Cognitive Modeling of Interference From Prospective-Memory Tasks

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Sebastian Horn

aus Tegernsee

Düsseldorf, Oktober 2012

Aus dem Institut für Experimentelle Psychologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referentin: Prof. Ute J. Bayen, Ph.D.

Koreferent: Prof. Dr. Jochen Musch

Tag der mündlichen Prüfung: 30. 11. 2012

Inhaltsverzeichnis

Zusammenfassung

Prospektive Gedächtnisaufgaben erfordern, sich zu einem bestimmten Zeitpunkt in der Zukunft an eine intendierte Handlung zu erinnern (z. B. auf dem Heimweg noch einen Einkauf zu erledigen). Charakteristisch für solche im Alltag relevante Aufgaben ist der selbstinitiierte Abruf aus dem Gedächtnis ohne explizite Hinweise. Um prospektives Erinnern im Laborparadigma zu untersuchen, bearbeiten Versuchspersonen eine fortlaufende Aufgabe (z. B. eine lexikalische Entscheidungsaufgabe). In diese Aufgabe werden vorher definierte prospektive Zielereignisse eingebettet, bei denen sich die Versuchspersonen erinnern sollen, eine andere Handlung auszuführen. Der Kosten- oder Interferenzeffekt (Hicks, Marsh, & Cook, 2005; Smith, 2003) bezieht sich auf die Beobachtung, dass die Latenzzeiten in der fortlaufenden Aufgabe oft erhöht sind, wenn prospektives Erinnern erforderlich ist. Solche Veränderungen in den Latenzzeiten wurden auch gefunden, bevor ein prospektives Zielereignis auftrat und wurden in theoretischen Arbeiten auf vorbereitende Aufmerksamkeit zurückgeführt, die für die Detektion von Zielereignissen funktional relevant sein soll (Smith, 2003). Gegenstand dieser Dissertation war die modellbasierte Analyse von Interferenzeffekten, um die zugrunde liegenden kognitiven Prozesse genauer zu identifizieren und voneinander zu trennen. Dabei wurden mit dem Diffusionsmodell (Ratcliff, 1978) simultan Veränderungen in der Geschwindigkeit und Genauigkeit der Bearbeitung der fortlaufenden Aufgabe berücksichtigt. In allen sechs Datensätzen wurde eine Erhöhung der Latenzzeiten durch die prospektive Gedächtnisaufgabe gefunden. Ein zentrales Modellierungsergebnis war, dass prospektive Aufgaben eine vorsichtigere Bearbeitung der fortlaufenden Aufgaben induzieren (d.h., eine Veränderung im Geschwindigkeits-Genauigkeits Kriterium a) und zu Verzögerungen in peripheren, nicht entscheidungsbezogenen Prozessen (Modellparameter T_{er}) führen können. Im Vergleich zwischen jüngeren und älteren Erwachsenen zeigte sich dabei ein qualitativ sowie quantitativ ähnliches Muster.

Abstract

Prospective-memory tasks require to remember intended actions at some point in the future (e.g., remembering to buy groceries on the way home). As a characteristic feature, these tasks involve self-initiated retrieval from memory without explicit reminders and are highly relevant in everyday life. To examine prospective memory in a laboratory paradigm, participants are engaged in an ongoing task (e.g., a lexical-decision task). Participants are then asked to remember to carry out another action on prespecified target events, which are embedded in the ongoing task. The cost- or interference effect (Hicks, Marsh, & Cook, 2005; Smith, 2003) refers to the finding that ongoing-task latencies are often increased with embedded PM tasks. Such changes in latencies have also been observed before target occurrence and have been regarded as an indicator for preparatory attentional processes that are functionally relevant for target detection (Smith, 2003). The aim of the present dissertation was a model-based analysis of interference effects to identify and disentangle the underlying cognitive processes, using Ratcliff's (1978) diffusion model that simultaneously takes into account ongoing-task speed and accuracy. In all six datasets, we found an increase in latencies with an embedded prospective-memory task. As a core modeling result, prospective-memory tasks induce more cautious ongoing-task decisions (i.e., a change in the speed-accuracy criterion a) and may lead to increases in peripheral nondecisional processes (model parameter T_{er}). A comparison between younger and older adults revealed qualitatively and quantitatively similar patterns.

