Load interaction which didn’t reach significance, F(5,40) = 2.1, p = .097, $\eta_p^2 = .201$. This suggested effect was driven substantially by a large difference at Frontal sites (64 ms, p = .002, $\eta_p^2 = .706$), and further differences at Fronto-central (34 ms, p = .008, $\eta_p^2 = .605$) and Parietal (51 ms, p = .05, $\eta_p^2 = .394$) sites. No other effects of Load on P3b latency approached significance. Latencies for each Load condition are plotted at the Antpos composites in Figure 7.12 (following page).

N2 Latency: For N2 latency there were no significant effects involving Load. Latencies for each Load condition are plotted at the Antpos composites in Figure 7.13, which illustrates almost identical behaviour in both conditions.

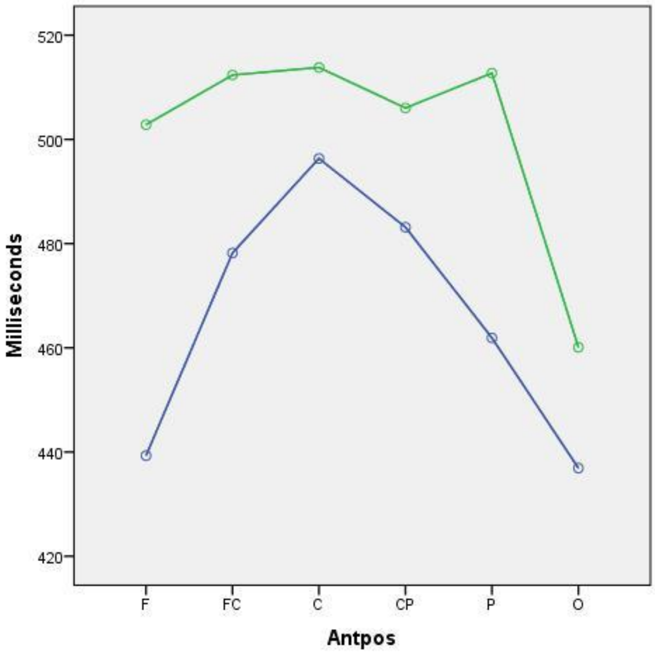


Figure 7.12. P3b latencies collapsed across Item at each Antpos composite, by Load. Blue trace = No-load condition, Green trace = Load condition.

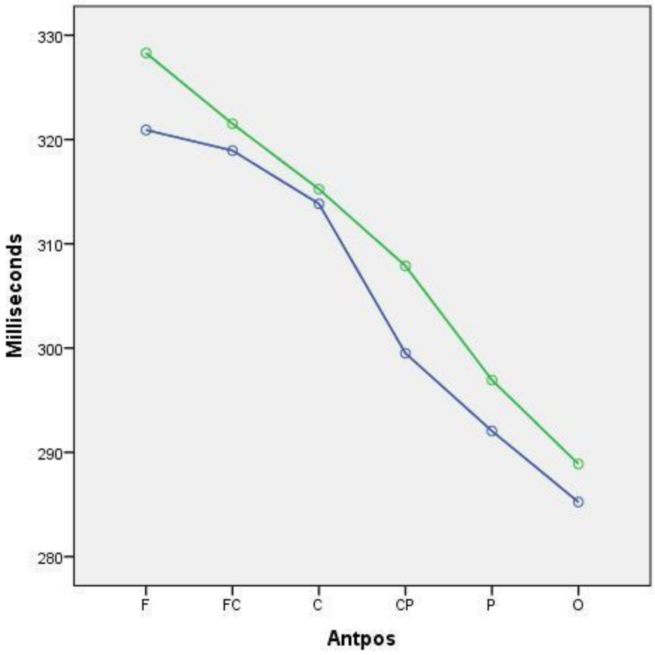


Figure 7.13. N2 latencies collapsed across Item at each Antpos composite, by Load. Blue trace = No-load condition, Green trace = Load condition.

The P3b latency data support Hypothesis 1. In terms of the findings so far in this thesis, and the substantial literature suggesting that increased demand on WM resources results in delayed context updating reflected by increased P3b latency, the P3b data are consistent with the inference that holding a spatial memory load while subitizing prolongs the process, and so the idea that WM resources are involved in subitizing. However, the lack of any associated Load related increase in N2 latencies does not support Hypothesis 2, and calls into question the timing relationship between the two components that was proposed from the results of the first two experiments.

7.3.2.2.2 P3b and N2 Amplitude

P3b Amplitude: P3b amplitude was described by a moderate Antpos by Load interaction, F(5,40) = 5.3, p = .001, η_p^2 = .396. Surprisingly, this effect appears driven by *greater* amplitude for the Load condition versus No-load at Frontal (2.1μv, p = .06, η_p^2 = .247), and Fronto-central (1.6μv, p = .098, η_p^2 = .213) although neither contrast reached significance. No other effects were significant. Amplitudes for each Load condition are plotted at the Antpos composites in Figure 7.14.

N2 Amplitude: N2 amplitude was described by a moderate Antpos by Load interaction, F(5,40) = 3.2, p = .017, η_p^2 = .283. This effect appears driven by *greater negativity* for the Load condition versus No-load not anteriorly, but at Occipital sites (1.7μv, p = .56, η_p^2 = .341), while numeric differences at Parietal (1.4μv, p = .196, η_p^2 = .199) and Centro-parietal (1.3μv, p = .185, η_p^2 = .208) sites did not approach significance. No other effects were significant. Amplitudes for each Load condition are plotted at the Antpos composites, in Figure 7.15.

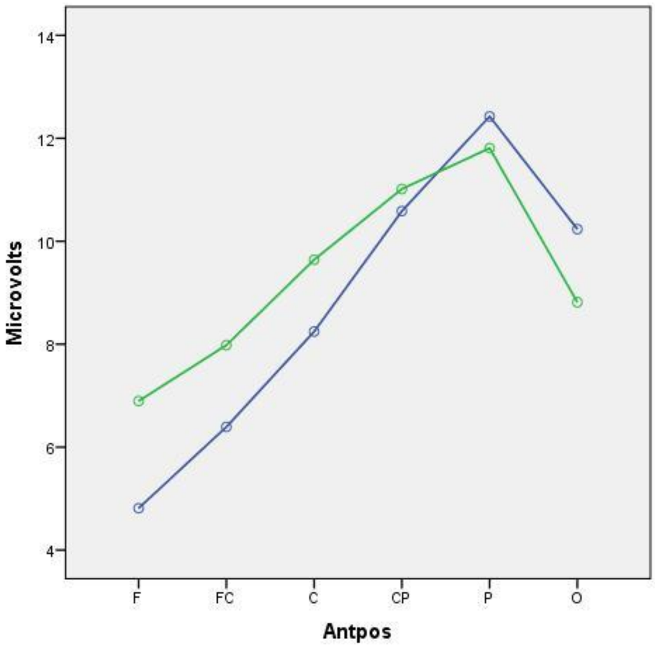


Figure 7.14. P3b amplitudes collapsed across Item at each Antpos composite, by Load. Blue trace = No-load condition, Green trace = Load condition.

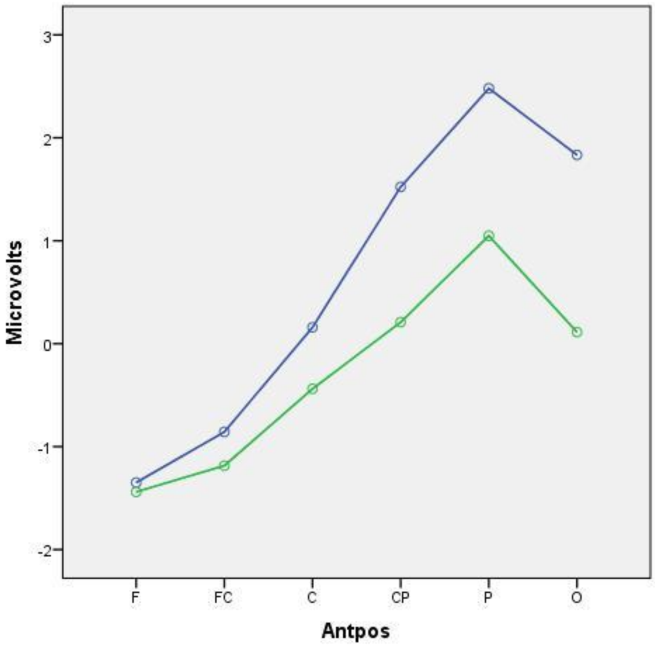


Figure 7.15. N2 amplitudes collapsed across Item at each Antpos composite, by Load. Blue trace = No-load condition, Green trace = Load condition.

This amplitude behaviour of the P3b and N2 components does not support either Hypothesis 3, or Hypothesis 4. It is entirely inconsistent with the pattern seen throughout the research reported in this thesis, where increased processing load has reliably been associated with reductions in P3b amplitude, and relative N2 amplitude. Additionally, the co-occurrence of these reductions has been mutually at more anterior sites. Here we see increased positivity for the P3b anteriorly, while in the N2 window the amplitude is reduced for the Load condition at the most posterior sites, with no effect at all anteriorly. This issue will be expanded on in the discussion.