1. Einführung

Im Jahr 2002 suchte ein 59-jähriger Patient mit starken Schmerzen im Unterleib eine Klinik auf. Computertomografische Aufzeichnungen ergaben schließlich, dass behandelnde Ärzte bei einer vorangegangenen Operation schlichtweg vergessen hatten, ein 16 cm langes Operationsbesteck wieder aus dem Bauchraum des Patienten zu entfernen (Dembitzer & Lai, 2003).

Gedächtnisfehler, bei denen intendierte Handlungen wieder vergessen werden, sind den meisten Menschen auch aus dem Alltag gut bekannt. In empirischen Studien ließen sich etwa die Hälfte *aller* berichteten Gedächtnisprobleme dieser Kategorie zuordnen (Crovitz & Daniel, 1984): Wir vergessen einen vereinbarten Termin, einer Email den Anhang beizufügen, oder abends noch ein Medikament einzunehmen. Aus dieser Perspektive scheint es daher bemerkenswert, dass Menschen in der Lage sind, sich nach Tagen oder Wochen zum richtigen Zeitpunkt an eine vormals gefasste Intention zu erinnern. Wie schon von Freud (1901/2009; S. 214) bemerkt, scheint es manchmal, als ob ein gefasster Vorsatz „schlummert […] bis die Zeit seiner Ausführung herannaht". Tatsächlich haben sich Philosophen, Mediziner, und Psychologen schon früh mit dem Erinnern von Absichten und unerledigten Handlungen beschäftigt (z. B. Lewin, 1926; Zeigarnik, 1927). Eine systematischere Erforschung dieser Phänomene erfolgte allerdings erst in den letzten Jahrzehnten in der kognitiven Psychologie unter dem Begriff *prospektives Gedächtnis* (PG; Meacham & Leiman, 1975). Die ersten empirischen Arbeiten dazu (z. B. Loftus, 1971) sind meist der Feldforschung in natürlichen Kontexten zuzuordnen. So wurden Versuchspersonen gebeten, sich zu bestimmten Zeitpunkten telefonisch zu melden (Maylor, 1990) oder eine Postkarte an das Institut zurück zu schicken (Patton & Meit, 1993), während später eine Vielzahl von Publikationen unter Verwendung eines experimentellen Laborparadigmas (Einstein & McDaniel, 1990) entstand, wie auch in der vorliegenden Arbeit.

Gegenstand dieser Dissertation ist die mathematisch-formale Modellierung und experimentelle Untersuchung von *Kosten- oder Interferenzeffekten* (Hicks, Marsh, & Cook, 2005; Smith, 2003) bei prospektiven Gedächtnisaufgaben. Einleitend wird zunächst die experimentelle Versuchsanordnung beschrieben und ein kurzer Überblick über grundlegende Charakteristika von prospektiven Gedächtnisaufgaben, über Interferenzeffekte, und über aktuelle theoretische Ansätze gegeben. Außerdem werden die mathematischen Modellierungstechniken umrissen, die in den Publikationen dieser Dissertation zur Anwendung kamen, und die wichtigsten empirischen Ergebnisse zusammengefasst.

2. Prospektive Gedächtnisaufgaben

Prospektive Gedächtnisaufgaben erfordern, sich an eine Intention zu einem bestimmten Ereignis oder Zeitpunkt in der Zukunft zu erinnern und die intendierte Handlung auszuführen (McDaniel & Einstein, 2007). Das prospektive Gedächtnis wird also konzeptionell nicht als ein Gedächtnis*system* aufgefasst (so wie dies z. B. von Tulving, 1983, für das episodische oder semantische Gedächtnis postuliert wurde), sondern lässt sich am besten als *Aufgabe* definieren, die ganz bestimmte Anforderungen beinhaltet und aus mehreren Phasen besteht. Ähnlich wie beispielsweise das Lesen von Texten eine im Alltag hochrelevante Aufgabe darstellt, an der verschiedenste kognitive Funktionen beteiligt sind (Kliegl, 2007), so sind auch bei prospektiven Gedächtnis*aufgaben* neben reinen Gedächtnisprozessen eine Vielzahl weiterer kognitiver Funktionen beteiligt (z. B. vorbereitende Aufmerksamkeitsprozesse; vgl. Smith & Bayen, 2004). Dabei wird prospektives Erinnern oft als ein aus vier Phasen bestehender Prozess betrachtet: (a) die Phase der Intentionsbildung oder -enkodierung, (b) der Intentionsspeicherung, (c) der Intentionsinitiierung, und (d) der Ausführung der intendierten Handlung (Kliegel, Martin, McDaniel, & Einstein, 2002). Im Folgenden werden in Anlehnung an

McDaniel und Einstein (2007) wichtige Charakteristika prospektiver Gedächtnisaufgaben dargestellt, die sie von anderen Gedächtnisaufgaben abgrenzen.

Wenn sich Menschen an vormals intendierte Handlungen erinnern, werden Informationen aus dem Gedächtnis abgerufen, die bei der Enkodierung Teil bewusster Handlungsplanung waren und damit prinzipiell *deklarativ* sind (d.h. in einem Symbolsystem potentiell mitteilbar; vgl. Squire, 2004). Die Voraussetzung der bewussten Handlungsplanung (Miller, Galanter, & Pribram, 1960; Shallice & Burgess, 1991) bei der Enkodierung schließt somit *nondeklarative* Information (wie etwa beim assoziativen Lernen, prozeduralen Lernen, und Priming) für den Intentionsabruf aus. So wäre beispielsweise die Untersuchung reflexiver oder automatischer Handlungen auf konditionierte Reize hin kein Gegenstand der prospektiven Gedächtnisforschung.

Ein weiteres Charakteristikum prospektiver Gedächtnisaufgaben ergibt sich aus der zeitlichen *Verzögerung* (dem Retentionsintervall) zwischen der Intentionsbildung und der Möglichkeit zur Handlungsausführung. Manchmal müssen Intentionen nur über kurze Zeit behalten werden (beispielsweise, wenn wir eine E-Mail schreiben, der vor dem Abschicken noch eine Datei angehängt werden soll). Wenn wir andererseits morgens bemerken, dass Getränke eingekauft werden müssen, die wir abends nach der Arbeit im Supermarkt besorgen wollen, dann fallen in dieses längere Retentionsintervall tagsüber eine Vielzahl weiterer Aktivitäten. Nach Graf und Uttl (2001) unterscheiden sich solche Aufgaben durch das Ausmaß, in dem Intentionen bewusst und kontinuierlich im Arbeitsgedächtnis repräsentiert sind. In Vigilanzaufgaben antworten Personen demnach kontrolliert, wenn bestimmte Ereignisse in der Zukunft auftreten, wobei die entsprechende Absicht vorwiegend im Fokus der Aufmerksamkeit steht. Von prospektiven Gedächtnisaufgaben im engeren Sinn („PM proper“; S. 438), gehen Graf und Uttl in ihrer Taxonomie nur dann aus, wenn die Absicht den Fokus der Aufmerksamkeit für eine bestimmte Zeit verlassen hat. Bislang gibt es kaum Arbeiten, die diese postulierten Unterschiede empirisch untersucht haben (Brandimonte, Ferrante, Feresin, &

Delbello, 2001). In Studie 4 der vorliegenden Arbeit sind Czernochowski, Horn, und Bayen (2012) mit Hilfe ereigniskorrelierter Potentiale (EKPs) der Frage nachgegangen, ob zwischen Vigilanzaufgaben und prospektiven Gedächtnisaufgaben qualitative oder quantitative Unterschiede bestehen.