7.4 Discussion and Conclusions

Reaction time data: Before going on to discuss the ERP findings it is necessary to address the lack of interference effects exhibited by the RT data in both the verbal and spatial Load paradigms. A defining feature of dual task paradigms is that the secondary task *interferes* with performance on the primary task. Generally this results in degraded performance on the primary task and a measure of this performance decrement can provide an estimate of the amount of interference the secondary task has introduced, and thus a measure of its effectiveness in influencing the resources the primary task utilises. A secondary task which utilises substantial resources required for the primary task will produce large interference effects and one which utilises minimal shared resources will produce small ones. The lack of any significant interference effects in the RT data for both of these memory load paradigms begs the question whether the memory load did have any effect on processing during subitizing and so, of course, questions the veracity of making any inferences regarding Load related differences in the ERP data. However, the reliable increases in P3b latency between load conditions in both paradigms provide an objective indication that the Load manipulations were degrading enumeration by delaying context updating, and so some indication that the secondary tasks were in fact influencing performance. In this case it is plausible to suggest that the lack of significant RT interference effects could be due to using a button press response rather than a voice trigger to record RTs, with subjects perhaps often responding prior to actually being able to articualate the number word. This inaccuracy could effectively mask any small differences in RT and so hinder finding any interference effects.

ERP data: Throughout this thesis arguments have been put forward suggesting that distinct appearance of the N2 component of the LPC and subsequent amplitude reduction through the P3b window during subitizing could reflect load related demands on WM resources, and that these demands could be time consuming within a single episode of context updating during which subitizing occurs. In Experiments 1 and 2 this ERP profile became clearly evident at the final level of the subitizing slope, that is, with the addition of a third item to the array (and continued to evolve at a load of four items). The dominant feature of this morphology has been devolution of the fronto-central LPC with emergence of the distinct deflection of the N2 pulling the waveform sharply negative, in the process defining the P2-N2 complex and the continued amplitude reduction through to the subsequent P3b peak. These marked amplitude reductions at a load of three items have been interpreted as indicators of load related demands on WM resources, a position strengthened by associated increases in context updating time reflected by prolonged P3b latency. Experiment 3 showed that the N2 was not generated simply as a response to increased item density, or by processes involved with pattern matching, or paired association. It was only when three items needed to be *enumerated* that the N2 became distinct, and this was accompanied by longer N2 and P3b latencies than for the pattern matching condition. In terms of the literature presented in Chapter 2 the appearance of these indicators of processing load and time consuming processing would suggest that something specifically related to enumeration demands these resources during subitizing. However, this argument still relied on speculation that WM activity is generating the N2, and so causing devolution of the LPC.

The primary aim of the two experiments reported in this Chapter has been to directly address the question of the functional significance of this ERP profile as an indicator of WM activity. In Experiment 4 the addition of a secondary task requiring maintenance of verbal information in WM while enumerating generated behaviour of the LPC that was firmly consistent with this interpretation of the morphology of the LPC. Context updating took longer when enumerating even one, or two items under memory load, and there was accompanying P3b amplitude reduction relative to the No-load condition. Additionally, the grand-average waveforms showed a more distinct, large N2 for the Load condition than for the No-load condition for even one or two items, and there was clear statistical support for the addition of verbal WM load being related to decreased amplitude in the N2 window.

Experiment 5 utilised a spatial WM load. The four hypotheses tested here assessed how this LPC morphology might reflect effortful WM activity as a result of carrying a spatial memory load during subitizing The results for this spatial Load paradigm are not generally consistent with this morphology of the LPC. Only Hypothesis 1 was supported—P3b latency was longer for the Load condition, predominant at Frontal sites, a distribution consistent with changes in the fronto-central LPC indicative of effortful processing in WM, and a finding that has emerged in each experiment so far—and there was no support for Hypothesis 2, with no associated increases in N2 latency. However, the lack of support for Hypotheses 3 and 4 addressing the amplitude behaviour of the P3b and the N2 raises interesting questions regarding the possibly different processes influenced by carrying a spatial rather than a verbal WM load during subitizing. There were significant P3b amplitude *increases* as a function of item load, predominant anteriorly, and while N2 amplitude was less for the Load versus the No-load condition, these differences were solely at the most posterior sites.

These are atypical patterns. Increased processing or memory load reliably leads to decreases in P3b amplitude, as was clearly evident in the verbal load paradigm, and so far all load related increases in negative going activity of the N2 have had anterior distributions. The *opposite* effects seen here are puzzling, and might indicate that holding a spatial versus a verbal Load while subitizing recruits different processes (generated by different brain regions) to undertake the dual tasks. As discussed in section 2.1.5.2, there are a range of ERP components which present in the 'N2' window and their functional significance is determined by eliciting conditions, and scalp distribution. One such component is the N2c, which has a posterior distribution for visual stimuli and is thought to reflect a subprocess of stimulus identification or classification (Daffner et al., 2000), particularly under conditions where a target stimulus must be chosen from conflicting distractor stimuli (Luck, 1995). Thus, it is thought that the N2c reflects the operation of an attentional filter. N2c amplitude increases with difficulty in stimulus categorisation or identification. It is possible that visual attentional resources are severely taxed by attempting to maintain the complex spatial stimuli in the current paradigm, and more of these early attentional resources are needed to choose, or identify, the spatial properties of the dot stimuli in the face of the competing internal stimuli held active in memory, thus making categorisation or identification more difficult and explaining the greater posterior negative activity for the Load condition. Given this scenario it is possible that attentional/WM resources normally used to bind the dot stimuli into WM, and which may be reflected by the anterior N2, are superceded by the required processes of identification and classification reflected by the N2c. If it is the negative going activity reflected by the anterior N2 which contributes to attenuating the amplitude of the positive P3b wave ('pulling it negative' as was suggested in section 2.1.5.1.2) then reduction of this activity could explain the increased P3b amplitude when subitizing under a spatial memory load. Bear in mind that in the activation model of WM outlined throughout this thesis attention and WM are intimately related constructs, and so this explanation appeals to increased utilisation of WM resources as driving the pattern of reduced posterior N2 amplitude and increased anterior P3b amplitude, and so suggests that this pattern is an indicator of effortful WM activity just as it is suggested of the more anterior devolution of the LPC.

Even without appealing to this speculative model to explain the waveform morphology in this spatial paradigm as reflecting WM activity during subitizing, there is strong evidence from the verbal memory load paradigm of WM involvement in enumerating, even only a single item. In the verbal paradigm carrying a memory load was associated with increased P3b latency and decreased P3b amplitude. These findings directly confirm an association between these features of P3b morphology and increasing demands on WM resources. As such, relative to enumerating with no load, the P3b latency and amplitude differences when enumerating just one item while holding a verbal information load active in WM suggest that WM resources are involved with subitizing not only two items, but even a single item. This view is consistent with the attentional blink findings of Olivers and Watson (2008), the dual-task findings of Vetter et al. (2008), and those of Poise et al. (2008) using time course of visual processing, which were reviewed in section 1.2.4, who all found effects of attention on subitizing one item.

CHAPTER 8

Summary and Conclusions

8.1.1 Summary

The primary aim of the research program reported in this thesis was to examine the role of WM activity during item enumeration of response contingent displays. The specific approach was to use ERPs to assess whether the characteristics of subitizing are determined by preattentive parallel processes, or by capacity constraints on serial WM activity. While there has been renewed interest over the last five years in examining the possibility that attentional resources were necessary to achieve successful subitizing (see section 1.2.4), strangely, no comprehensive studies to date have used ERPs to investigate this question. ERPs would seem like the perfect tool to investigate possible WM involvement in subitizing because they provide online information about processing activity from stimulus presentation right through to response selection and execution. Robust ERP components such as the P3b appear to change as a function of processing load and task difficulty. This is presumably due to demands on immediate processing resources, which are likely to fall within the construct of WM.

The conceptual framework of this research project was derived as a result of initial speculation regarding a possible relationship between increasing P3b peak latency across the commonly identified subitizing slope, and the small, non-equal numeric increases in RT which define the increments of that often non-linear slope. The temporal window in which the P3b might peak during some attention demanding task could stretch from ~400 ms to ~650 ms post-stimulus, which encompasses the window during which subitizing of three or four items is generally completed in the adult population. This led to the speculation that if increases in P3b latency did occur across the slope they may indicate involvement of time consuming WM activity in determining the non-linear slope, and thus in subitizing.

Another property of P3b latency forced this idea to be refined further with the concept of a *processing episode*. P3b latency only increases so far, as a function of task difficulty, before it reaches a point where the processing required to complete the task (achieve task relevant context updating) cannot be completed within a single episode of context updating in WM. In these circumstances RTs continue to increase, and while P3b components are still generated their latency does not continue to increase. These characteristics suggested that the period up until the P3b peak might reflect a single episode of context updating in WM which for smaller information loads involves processing task relevant information through to task completion, and thus can be seen as a single processing episode. However, there is a finite limitation on the amount of information that can be processed within the period of a single processing episode. Given that the time course of a processing episode and the time it takes to subitize fall within the same window, if WM resources are involved in subitizing then there could be a relationship between the capacity of a processing episode and the well established subitizing limit of (mostly) three or four items.

A model of WM was developed to explain what may be occurring over those few hundred milliseconds of a processing episode. In contrast to a common understanding of WM activity involving information being held active and manipulated over some seconds, the *activation model* views WM as a dynamic system involved in the moment by moment binding of representations into integrated representations (chunks) which become contents of consciousness. This idea of integration or binding together of representations as a mechanism underlying cognitive processing operations in WM is consistent with the role Baddeley (2000) proposed for the episodic buffer component of his multi-component model of WM. Importantly, this type of dynamic integration occurs during the course of processing information through to consciousness, and can therefore be seen to operate prior to and in service of integrated representations arriving in consciousness (e.g., the numerosity of an array) and signalling the end of an episode of context updating.

Within the activation model of WM, the *serial counting explanation* was put forward to offer a perspective on the type of processing operations which might underlie enumeration within the subitizing slope, and the type of WM capacity that may be reflected by the subitizing limitation. In terms of enumeration it was proposed that the operations which proceed during the course of a processing episode might involve keeping representations of the item(s) active while assigning quantification identity to them, followed by binding the resulting intermediate products in terms of some addition algorithm. Once additional items cause these intermediate operations to proliferate, limitations on the activation available to keep representations active and to perform binding operations on them in the course of a single processing episode would emerge, reflected by prolonged RT.

While P3b latency considerations inspired this project there are other features of P3b behaviour which might reflect time consuming, effortful processing in WM. In some paradigms which ostensibly load WM and which have levels of increasing difficulty, as difficulty increases the relatively unitary P3b peak seen in simple attentional paradigms can show great increases in latency which are associated with major changes in the waveforms' morphology. This *devolution of the fronto-central P3b*, entailing the emergence of the N2 deflection interceding between the prior P2 and the following P3b peaks, was introduced in detail in Chapter 2. While there is an extensive literature relating P3b behaviour to types of processing difficulty, and time-course of processing involved in context updating in WM, the picture for the N2 was far more diverse, with the component being claimed as an indicator of more paradigmatically constrained behaviour, such as inhibition (e.g., Tillman & Wiens, 2011) or visual object identification and memory search (e.g., Schendan & Kutas, 2002). An attempt was made to interpret the range of different circumstances under which this type of negative deflection in the N2 window reflected processing demands requiring integrative processing in WM. Thus, the second focus of this research program was to assess the functional significance of the N2, and add to the literature on functional significance of the P3b.

A straightforward research design was sufficient to address both of these questions. Because there had been no previous ERP research examining item enumeration using response contingent displays the first step was to identify the ERP profile generated by enumerating when subitizing items under these display conditions, and assess the characteristics of this profile for any explanation of the subitizing slope (Experiments 1 & 2). This profile could then be assessed in the context of competing explanations of subitizing (pattern recognition and associative processes), while controlling for the possible confound of item load (Experiment 3). The final step was to contrast behaviour of this ERP profile while enumerating under conditions which directly loaded WM, and those that didn't (Experiments 4 & 5). This allowed component behaviour of the LPC to be directly related to WM activity, or not.

Experiment 1 presented subjects with a simple item enumeration task, with the aim of firstly identifying any behaviour of the LPC that might reflect the operations which occur during subitizing, and when subitizing is exceeded, and secondly, to assess whether there were any indications of serial WM activity across the subitizing slope. RT showed there was a non-linear subitizing slope for this sample. Across the subitizing slope the LPC changed morphology in response to the number of units of information that had to be accurately enumerated. As item level increased across the slope, the behaviour of the LPC components were consistent with what would be expected from increased processing demands on WM. P3b latency increased between arrays of one and two items, showing that it takes longer to accomplish the necessary operations to enumerate an array of two items than for a single item. Importantly, at the step to three items P3b latency increased again, this time by a substantially larger increment. These findings suggested that the two RT increments of the non-linear subitizing slope reflected meaningful increases in processing time in WM that supported the serial counting explanation of the subitizing slope over the list search explanation. Additionally, when information load rose from two to three units the lengthy increase in P3b latency was preceded by a substantial negative deflection at both the frontal and central sites that appeared as part of a P2-N2 complex. This was followed by a substantial decrease in amplitude of the delayed P3b. Finally, timing of the N2 peak and the P3b peak were both related to increases in RT across the slope.

This constellation of changes in the LPC morphology at the final level of the subitizing slope were together consistent with what would be expected from increased demand on WM resources, and so suggested that the defining feature of the non-linear subitizing slope, a numerically larger RT step between two items and three items than between one item and two, may be determined by serial processing in WM that subserves quantification. The indication that this type of WM activity occurs during subitizing (as subitizing is defined by RT) is theoretically consistent with the suggestion that the subitizing limit may be related to limitations in capacity to engage in serial binding and maintenance operations across a single episode of context updating. That is, by limitations in chunking capacity.

Finally, examination of the waveforms for the two subitizing span groups (those with an RT derived span of 3 items, and those with a span of 4 items) showed strong similarity between the groups in morphology across the first four item levels, thus indicating a robust effect considering half the number of observations contributed to the high-span group’s waves. At three items the addition of an extra unit of information was associated with observable amplitude reduction for *both* span-groups, with divergence from the waves for two items clearly beginning around 200 ms post-stimulus for both groups. The key finding from these data was that amplitude reduction beginning around 200 ms post-stimulus and appearance of the distinct N2 component became evident at a load of three items for *both groups*. Thus, instead of there being a discrete change in morphology directly related to the subitizing *span*, the indication was that there may be a distributed effect on the N2 amplitude as a function of item load, and subitizing span, with change in the N2 amplitude occurring over two item levels for the high-span group, and over only one overlapping level (i.e., three items) for the low-span group.

Overall, the data from Experiment 1 identified an ERP profile across the subitizing slope which illustrated marked devolution of the LPC as a function of increased item load, the most obvious feature of which was a dramatic change in morphology at a load of three items with the appearance of a large N2 deflection defining a P2-N2 complex, and followed by prolonged P3b latency. The profile provided strong indication from the P3b latency behaviour that serial WM activity beyond that required for list search alone is implicated in generating the non-linear subitizing slope evident in this sample. This general ERP profile was described by strong statistical effects, and was robust, being clearly evident in average waves from the span-group subsets. Finally, these load related changes in morphology at three items were evident for people with RT derived subitizing spans of three *and* of four items, indicating that the loading of WM resources may be distributed across more than a single item level (at least for high-span people).

Experiment 2 was a replication of Experiment 1 using the (then) newly acquired Neuroscan amplifiers. This experiment confirmed the robust ERP profile presented by the LPC across the subitizing slope, and through to four items (when half of this sample were still subitizing), with effect sizes again being strong. The comprehensive scalp coverage allowed with this equipment also confirmed the proposed anterior distribution of the negative going activity in the N2 window which defines devolution of the LPC across the subitizing slope, which was evident in the frontal and central sites examined in Experiment 1. This sample also generated a non-linear subitizing slope, where contrary to the findings from Experiment 1, P3b peak latency did not increase significantly between one item and two, but only at the step from two to three items. Thus, the argument suggesting that the non-linear subitizing slope is determined by gradually increasing load on WM resources during subitizing could not be supported as it was in Experiment 1. However, there was a clear pattern of numeric latency increases from one to four items, described by significant, very strong linear contrasts and maximum at Frontal, and Fronto-central sites. The latency increments of the P3b peak were sensitive to the RT increments from two to three items, and from three to four items. This sensitivity was maximal at fronto-central and central sites. At these same two steps the latency increases of the N2 peak were sensitive to those of the P3b, maximal fronto-centrally, and they were both significant. The amplitude data across the first four items in this experiment was also generally consistent with the findings from Experiment 1, although the movement in the N2 window was not as pronounced. Nevertheless, average amplitude in the N2 window showed a strong trend to reduce at both three and four items, and P3b amplitude reliably reduced at three and at four items. Looking just within the subitizing slope this pattern across the first three items was consistent with the interpretation from Experiment 1 that the primary indicator of the onset of processing load is the amplitude decrease beginning around 200 ms post-stimulus with the N2 and extending through the P3b window. This supposition was strongly supported by both span-groups' difference wave analysis, which showed evidence at three items of the appearance of a negative going difference component beginning around 200 ms post-stimulus that extended through both the N2 and P3b windows.

These first two experiments identified that a robust ERP profile consistent with effortful, time consuming WM activity is generated across the subitizing slope during response contingent item enumeration. When a load of three items must be enumerated the relatively unitary P3b component evident at a load of one and two items devolves dramatically (in Experiment 1, less so in Experiment 2) into the LPC. A distinct N2 deflection becomes evident, creating a P2-N2 complex which is accompanied by increased latency, and continued reduced amplitude of the following P3b peak. This same morphology is evident at a load of four items. Additionally, P3b latency showed strong evidence of steady and continued increase right through the subitizing slope and through to a load of four items, thus suggesting some sort of additional, time consuming processing with each additional item. However, while both samples generated a non-linear subitizing slope, and had large significant increases in P3b latency at the step to three items, only in Experiment 1 was there also clear evidence of significantly longer context updating for two items versus one. As such, while the ERP profile suggests load related WM activity is occurring during subitizing, there is only limited support for the serial counting explanation of the subitizing slope over the list search explanation.

The predominant feature of the LPC morphology identified in these first experiments was the occurrence of the major negative deflection, the N2, and reduced amplitude in the following P3b window during subitizing a load of three items. The P3b latency increases seen concurrent with appearance of the distinct N2, and the timing stability of the preceding P2 peak, support the interpretation of the N2 as reflecting time consuming processes that precede and prolong context updating in WM. An obvious question was does the item level at which such load related processes become active vary as a function of RT derived subitizing span? In other words, does the N2 reflect a process that has a discrete relationship with subitizing span? A final important finding from these experiments was that for people who had RT derived subitizing spans of three, or of four items, the morphology of the LPC at three items included an obvious N2 component, with reduced amplitude in the P3b window. Therefore it is likely that whatever load related processes the N2 might reflect, these are distributed across these item levels independent of RT derived subitizing span, and so will be recruited within the subitizing slope no matter what an individual’s span. This finding allowed the following experiments, which addressed activity across the subitizing slope, to include participants of any span size.