2.1 Prospektive und retrospektive Komponente

Nach Einstein und McDaniel (1990; siehe auch Einstein & McDaniel, 1996) liegen prospektiven Gedächtnisaufgaben stets zwei Komponenten zugrunde. Die prospektive Komponente umfasst die kognitiven Prozesse, die zur Initiierung der Intention führen, *dass* wir etwas ausführen wollten. Die retrospektive Komponente umfasst zwei Aspekte: (a) den Abruf des Intentionsinhalts und (b) Wiedererkennen des relevanten Kontexts bzw. Ereignisses zur Handlungsausführung (d.h. Erinnern *was* wir ausführen wollten und *wann* das geeignete Zielereignis eintritt; Smith & Bayen, 2006). Die Gedächtnisprozesse der retrospektiven Komponente sind folglich identisch mit denjenigen, die auch bei klassischen Methoden der Gedächtnisprüfung (wie der geförderten Reproduktion oder beim Wiedererkennen; vgl. Bredenkamp & Erdfelder, 1996) untersucht werden. Der entscheidende Unterschied zu klassischen Gedächtnisaufgaben besteht in den zusätzlichen Anforderungen der prospektiven Komponente, nicht nur Information aus der Vergangenheit abzurufen, sondern sich zunächst zu erinnern, *dass* etwas abgerufen werden muss („remembering to remember“; Harris, 1984). Aufgrund der hohen alltäglichen Relevanz dieses letzteren Aspektes (Crovitz & Daniel, 1984) hat sich die Erforschung prospektiven Erinnerns besonders auf die Prozesse der prospektiven Komponente und die Abgrenzung zu rein retrospektiven Prozessen konzentriert. So haben Smith und Bayen (2004; Horn, Bayen, Smith, & Boywitt, 2011) mithilfe eines multinomialen Verarbeitungsbaummodells (Batchelder & Riefer, 1999; Erdfelder et al., 2009) gezeigt, dass das Zusammenspiel beider Komponenten valide trennbar ist und der beobachtbaren prospektiven

Gedächtnisleistung zugrunde liegt. An dieser Stelle sei darauf hingewiesen, dass potentiell jede (deklarative) Gedächtnisaufgabe sowohl prospektive als auch retrospektive Anteile aufweisen kann (siehe McDaniel & Einstein, 2007). Trotzdem würden viele Gedächtnistests nicht als prospektive Aufgaben gelten, weil die Anforderungen der prospektiven im Vergleich zu denen der retrospektiven Komponente minimal sind. Beispielsweise erfordern viele Untersuchungen zum episodischen Gedächtnis, vorher gelernte Information zu einem bestimmten Zeitpunkt in der Zukunft zu reproduzieren. Der Kontext der Untersuchung, die Instruktionen, oder die Versuchsleitung machen allerdings klar, wann die relevante Abrufphase beginnt. Im Gegensatz dazu muss im Laborparadigma zum prospektiven Gedächtnis (siehe Abschnitt 3) die Erinnerung an eine intendierte Handlung *selbstinitiiert* erfolgen, ohne dass das Individuum durch eine externe Quelle explizit in einen Abrufmodus („retrieval mode"; Tulving; 1983) versetzt wird. Craik (1986) hat daher vermutet, dass prospektives Erinnern aufgrund geringerer Abrufhilfe durch externe Hinweisreize ressourcenintensive, selbstinitiierte Prozesse erfordert, die zu besonders deutlichen Leistungsunterschieden im höheren Erwachsenenalter führen sollten (siehe Smith & Bayen, 2006, oder Smith, Horn, Bayen, 2012, für eine weiterführende Diskussion zu Altersunterschieden im prospektiven Gedächtnis).

2.2 Ereignisbasierte und zeitbasierte Aufgaben

In Abhängigkeit von der kritischen Situation, in der die Intentionsinitiierung erforderlich wird, werden *ereignisbasierte* und *zeitbasierte* prospektive Gedächtnisaufgaben unterschieden (Einstein & McDaniel, 1990, 1996). In ereignisbasierten Aufgaben ist das kritische Zeitfenster für die Ausführung der intendierten Handlung mit dem Eintreten eines bestimmten Zielereignisses verknüpft. So könnten wir beispielsweise einer Kollegin eine wichtige Nachricht mitteilen, sobald wir diese auf der Arbeit antreffen. Charakteristisch für ereignisbasierte Aufgaben ist somit die Existenz eines externen (prospektiven) Hinweisreizes in der Umwelt, sobald die intendierte Handlung relevant wird. In zeitbasierten Aufgaben hingegen ist der kritische Moment für die Intentionsinitiierung alleine durch die Zeit definiert (z. B. nach 45 Minuten den Kuchen wieder aus dem Ofen nehmen oder um 19 Uhr zu einer Verabredung erscheinen). Zeitbasierte Intentionen können also unabhängig von den kontextuellen Gegebenheiten relevant werden, ohne dass in der Umgebung Hinweisreize existieren. Gegenstand der vorliegenden Arbeit sind ereignisbasierte prospektive Gedächtnisaufgaben.