Experiment 3 went on to investigate whether processes involved with pattern matching and association may underlie subitizing, as suggested by Mandler and Shebo (1982), and Logan and Zbrodoff (2003). If this is the case then the negative going activity seen at three items in the first two experiments might in fact be reflecting configural binding operations and associative processes and not serial quantification operations in WM, as the serial counting explanation would suggest. Indeed, configural binding operations and associative processes may also occur in WM, but the flat RT function over the subitizing range that is associated with enumerating canonical patterns (see section 1.2.7) suggests that as item level increases no additional, time consuming demands are made on WM resources. The experimental paradigm also allowed control for another possible determinant of this negative activity, increased perceptual load. This has been shown to reduce amplitude over the P3b window (Lorist, 1996). In the item enumeration task increasing the number of items to be enumerated also increases the perceptual load of a trial, independent of any additional processing demands that the requirement to engage in enumeration might introduce.

The primary aim of Experiment 3 was to assess whether processing arrays of three items in terms of their configuration also generated the predominantly anterior negativity in the N2 and P3b windows that had been generated during item enumeration. A second aim was to see whether any such negativity was also accompanied by prolonged context updating, reflected by P3b peak latency, which would indicate additional time consuming operations and thus serial processing. At a load of three items these results showed clear negative going activity within the N2 window that was evident when the items must be enumerated, but *not* evident when the items needed to be processed only in terms of their configuration. The effect of requiring quantification of the array elements was very strong. Additionally, P3b latencies at three items were longer for the enumeration condition than they were for the pattern condition, suggesting more task difficulty, and strengthening the relationship between increased task difficulty, emergence of the N2, and prolonged P3b latency. Thus, there was no support for the explanation that the negative going activity in the N2 window might reflect processes related to configural processing. However, the findings provided clear support for the idea that the negative going activity in the N2 window which manifests in the waveforms at three items reflects the operation of quantification operations that are associated with item enumeration while subitizing, as defined by RT, was still occurring.

The final two experiments directly addressed the question of the functional significance of this ERP profile as an indicator of WM activity. The logic underlying the approach taken in these experiments was straightforward. Subjects were required to enumerate arrays of simple items—the primary task—while sometimes performing secondary tasks that placed a concurrent load (of either verbal or spatial information) on WM resources. Because the memory load was already occupying WM resources when subitizing had to proceed, any differences in morphology between the two memory load conditions at a particular item level would be attributable to the functional effect of WM load. Conversely, any effects of memory load that emerge would suggest that WM resources are involved with enumerating the number of items at which the effect occurred.

In Experiment 4 the verbal memory load generated clear effects on the fronto-central LPC morphology when enumerating even a single item. Context updating was prolonged, and P3b amplitude reduced relative to enumerating an item with no concurrent memory load. There was also a distinct N2 component, of similar magnitude to that observed in early experiments when enumerating arrays of three items. P3b latency was also longer, and its amplitude less, when enumerating two item arrays under memory load. The same distinct N2 was also evident for the memory-load but not the no-load condition for these two item arrays. Mean amplitudes across the anterior scalp in the N2 window were consistently lower in the

Load condition, with moderate effect sizes. So, the component was clearly evident and there was definitive statistical support for the claim that large N2 deflections accompanied prolonged context updating and reduced P3b amplitude while subitizing under conditions of verbal memory load. Overall, the addition of a verbal information load in WM while subitizing one or two items generated changes in the ERP behaviour that were similar to those seen at a load of three items when enumeration occurs with no additional memory load, and which are indicative of load related, time consuming processing in WM.

In Experiment 5, when spatial memory load was imposed during enumeration, clear effects of memory load were evident on P3b latency with context updating prolonged substantially. However, while there were amplitude changes in both the P3b and N2 windows these were not in a pattern consistent with the fronto-central LPC morphology examined throughout this thesis and which has been so robust across the other four experiments. P3b amplitude increased with Load, and significant relative reductions in N2 amplitude only occurred at posterior sites. It was speculated that this pattern could reflect the operation of another N2 component, the N2c, which has a posterior distribution and reflects operations involved in visual WM relating to stimulus categorisation and individuation, and a mechanism relating the functional interaction of the N2c and P3b to WM activity was proposed. This model suggested that maintaining spatial stimuli while enumerating might recruit different processes, and hence brain regions than those required to enumerate while holding verbal information active. In any case, within the LPC model examined in this thesis the spatial paradigm exhibited Load related increases in context updating time but no other indicators of WM activity during subitizing.

8.1.2 Functional Significance of the LPC

Before moving on to discuss the implications of these findings for understanding subitizing it is necessary to clarify what the data suggest about the functional significance of the LPC components across the subitizing slope. P3b latency (a) increased as a function of increasing item load (most notably at the step from two to three items); (b) was longer when having to enumerate rather than pattern match; and (c) increased as a function of memory load, when only a single item had to be enumerated. These load related latency increases were in each case accompanied by P3b amplitude reductions. Overall, the fact that the P3b was sensitive to changes in item load *and* to changes in concurrent memory load supports the functional significance argued for in Chapter 2, where it was suggested that P3b latency and amplitude changes reflected time course and processing demands of load related operations in WM.

A key feature of the argument put forward in Chapter 2 relating the LPC to WM activity was the idea of *devolution* of the relatively unitary P3b under conditions of WM load. In this research program the latency changes seen in the P3b were accompanied by clear devolution of the P3b into the LPC, with the appearance of the distinct N2 deflection once three items needed to be subitized. There was some evidence that timing of the N2 peak is associated with prolonged context updating during subitizing (i.e., increased as a function of increasing item load), but the most robust feature of the N2 presentation was reliable, load related amplitude reduction when subitizing a load of three items. This same behaviour (without significant statistical support) occurred when only one item had to be subitized when there was a concurrent memory load (whether verbal or spatial) already drawing on WM resources. Appearance of the N2 component being conditional on either item load or

WM load leads to the obvious inference that the N2 component reflects some WM related process(es) that respond to resource demands, and that these resources are recruited as item load increases during subitizing.

The final component in the devolved LPC is the P2. In this research program neither P2 latency nor amplitude changed as a function of item load, or of memory load. The P2 peak is defined when the N2 pulls the waveform in a negative direction under conditions of either item, or memory load, creating what has been referred to as the P2-N2 complex. The waveform comparison finishing Chapter 7 showed quite clearly that the timing characteristics of the P2 morphology are very regular even when different types of information are loading WM, or when item load is increased from two to three items. Thus, in line with arguments put forward in section 2.1.5.3, the functional significance of the P2 in processing episodes reflected by the unfolding of the LPC may be as part of a time constrained input system to feed representations forward for further integrative processing.

In summary, in this research program there is clear evidence of devolution of the LPC under conditions of increasing item load, and memory load. There is convincing evidence suggesting that the increases in P3b latency and decreases in P3b amplitude that accompany this devolution are reflecting increases in demands for WM resources. At the same time, there is evidence suggesting that the N2 component that defines the morphology of the LPC devolution is also reflecting load related processes related to WM activity. The P2 component shows no variation as a function of any demands made by increasing either item load or memory load.

8.1.3 Conclusions

Accepting these interpretations for the functional significance of the N2 and the P3b in the context of item enumeration, the results of this research program are clear. Relying just on behaviour of the P3b to begin with, revisiting the speculations from Experiments 1 and 2 we can conclude that time consuming WM activity is certainly involved with subitizing three items, and this activity appears to be beyond what would be required solely to search an ordered list of number names. This conclusion provides some explanation for the non-linear subitizing slope, and some support for the serial counting explanation of the slope over the list search explanation. Revisiting Experiment 4 there is the further suggestion that WM resources are also critical in subitizing even a single item. This conclusion is strengthened by reference to the behaviour of the N2 component in both Experiments 4 and 5. This clear relationship between behaviour of the LPC which reflects load related WM activity, and subitizing, also supports the core speculation underlying the logic of this research program—that is, that subitizing limits might be related to limits in WM capacity to engage in moment by moment maintenance and binding activities, a concept that has been referred to throughout this thesis as chunking capacity.

These conclusions have substantial implications in terms of models of subitizing during item enumeration. Simply finding that WM *is* involved in enumerating small collections of items rules out explanations that appeal to modular facilitation (e.g., Butterworth, 1999; Gallistel & Gelman, 2000), at least when accuracy and not estimation is stressed and displays are response contingent. Explanations appealing to perceptual mechanisms as the determining bottleneck can also be called into question by these findings. For example, there may well be only three or four FINSTs, as Pylyshyn (1989) suggests, thus giving the visual system an upper capacity of three or four items that can be indexed without necessity for ‘emptying’ and reassignment. However, the indication that WM resources are active while subitizing even one item, and the reliable indications of effortful WM processing while subitizing three items suggests that it might be at the stage of quantification in WM, not perceptual indexing, that processing constraints manifest. Similarly, the results of the pattern matching paradigm from Experiment 3 gave direct evidence against these sorts of processes underlying subitizing. While canonical or overlearnt patterns *do* influence characteristics of subitizing such as the removal of a subitizing slope (e.g., Wender and Rothkegel, 2000), under conditions where items appear in random configurations the current evidence suggests that WM resources subserving quantification processes, not pattern matching or associative processes, are recruited during successful subitizing. This position is consistent with the serial counting explanation of subitizing.

This raises the question of what sort of WM resources are recruited to achieve subitizing. Using secondary tasks that imposed either verbal or spatial information load on WM addresses the possible fractionation of WM (discussed in section 1.1.1.1) in terms of verbal and spatial domains of processing. Presenting domain specific stimuli in the secondary tasks therefore allowed some speculation regarding what types of WM activity might be recruited to achieve subitizing. For example, if subvocalisation is involved in subitizing but shifts in spatial attention aren't, subitizing while holding a verbal but not spatial information load in WM should lead to interference effects. The converse would hold should spatial attention but not subvocalisation be critical for subitizing to proceed. In the data from Experiments 4 and 5, however, there were strong indications of interference effects (in terms of the distinct N2) at all item levels (predominant for one and two item arrays) *for both verbal and spatial WM load.* An immediate interpretation of this might be that both verbal and spatial WM resources are recruited to subitize even a single item. However, assuming that the complexity of both types of memory stimuli was sufficient to tax WM (the interference effects suggest that this was the case), another explanation could be that even though verbal (for example) WM resources are not actually *directly* involved in subitizing, the verbal task taxes enough *shared activation* to disrupt spatial processing (which for argument's sake let us say *is* involved). This position would be consistent with the shared resource model put forward by the activation model (of course, the example could hold equally as well for spatial information). Thus, these findings do not support the inference that both verbal and spatial WM resources are recruited during subitizing. They do however provide support for the core feature of the activation model of WM referred to throughout this thesis—that is, that there is a fundamental resource of activation which WM operations draw on to accomplish maintenance and processing activities.

In conclusion, this research program has provided, to the author's knowledge, the first comprehensive examination of subitizing during response contingent item enumeration using ERPs as the dependent variables. The research design allowed inferences to be made regarding the functional significance of the N2 and P3b ERP components of the LPC as indicators of WM activity during subitizing. The behaviour of both the N2 and P3b throughout these five experiments has been consistent with the conclusion that, at least in the response contingent paradigm used in this program, WM resources are utilised during subitizing, supporting the serial counting explanation, and thus single-process views of subitizing.

REFERENCES

Akyürek, E. G., Leszczynski, M., & Schubo, A. (2010). The temporal locus of the interaction between working memory consolidation and the attentional blink. *Psychophysiology, 47*(6), 1134-1141.

Allen, R., & McGeorge, P. (2008). Enumeration: Shape information and expertise. *Acta Psychologica, 129*(1), 26-31.

Ashley, V., Vuilleumier, P., & Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *Neuroreport, 15*(1), 211-216.

Atchley, P., Jones, S. E., & Hoffman, L. (2003). Visual marking: A convergence of goal- and stimulus-driven processes during visual search. *Perception & Psychophysics, 65*(5), 667-667.

Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4: A new look at visual numerosity judgements. *Perception, 5,* 327-334.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*(2), 82-90.

Azizian, A., Freitas, A. L., Parvaz, M. A., & Squires, N. K. (2006). Beware misleading cues: perceptual similarity modulates the N2/P3 complex. *Psychophysiology, 43*(3), 253-260.

Baddeley, A. D. (1986). *Working memory*. New York, NY, US: Clarendon Press/Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417-423.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). New York, NY, USA: Cambridge University Press.

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception & Psychophysics, 50*(6), 555-564.

Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology, 42*(1), 33-42.

Basak, C., & Verhaeghen, P. (2003). Subitizing speed, subitizing range, counting speed, the Stroop effect, and aging: Capacity differences and speed equivalence. *Psychology & Aging, 18*(2), 240-249.

Bastiaansen, M. C. M., & Brunia, C. H. M. (2001). Anticipatory attention: An event-related desynchronization approach. *International Journal of Psychophysiology, 43*(1), 91-107.

Bentin, S., & McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 130-149.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav Brain Sci, 30*(5-6), 481-499; discussion 499-548.

Bodner, G. E., & Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory and Language, 37*(2), 268-293.

Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clin Neurophysiol, 112*(12), 2224-2232.

Boles, D. B., Phillips, J. B., & Givens, S. M. (2007). What dot clusters and bar graphs reveal: Subitizing is fast counting and subtraction. *Perception and Psychophysics, 69*(6), 913-922.

Bowles, R. P., & Salthouse, T. A. (2003). Assessing the age-related effects of proactive interference on working memory tasks using the Rasch model. *Psychology and Aging, 18*(3), 608-615.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review, 64*, 205-215.

Broadbent, D. E. (1958). *Perception and communication*: London, Pergamon Press. (1958). v, 338 pp.

Brookhuis, K. A., Mulder, G., Mulder, L. J., & Gloerich, A. B. (1983). The P3 complex as an index of information processing: the effects of response probability. *Biological Psychology, 17*(4), 277-296.

Brookhuis, K. A., Mulder, G., Mulder, L. J., Gloerich, A. B., van Dellen, H. J., van der Meere, J. J., et al. (1981). Late positive components and stimulus evaluation time. *Biol Psychol, 13,* 107-123.

Bruder, G., Kayser, J., Tenke, C., Rabinowicz, E., Friedman, M., Amador, X., et al. (1998). The time course of visuospatial processing deficits in schizophrenia: An event-related brain potential study. *Journal of Abnormal Psychology, 107*(3), 399-411.

Bruin, K. J., Wijers, A. A., & van Staveren, A. S. (2001). Response priming in a go/nogo task: do we have to explain the go/nogo N2 effect in terms of response activation instead of inhibition? *Clin Neurophysiol, 112*(9), 1660-1671.

Brumaghim, J. T., Elorman, R., Strauss, J., Lewine, J. D., & Goldstein, M. G. (1987). Does Methylphenidate Affect Information Processing? Findings From Two Studies on Performance and P3b Latency. *Psychophysiology, 24*(3), 361-373.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision, 10*(6), 1-10.

Butterworth, B. (1999). *What counts - How every brain is hardwired for math*. New York: The Free Press.

Cadwell, J. A., & Villarreal, R. A. (1999). Electrophysiologic equipment and electrical safety. In M. J. Aminoff (Ed.), *Electrodiagnosis in clinical neurology* (4th ed ed., pp. pp. 15–33). New York: Churchill Livingstone.

Cameron, A. M., Geffen, G. M., Kavanagh, D. J., Wright, M. J., McGrath, J. J., & Geffen, L. B. (2003). Event-related potential correlates of impaired visuospatial working memory in schizophrenia. *Psychophysiology, 40*(5), 702-715.

Camos, V., & Tillmann, B. (2008). Discontinuity in the enumeration of sequentially presented auditory and visual stimuli. *Cognition, 107*(3), 1135-1143.

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional Imaging of Numerical Processing in Adults and 4-y-Old Children. [doi:10.1371/journal.pbio.0040125]. *PLoS Biol, 4*(5), e125 EP -.

Carretti, B., Mammarella, I., & Borella, E. (2011). Age differences in proactive interference in verbal and visuospatial working memory. *Journal of Cognitive Psychology*, 1-13.

Chapman, R. M., & Bragdon, H. R. (1964). Evoked responses to numerical and non-numerical visual stimuli while problem solving. *Nature, 203*, 1155-1157.

Chi, M. T., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology, 19*(3), 434-439.

Christensen, C. A., Ivkovich, D., & Drake, K. J. (2001). Late positive ERP peaks observed in stimulus-response compatibility tasks tested under speed-accuracy instructions. *Psychophysiology, 38*(3), 404-416.

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist, Vol 49*(12), 997-1003.

Coles, M. G. H., & Rugg, M. D. (1995). Event-related brain potentials: An introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford: Oxford University Press.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*(4), 354-373.

Corbetta, M., Shulman, G. L., Miezin, F. M., & Petersen, S. E. (1995). Superior Parietal Cortex Activation During Spatial Attention Shifts and Visual Feature Conjunction. *Science, 270*(5237), 802-805.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*(1), 21-58.

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). New York, NY, USA: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences, 24*(1), 87-185.

Daffner, K. R., Mesulam, M. M., Scinto, L. F. M., Calvo, V., Faust, R., & Holcomb, P. J. (2000). An electrophysiological index of stimulus unfamiliarity. *Psychophysiology, 37*(6), 737-747.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466.

Danos, P., Guich, S., Abel, L., & Buchsbaum, M. S. (2001). EEG alpha rhythm and glucose metabolic rate in the thalamus in schizophrenia. *Neuropsychobiology, 43*(4), 265-272.

Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & et al. (1990). Approach/withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of Personality and Social Psychology, 58*(2), 330-341.

Davidson, R. J., Jackson, D. C., & Larson, C. L. (2000). Human electroencephalography. New York, NY: Cambridge University Press.

Deacon, D., Breton, F., Ritter, W., & Vaughan Jr., H. G. (1991). The Relationship Between N2 and N400: Scalp Distribution, Stimulus Probability, and Task Relevance. *Psychophysiology, 28*(2), 185-200.

Dean, R. S., Garabedian, A. A., & Yekovich, F. R. (1983). The effect of modality shifts on proactive interference in long-term memory. *Contemporary Educational Psychology, 8*(1), 28-45.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

Dehaene, S., & Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultanagnosis patients. *Journal of Experimental Psychology: Human Perception & Performance, 20*(5), 958-975.

Deldin, P. J., Deveney, C. M., Kim, A. S., Casas, B. R., & Best, J. L. (2001). A Slow Wave Investigation of Working Memory Biases in Mood Disorders. *Journal of Abnormal Psychology, 110*(2), 267-281.

Demeyere, N., Lestou, V., & Humphreys, G. W. (2010). Neuropsychological evidence for a dissociation in counting and subitizing. *Neurocase, 16*(3), 219-237.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*, 80-90.

Deveney, C. M., & Deldin, P. J. (2004). Memory of Faces: A Slow Wave ERP Study of Major Depression. *Emotion, 4*(3), 295-304.

Deyzac, E., Logie, R. H., & Denis, M. (2006). Visuospatial working memory and the processing of spatial descriptions. *British Journal of Psychology, 97*(2), 217-243.

Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: Mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology, 41*(5), 665-678.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*(3), 357-427.