3. Laborparadigma

Einstein und McDaniel (1990) haben in einer einflussreichen Arbeit ein Paradigma eingeführt, um prospektives Erinnern experimentell im Labor zu untersuchen. Ein Kerngedanke war dabei, Abrufinstruktionen durch eine externe Quelle zu vermeiden (im Gegensatz zu typischen episodischen Gedächtnistests) und auch innerhalb kürzerer Experimente die Anforderungen der prospektiven Komponente zu maximieren. In Analogie zu alltäglichen Situationen wird die prospektive Gedächtnisaufgabe dabei in andere Aktivitäten eingebettet: So sind die Versuchspersonen nach der Intentionsenkodierung mit einer fortlaufenden Aufgabe („ongoing task“) beschäftigt, während dann zu bestimmten Zeitpunkten oder Ereignissen die intendierte Handlung erinnert und ausgeführt werden muss. Typischerweise wird für diesen Fall gefordert, die fortlaufende Aufgabe entweder zu unterbrechen und stattdessen die prospektive Handlung auszuführen (Smith & Bayen, 2004) oder diese Handlung zusätzlich auszuführen (um potentielle Aufgabenwechselkosten zu vermeiden; Cohen, Jaudas, & Gollwitzer, 2008). In diesem Zusammenhang wurden auch Gemeinsamkeiten mit Doppelaufgaben und Aufgabenwechsel untersucht (Bisiacchi, Schiff, Ciccola, & Kliegel, 2009; vgl. Pashler, 1994), wobei im prospektiven Gedächtnisparadigma die Zweitaufgabe deutlich seltener relevant wird (z. B. in 4% aller Experimentaldurchgänge; Marsh, Hicks, Cook, Hansen, & Pallos, 2003) und auch vergessen werden oder unbemerkt bleiben kann.

Um zu vermeiden, dass Versuchspersonen kontinuierlich an die prospektive Gedächtnisaufgabe denken und die Intention konstant im Arbeitsgedächtnis repräsentiert ist (Einstein & McDaniel, 1990), wird nach der initialen Intentionsenkodierung manchmal ein Verzögerungsinterall mit Distraktoraufgaben eingesetzt. Neuere Arbeiten zeigen aber, dass solche Verzögerungsintervalle zu entgegengesetzten Effekten führen können und eine Verlängerung der fortlaufenden Aufgabe effektiver sein kann, um kontinuierliches Memorieren zu vermeiden (Martin, Brown, & Hicks, 2011). Tabelle 1 fasst die Schritte einer typischen prospektiven Laboraufgabe zusammen.

Tabelle 1

Übersicht: Laborparadigma für prospektive Gedächtnisaufgaben

1. Einführung einer fortlaufenden Aufgabe *(„ongoing task“)* mit mehreren Übungsdurchgängen. (Verwendete fortlaufende Aufgaben in bisherigen Studien z. B.: lexikalische Entscheidungen, Kurzzeitgedächtnisaufgaben, Rating-Aufgaben, Farbübereinstimmungsaufgaben, bekannte Gesichter benennen, Satzverifikation, etc.)
2. Instruktionen zur prospektiven Gedächtnisaufgabe (z. B., „Drücken Sie eine bestimmte Taste, wenn Sie während der fortlaufenden Aufgabe das Wort *‚Tiger‘* sehen“).
3. Distraktorintervall, in dem die Versuchspersonen mit einer anderen Tätigkeit beschäftigt werden (z. B., Fragebogen ausfüllen, Reaktionszeitaufgabe, etc.)
4. Wiedereinführung der fortlaufenden Aufgabe. Die Versuchspersonen werden nicht mehr an die prospektive Gedächtnisaufgabe erinnert und bearbeiten Durchgänge der fortlaufenden Aufgabe.
5. Die kritischen prospektiven Zielereignisse (z. B., bestimmte Wörter wie *‚Tiger‘*) treten eingebettet in die fortlaufende Aufgabe auf. Das prospektive Gedächtnis, daraufhin eine zugewiesene intendierte Handlung auszuführen, wird getestet.