Donchin, E., & Lindsley, D. B. (1965). Average evoked potentials and reaction times to visual stimuli. *Electroencephalography and clinical neurophysiology*, 217-223.

Donchin, E., Wicke, J. D., & Lindsley, D. B. (1963). Cortical evoked potentials and perception of paired flashes. *Science, 141*, 1285-1286.

Dümpelmann, M., Ball, T., & Schulze-Bonhage, A. (2011). sLORETA allows reliable distributed source reconstruction based on subdural strip and grid recordings. [10.1002/hbm.21276]. *Human Brain Mapping*, n/a.

Effern, A., Lehnertz, K., FernÃ¡ndez, G., Grunwald, T., David, P., & Elger, C. E. (2000). Single trial analysis of event related potentials: Non-linear de-noising with wavelets. *Clinical Neurophysiology, 111*(12), 2255-2263.

Egeth, H. E., Leonard, C. J., & Palomares, M. (2008). The role of attention in subitizing: Is the magical number 1? *Visual Cognition, 16*(4), 463-473.

Eimer, M., & Schroeger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology, 35*(3), 313-327.

Emery, L., Hale, S., & Myerson, J. (2008). Age differences in proactive interference, working memory, and abstract reasoning. *Psychology and Aging, 23*(3), 634-645.

Engle, R. W. (2001). What is working memory capacity? In H. L. Roediger, III & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder. Science conference series* (pp. 297-314). Washington, DC: American Psychological Association.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). New York, NY, USA: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297). New York, NY, USA: Cambridge University Press.

Fabiani, M., Gratton, G., & Coles, M. G. H. (2000). Event-related brain potentials. New York, NY: Cambridge University Press.

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. In P. K. Ackles, J. R. Jennings & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 2, pp. 1-78). Greenwich, Connecticut: Jai Press.

Falkenstein, M., Hohnsbein, J., & Hoormann, J. r. (1994). Time pressure effects on late components of the event-related potential (ERP). *Journal of Psychophysiology, 8*(1), 22-30.

Falkenstein, M., Hoormann, J. r., Hohnsbein, J., & Kleinsorge, T. (2003). Short-term mobilization of processing resources is revealed in the event-related potential. *Psychophysiology, 40*(6), 914-923.

Ferrand, L., Grainger, J., & Segui, J. (1994). A study of masked form priming in picture and word naming. *Memory and Cognition, 22*(4), 431-441.

Ferree, T. C., Luu, P., Russell, G. S., & Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clin Neurophysiol, 112*(3), 536-544.

Fleming, K., Goldberg, T. E., Blinks, S., Randolph, C., Gold, J. M., & Weinberger, D. R. (1997). Visuospatial working memory deficits in patients with schizophrenia. *Biological Psychiatry*(41), 43-49.

Fodor, J. (1986). The modularity of mind. In Z. W. Pylyshyn & W. Demopoulos (Eds.), *Meaning and cognitive structure: Issues in the computational theory of mind. Theoretical issues in cognitive science.* (pp. 3-18). Westport, CT: Ablex Publishing.

Folk, C. L., Egeth, H. E., & Kwak, H.-w. (1988). Subitizing: Direct apprehension or serial processing? *Perception & Psychophysics, 44*(4), 313-320.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45*(1), 152-170.

Forster, K. I. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research, 27*(2), 203-233.

Fox, A. M., Michie, P. T., Wynne, C. D., & Maybery, M. T. (2000). ERP correlates of response inhibition to elemental and configural stimuli in a negative patterning task. *Clin Neurophysiol, 111*(6), 1045-1053.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews, 25*(4), 355-373.

Gallistel, C. R., & Gelman, R. (1991). Subitizing: The preverbal counting process. In W. Kessen, A. Ortony & F. I. M. Craik (Eds.), *Memories, thoughts, and emotions: Essays in honor of George Mandler.* (pp. 65-81). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4*(2), 59-59-65.

García-Larrea, L., & Cézanne-Bert, G. (1998). P3 positive slow wave and working memory load: A study on the functional correlates of slow wave activity. *Electroencephalography & Clinical Neurophysiology: Evoked Potentials, 108*(3), 260-273.

Geffen, G. M., Wright, M. J., Green, H. J., Gillespie, N. A., Smyth, D. C., Evans, D. M., et al. (1997). Effects of memory load and distraction on performance and event-related slow potentials in a visuospatial working memory task. *Journal of Cognitive Neuroscience, 9*(6), 743-757.

Gilhooly, K. J., Wynn, V., Philips, L. H., Llogie, R. H., & Della Sala, S. (2002). Visuo-spatial and verbal working memory in the five-disc Tower of London task: An individual differences approach. *Thinking & Reasoning, 8*(3), 165-178.

Gunter, T. C., Friederici, A. D., & Hahne, A. (1999). Brain responses during sentence reading: Visual input affects central processes. *Neuroreport: For Rapid Communication of Neuroscience Research, 10*(15), 3175-3178.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13*(3), 339-356.

Hale, S., Bronik, M. D., & Fry, A. F. (1997). Verbal and Spatial Working Memory in School-Age Children: Developmental Differences in Susceptibility to Interference. [doi:]. *Developmental Psychology, 33*(2), 364-371.

Han, S. (2000). Hierarchical processing and level-repetition effect as indexed by early brain potentials. *Psychophysiology, 37*(6), 817-830.

Hitch, G. J., & Baddeley, A. D. (1976). Verbal reasoning and working memory. *The Quarterly Journal of Experimental Psychology, 28*(4), 603-621.

Hoen, M., & Dominey, P. F. (2000). ERP analysis of cognitive sequencing: A left anterior negativity related to structural transformation processing. *Neuroreport: For Rapid Communication of Neuroscience Research, 11*(14), 3187-3191.

Hoffman, J. E., Simons, R. F., & Houck, M. R. (1983). Event-Related Potentials During Controlled and Automatic Target Detection. *Psychophysiology, 20*(6), 625-632.

Hopkins, R. H. (1973). Presentation modality and proactive interference in short-term retention using a mixed-modality distractor task. *Memory & Cognition, 1*(4), 439-442.

Hopkins, R. H., Edwards, R. E., & Cook, C. L. (1973a). Presentation modality, distractor modality, and proactive interference in short-term memory. [doi: 10.1037/h0034379]. *Journal of Experimental Psychology, 98*(2), 362-367.

Hopkins, R. H., Edwards, R. E., & Cook, C. L. (1973b). Presentation modality, distractor modality, and proactive interference in short-term memory. *Journal of Experimental Psychology, 98*(2), 362-367.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and Tracking Difficulty: Evidence For Multiple Resources in Dual-Task Performance. *Psychophysiology, 17*(3), 259-273.

Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots.

Jerger, K., Biggins, C., & Fein, G. (1992). P50 suppression is not affected by attentional manipulations. *Biological Psychiatry, 31*(4), 365-377.

Jevons, W. S. (1871). The power of numerical discrimination. *Nature, 3*, 281-282.

John, E. R., Easton, P., & Isenhart, R. (1997). Consciousness and cognition may be mediated by multiple independent coherent ensembles. *Consciousness and Cognition: An International Journal, 6*(1), 3-39.

Johnson, R., Jr. (1986). A Triarchic Model of P300 amplitude. *Psychophysiology, 23*(4), 367-384.

Johnson, R., Jr. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. In P. K. Ackles, J. R. Jennings & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 3, pp. 69-137). Greenwich, Connecticut: Jai Press.

Jonkman, L., Lansbergen, M., & Stauder, J. (2003). Developmental differences in behavioral and event-related brain responses associated with response preparation and inhibition in a go/nogo task. *Psychophysiology, 40*(5), 752-761.

Kafer, K. L., & Hunter, M. (1997). On testing the face validity of planning/problem-solving tasks in a normal population. *Journal of the International Neuropsychological Society, 3*(02), 108-119.

Kamijo, K., & Takeda, Y. (2009). Subitizing requires more attentional resources in older adults. *Japanese Journal of Physiological Psychology and Psychophysiology, 27*(3), 199-206.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(2), 336-358.

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology, 47*(5), 888-904.

Karayanidis, F., & Michie, P. T. (1996). Frontal processing negativity in a visual selective attention task. *Electroencephalography & Clinical Neurophysiology, 99*(1), 38-56.

Karayanidis, F., & Michie, P. T. (1997). Evidence of visual processing negativity with attention to orientation and color in central space. *Electroencephalography & Clinical Neurophysiology, 103*(2), 282-297.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology, 62*(0002-9556, 0002-9556), 498-498-525.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*(5), 580-602.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7*(3), 376-395.

Klahr, D. (1973a). A production system for counting, subitizing and adding. Oxford, England: Academic.

Klahr, D. (1973b). Quantification processes. In W. D. Chase (Ed.), *Visual information processing* (pp. 3-34). Oxford, England: Academic.

Klahr, D., & Wallace, J. G. (1973). The role of quantification operators in the development of conservation of quantity.

Klimesch, W. (1997). EEG-alpha rhythms and memory processes. *International Journal of Psychophysiology, 26*(1-3), 319-340.

Kok, A. (1986). Effects of degradation of visual stimulation on components of the event-related potential (ERP) in go/nogo reaction tasks. *Biol Psychol, 23*(1), 21-38.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38*(3), 557-577.

Kok, A., Ramautar, J. R., De Ruiter, M. B., Band, G. P. H., & Ridderinkhof, K. R. (2004). ERP components associated with successful and unsuccessful stopping in a stop-signal task. *Psychophysiology, 41*(1), 9-20.

Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology, 33*(3), 282-294.

Kramer, A., Schneider, W., Fisk, A., & Donchin, E. (1986). The effects of practice and task structure on components of the event-related brain potential. *Psychophysiology, 23*(1), 33-47.

Kramer, A., Strayer, D. L., & Buckley, J. (1991). Task Versus Component Consistency in the Development of Automatic Processing: A Psychophysiological Assessment. *Psychophysiology, 28*(4), 425-437.

Krause, C. M. (2003). Brain electric oscillations and cognitive processes.

Kroll, N. E., Bee, J., & Gurski, G. (1973). Release of proactive interference as a result of changing presentation modality. [doi: 10.1037/h0034293]. *Journal of Experimental Psychology, 98*(1), 131-137.

Kucian, K., Loenneker, T., Dietrich, T., Dosch, M., Martin, E., & von Aster, M. (2006). Impaired neural networks for approximate calculation in dyscalculic children: a functional MRI study. *Behavioral and Brain Functions, 2*, 31-31-31.

Kunar, M. A., Humphreys, G. W., Smith, K. J., & Hulleman, J. (2003). What is "marked" in visual marking? Evidence for effects of configuration in preview search. *Perception & Psychophysics, 65*(6), 982-982-996.

Kutas, M., & Federmeier, K. D. (2009). N400. *Scholarpedia, 4*(10).

Kutas, M., & Iragui, V. (1998). The N400 in a semantic categorization task across 6 decades. *Electroencephalography & Clinical Neurophysiology: Evoked Potentials, 108*(5), 456-471.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science, 197*(4305), 792-795.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence, 14*(4), 389-433.

Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences. Special Issue: Frontiers of learning and individual differences: Paradigms in transition, 2*(2), 129-160.

Larson, C. L., Davidson, R. J., Abercrombie, H. C., Ward, R. T., Schaefer, S. M., Jackson, D. C., et al. (1998). Relations between PET-derived measures of thalamic glucose metabolism and EEG alpha power. *Psychophysiology, 35*(2), 162-169.

Laskaris, N. A., & Ioannides, A. A. (2002). Semantic geodesic maps: A unifying geometrical approach for studying the structure and dynamics of single trial evoked responses. *Clinical Neurophysiology, 113*(8), 1209-1226.

Logan, G., & Zbrodoff, N. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review, 10*(3), 676-682.

Logie, R. H., & Baddeley, A. D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(2), 310-326.

Lorist, M. M., Snel, J., Kok, A., & Mulder, G. (1996). Acute effects of caffeine on selective attention and visual search processes. *Psychophysiology, 33*(4), 354-361.

Low, K. A., Leaver, E. E., Kramer, A. F., Fabiani, M., & Gratton, G. (2009). Share or compete? Load-dependent recruitment of prefrontal cortex during dual-task performance. *Psychophysiology, 46*(5), 1069-1079.

Luck, S. J. (1995a). Multiple mechanisms of visual-spatial attention: Recent evidence from human electrophysiology. *Behavioural Brain Research, 71*(1-2), 113-123.

Luck, S. J. (1995b). Multiple mechanisms of visual-spatial attention: Recent evidence from human electrophysiology. *Behavioural Brain Research, 71*(1-2), 113-123.

Luck, S. J., & Hillyard, S. A. (1990). Electrophysiological evidence for parallel and serial processing during visual search. *Percept Psychophys, 48*(6), 603-617.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31*(3), 291-308.

Luck, S. J., & Hillyard, S. A. (1995). The role of attention in feature detection and conjunction discrimination: An electrophysiological analysis. *International Journal of Neuroscience. Special M. Russell Harter memorial issue: Progress in visual information processing, 80*(1-4), 281-297.

Luck, S. J., Kappenman, E. S., Fuller, R. L., Robinson, B., Summerfelt, A., & Gold, J. M. (2009). Impaired response selection in schizophrenia: evidence from the P3 wave and the lateralized readiness potential. *Psychophysiology, 46*(4), 776-786.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General, 130*(2), 199-207.

Magliero, A., Bashore, T. R., Coles, M. G. H., & Donchin, E. (1984). On the Dependence of P300 Latency on Stimulus Evaluation Processes. *Psychophysiology, 21*(2), 171-186.

Makeig, S., Westerfield, M., Jung, T.-P., Covington, J., Townsend, J., Sejnowski, T. J., et al. (1999). Functionally independent components of the late positive event-related potential during visual spatial attention. *Journal of Neuroscience, 19*(7), 2665-2680.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111*(1), 1-22.

Martinez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., et al. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neuroscience, 2*(4), 364-369.

McCarthy, G., & Donchin, E. (1981). A metric for thought: A comparison of P300 latency and reaction time. *Science, 211*(4477), 77-80.

McGrath, J., Chapple, B., & Wright, M. (2001). Working memory in schizophrenia and mania: Correlation with symptoms during the acute and subacute phases. *Acta Psychiatria Scandinavica*(103), 181-188.

Mecklinger, A., Kramer, A. F., & Strayer, D. L. (1992). Event related potentials and EEG components in a semantic memory search task. *Psychophysiology, 29*(1), 104-119.

Mecklinger, A., Kramer, A. F., & Strayer, D. L. (1992). Event related potentials and EEG components in a semantic memory search task. *Psychophysiology, 29*(1), 104-119.

Mecklinger, A., & Müller, N. (1996). Dissociations in the processing of "what" and "where" information in working memory: An event-related potential analysis. *Journal of Cognitive Neuroscience, 8*(5), 453-473.

Mecklinger, A., & Pfeifer, E. (1996). Event-related potentials reveal topographical and temporal distinct neuronal activation patterns for spatial and object working memory. *Cognitive Brain Research, 4*(3), 211-224.

Michalski, A. (1998). Novel environment as a stress-inducing factor. An event-related potentials study. *Acta Neurobiol Exp (Wars), 58*(3), 199-205.

Missonnier, P., Leonards, U., Gold, G., Palix, J., Ibanez, V., & Giannakopoulos, P. (2003). A new electrophysiological index for working memory load in humans. *Neuroreport: For Rapid Communication of Neuroscience Research, 14*(11), 1451-1455.

Miyake, A., & Shah, P. (1999a). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442-481). New York, NY, USA: Cambridge University Press.

Miyake, A., & Shah, P. (Eds.). (1999b). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY, USA: Cambridge University Press.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral & Brain Sciences, 13*(2), 201-288.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*(1), 1-21.

Nan, Y., Knösche, T. R., & Luo, Y.-J. (2006). Counting in everyday life: Discrimination and enumeration. *Neuropsychologia, 44*(7), 1103-1113.

Nittono, H., Shibuya, Y., & Hori, T. (2007). Anterior N2 predicts subsequent viewing time and interest rating for novel drawings. *Psychophysiology, 44*(5), 687-696.

Olivers, C. N. L., & Watson, D. G. (2003). Visual marking inhibits singleton capture. *Cognitive Psychology, 47*(1), 1-42.

Olivers, C. N. L., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition, 16*(4), 439-462.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin, 116*(2), 220-244.

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: selected studies of the Event-Related Potential. *Int J Med Sci, 2*(4), 147-154.

Pesciarelli, F., Kutas, M., Dell'acqua, R., Peressotti, F., Job, R., & Urbach, T. P. (2007). Semantic and repetition priming within the attentional blink: an event-related brain potential (ERP) investigation study. *Biol Psychol, 76*(1-2), 21-30.

Pfefferbaum, A., Christensen, C., Ford, J. M., & Kopell, B. S. (1986). Apparent response incompatibility effects on P3 latency depend on the task. *Electroencephalography & Clinical Neurophysiology, 64*(5), 424-437.

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*.

Piazza, M., Mechelli, A., Butterworth, B., & Price, C. J. (2002). Are Subitizing and Counting Implemented as Separate or Functionally Overlapping Processes? *NeuroImage, 15*, 435-446.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*(2), 127-152.

Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. L. (2009). One, two, three, many - Subitizing in active touch. *Acta Psychologica, 131*(2), 163-170.

Plaisier, M. A., & Smeets, J. B. J. (2011). Haptic subitizing across the fingers. *Attention, Perception, and Psychophysics, 73*(5), 1579-1585.

Plaisier, M. A., Tiest, W. M. B., & Kappers, A. M. L. (2010). Grabbing subitizing with both hands: Bimanual number processing. *Experimental Brain Research, 202*(2), 507-512.

Poiese, P., Spalek, T. M., & Di Lollo, V. (2008). Attentional involvement in subitizing: Questioning the preattentive hypothesis. *Visual Cognition, 16*(4), 474-485.

Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain & Cognition, 56*(1), 5-13.

Potts, G. F., & Tucker, D. M. (2001). Frontal evaluation and posterior representation in target detection. *Cognitive Brain Research, 11*(1), 147-156.

Pratt, N., Willoughby, A., & Swick, D. (2011). Effects of working memory load on visual selective attention: behavioral and electrophysiological evidence. *Frontiers of Human Neuroscience, 5*, 57.

Pritchard, W. S., Shappell, S. A., & Brandt, M. E. (1991). Psychophysiology of the N200/N400: A review and classification scheme. In J. R. Jennings & P. K. Ackles (Eds.)*, Advances in psychophysiology: A research annual*. London: Jessica Kingsley.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*(1), 65-97.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*(1-2), 127-158.

Ragot, R. (1984). Perceptual and motor space representation: An event-related potential study. *Psychophysiology, 21*(2), 159-170.

Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition, 107*(1), 82-104.

Ramautar, J. R., Kok, A., & Ridderinkhof, K. R. (2004). Effects of stop-signal probability in the stop-signal paradigm: the N2/P3 complex further validated. *Brain and Cognition, 56*(2), 234-252.

Repp, B. H. (2007). Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts. *Perception and Psychophysics, 69*(4), 529-543.

Ritter, W., Simson, R., & Vaughan, H. G. (1983). Event-related potential correlates of two stages of information processing in physical and semantic discrimination trials. *Psychophysiology, 20*(2), 168-179.

Ritter, W., Simson, R., Vaughan, H. G., Jr., & Friedman, D. (1979). A brain event related to the making of a sensory discrimination. *Science, 203*(4387), 1358-1361.

Ritter, W., Simson, R., Vaughan, H. G., & Macht, M. (1982). Manipulation of event-related potential manifestations of information processing stages. *Science, 218*(4575), 909-911.

Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., & Münte, T. F. (2002). Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia, 40*(7), 778-787.

Rösler, F., & Heil, M. (1991). Toward a functional categorization of slow waves: Taking into account past and future events. *Psychophysiology, 28*(3), 344-358.

Rösler, F., Heil, M., Bajric, J., Pauls, A. C., & Hennighausen, E. (1995). Patterns of cerebral activation while mental images are rotated and changed in size. *Psychophysiology, 32*(2), 135-149.

Rösler, F., Heil, M., & Glowalla, U. (1993). Monitoring retrieval from long-term memory by slow event-related brain potentials. *Psychophysiology, 30*(2), 170-182.

Ruchkin, D. S., Canoune, H. L., Johnson, R., & Ritter, W. (1995). Working memory and preparation elicit different patterns of slow wave event-related brain potentials. *Psychophysiology, 32*(4), 399-410.

Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences, 26*(6), 709-777.

Rugg, M. D., & Coles, M. G. H. (1995a). *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford: Oxford University Press.

Rugg, M. D., & Coles, M. G. H. (1995b). The ERP and cognitive psychology: Conceptual issues. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford: Oxford University Press.

Sagi, D., & Julesz, B. (1985a). Detection versus discrimination of visual orientation. *Perception, 14*(5), 619-628.

Sagi, D., & Julesz, B. (1985b). "Where" and "what" in vision. *Science (New York, N.Y.), 228*(4704), 1217-1217-1219.

Salisbury, D. F. (2004). Semantic memory and verbal working memory correlates of N400 to subordinate homographs. *Brain & Cognition, 55*(2), 396-399.

Saltzman, I. J., & Garner, W. R. (1948). Reaction time as a measure of span of attention. *Journal of Psychology: Interdisciplinary and Applied, 25*(0022-3980, 0022-3980), 227-227-241.

Sathian, K., Simon, T. J., Peterson, S., Patel, G. A., Hoffman, J. M., & Grafton, S. T. (1999). Neural evidence linking visual object enumeration and attention. *Journal of Cognitive Neuroscience, 11*(1), 36-51.

Schendan, H. E., & Kutas, M. (2002). Neurophysiological evidence for two processing times for visual object identification. *Neuropsychologia, 40*(7), 931-945.

Schendan, H. E., & Kutas, M. (2003). Time course of processes and representations supporting visual object identification and memory. *Journal of Cognitive Neuroscience, 15*(1), 111-135.

Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology, 37*(4), 473-484.

Schmitt, B. M., Rodriguez-Fornells, A., Kutas, M., & Münte, T. F. (2001). Electrophysiological estimates of semantic and syntactic information access during tacit picture naming and listening to words. *Neuroscience Research, 41*(3), 293-298.

Schneider, W., Eschman, A., & Zuccolotto, A. (2001). E-Prime (Version 1.2). Pittsburgh: Psychology Software Tools, Inc.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition, 80*(1-2), 159-177.

Schubotz, R., & Friederici, A. D. (1997). Electrophysiological correlates of temporal and spatial information processing. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 8*(8), 1981-1986.

Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology, 54*(1), 1-14.

Shah, P., & Miyake, A. (1999). Models of working memory: An introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 1-27). New York, NY, USA: Cambridge University Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190.

Shin, J., Talnov, A., Matsumoto, G., & Brankack, J. (2001). Hippocampal theta rhythm and running speed: A reconsideration using within-single trial analysis. *Neurocomputing: An International Journal, 38-40*, 1567-1574.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex, 6*(1), 11-20.

Smith, J. L., Smith, E. A., Provost, A. L., & Heathcote, A. (2009). Sequence effects support the conflict theory of N2 and P3 in the Go/NoGo task. *Int J Psychophysiol, 75*(3), 217-226.

Smolnik, R. d., Perras, B., Mölle, M., Fehm, H. L., & Born, J. (2000). Event-related brain potentials and working memory function in healthy humans after single-dose and prolonged intranasal administration of adrenocorticotropin 4-10 and desacetyl--melanocyte stimulating hormone. *Journal of Clinical Psychopharmacology, 20*(4), 445-454.

Spearman, C. (1923). The nature of "intelligence" and the principles of cognition.

Spitzer, M. (1993). The psychopathology, neuropsychology, and neurobiology of associative and working memory in schizophrenia. *European Archives of Psychiatry & Clinical Neurosciences*(243), 57-70.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*(3736), 652-654.

Strayer, D. L., & Kramer, A. F. (1990). Attentional requirements of automatic and controlled processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(1), 67-82.

Svenson, O., & Sjöberg, K. (1983). Speeds of subitizing and counting processes in different age groups. *Journal of Genetic Psychology, 142*(2), 203-211.

Tillman, C. M., & Wiens, S. (2011). Behavioral and ERP indices of response conflict in Stroop and flanker tasks. *Psychophysiology, 48*(10), 1405-1411.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*(6984), 751-751-754.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 459-478.

Trick, L. M. (1990). Subitizing and counting: Preattentive and attentive processing in visual enumeration. *Dissertation Abstracts International, 51*(4-B), 2095.

Trick, L. M. (1992). A theory of enumeration that grows out of a general theory of vision: Subitizing, counting, and FINSTs. Oxford, England: North-Holland.

Trick, L. M. (2005). The role of working memory in spatial enumeration: Patterns of selective interference in subitizing and counting. *Psychonomic Bulletin & Review, 12*(4), 675-681.

Trick, L. M., Enns, J. T., & Brodeur, D. A. (1996). Life span changes in visual enumeration: The number discrimination task. *Developmental Psychology, 32*(5), 925-932.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 331-351.

Trick, L. M., & Pylyshyn, Z. W. (1994a). Cueing and counting: Does the position of the attentional focus affect enumeration? *Visual Cognition, 1*(1), 67-100.

Trick, L. M., & Pylyshyn, Z. W. (1994b). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*(1), 80-102.

Tuholski, S. W., Engle, R. W., & Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. *Memory and Cognition, 29*(3), 484-492.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*(2), 127-154.

Van der Lubbe, R. H. J., & Verleger, R. (2002). Aging and the Simon task. *Psychophysiology, 39*(1), 100-110.

Van Oeffelen, M. P., & Vos, P. G. (1982a). Configurational effects on the enumeration of dots: Counting by groups. *Memory & Cognition, 10*(4), 396-396-404.

Van Oeffelen, M. P., & Vos, P. G. (1982b). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics, 32*(2), 163-170.

Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology, 34*(2), 131-156.

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *PLoS ONE, 3*(9).

Vos, S. H., Gunter, T. C., Kolk, H. H. J., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology, 38*(1), 41-63.

Wagner, P., Röschke, J., Grözinger, M., & Mann, K. (2000). A replication study on P300 single trial analysis in schizophrenia: Confirmation of a reduced number of 'true positive' P300 waves. *Journal of Psychiatric Research, 34*(3), 255-259.

Warren, H. C. (1897). The Reaction Time of Counting. (Stud. fr. Princeton Psychol. Lab.). *Princeton Contrib. to Psychol., II*, 99-99-121.

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology, 38*(6), 998-1003.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. San Antonio: Psychological Corporation.

Wender, K. F., & Rothkegel, R. (2000). Subitizing and its subprocesses. *Psychological Research, 64*(2), 81-92.

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science, 221*(4615), 1080-1082.

Wijers, A. A., Mulder, G., Okita, T., & Mulder, L. J. M. (1989). Event-Related Potentials During Memory Search and Selective Attention to Letter Size and Conjunctions of Letter Size and Color. *Psychophysiology, 26*(5), 529-547.

Wijers, A. A., Mulder, G., Okita, T., Mulder, L. J. M., & Scheffers, M. K. (1989). Attention to Color: An Analysis of Selection, Controlled Search, and Motor Activation, Using Event-Related Potentials. *Psychophysiology, 26*(1), 89-109.

Wijers, A. A., Okita, T., Mulder, G., Mulder, L. J., Lorist, M. M., Poiesz, R., et al. (1987). Visual search and spatial attention: ERPs in focussed and divided attention conditions. *Biol Psychol, 25*(1), 33-60.

Wijers, A. A., Otten, L. J., Feenstra, S., Mulder, G., & Mulder, L. J. M. (1989). Brain Potentials During Selective Attention, Memory Search, and Mental Rotation. *Psychophysiology, 26*(4), 452-467.

Wolach, I., & Pratt, H. (2001). The mode of short-term memory encoding as indicated by event-related potentials in a memory scanning task with distractions. *Clinical Neurophysiology, 112*(1), 186-197.

Wolters, G., Van Kempen, H., & Wijlhuizen, G.-J. (1987). Quantification of small numbers of dots: Subitizing or pattern recognition? *American Journal of Psychology, 100*(2), 225-237.

Xu, X., & Liu, C. (2008). Can subitizing survive the attentional blink? An ERP study. *Neuroscience Letters, 440*(2), 140-144.