Als konkretes Beispiel folgt die allgemeine Versuchsanordnung der vorliegenden Experimente: Hier waren die Versuchspersonen als fortlaufende Aufgabe mit *lexikalischen Entscheidungen* beschäftigt, bei denen so schnell und akkurat wie möglich angegeben werden soll, ob es sich bei einer präsentierten Buchstabenkette um ein Wort (z. B. „*Haus*“) oder kein Wort (z. B. „*Huas*“) der deutschen Sprache handelt. Diese binäre Entscheidungsaufgabe wurde sowohl in der prospektiven Gedächtnisforschung (z. B. Marsh et al., 2003; Smith, 2003) als auch für Reaktionszeitmodellierung (siehe Abschnitt 6; z. B. Ratcliff, Gomez, & McKoon,

2004) intensiv eingesetzt, so dass Vergleiche mit anderen Datensätzen möglich werden. Nach einer anfänglichen Übungsphase werden die Versuchspersonen dann aufgefordert, diese Aufgabe weiterhin zu bearbeiten, aber künftig auch daran zu denken, bei bestimmten Wörtern (z. B., „*Tiger*“) statt der lexikalischen Entscheidung eine andere Aufgabe auszuführen (z. B. die Taste „*F1*“ zu drücken; Einführung der prospektiven Gedächtnisaufgabe). Nach der Enkodierung der prospektiven Zielereignisse (z. B. einer Liste von Wörtern: „*Tiger*“, „*Hund*“, etc.) erfolgt dann die Bearbeitung einer lexikalischen Entscheidungsaufgabe mit eingebetteten Zielereignissen, wobei die sekundäre prospektive Aufgabe keinerlei weitere Erwähnung findet.

4. Theoretische Ansätze

Verschiedene theoretische Ansätze konzentrieren sich vorwiegend auf die kognitiven Prozesse der Intentionsinitiierung beim ereignisbasierten prospektiven Gedächtnis (Guynn, 2003; Marsh et al., 2003; McDaniel & Einstein, 2000; Smith, 2003). In diesem Zusammenhang wurde auch kontrovers diskutiert, ob die Detektion von prospektiven Zielereignissen auf automatischen oder kontrollierten Prozessen beruht (Einstein & McDaniel, 2010; Smith, 2010; vgl. Hasher & Zacks, 1979, oder Kahneman & Treisman, 1984, für Kriterien automatischer kognitiver Verarbeitung).

Einige theoretische Modelle gehen davon aus, dass Stimuli in der Umgebung die spontane, automatische Detektion des Zielereignisses auslösen und zum Abruf der intendierten Handlung führen. Basierend auf introspektiven Berichten von Versuchspersonen (McDaniel & Einstein, 2007; vgl. Bargh & Chartrand, 1999) erscheint diese Annahme zunächst intuitiv plausibel (beispielsweise wenn wir durch einen Briefkasten „spontan“ daran erinnert werden, noch die Postkarte einzuwerfen; siehe aber Nisbett & Wilson, 1977, oder Pongratz, 1998, für eine kritische Diskussion introspektiver Evidenz). So wurde postuliert, dass relevante Zielereignisse ein Gefühl der Vertrautheit („familiarity“; Einstein & McDaniel, 1996) oder Diskrepanz (Breneiser & McDaniel, 2006; vgl. Whittlesea & Williams, 2000) auslösen, welches in Folge zu einer kontrollierten Suche im Gedächtnisspeicher führt, warum das Ereignis besonders wahrgenommen wurde. Ferner wurde in Anlehnung an retrospektive Gedächtnismodelle (vgl. Moscovitch, 1994) postuliert, dass hinreichend starke Assoziationen zwischen Zielereignis und intendierter Handlung zu einer reflexiven Intentionsinitiierung führen können, sobald das relevante Ereignis eintritt (McDaniel, Guynn, Einstein, & Breneiser, 2004).

Aus einer alternativen theoretischen Perspektive erfordert Intentionsinitiierung meist exekutive Kontrolle (Craik, 1986; Marsh & Hicks, 1998). Insbesondere die *„preparatory attentional and memory processes theory“* (PAM; Smith, 2003, 2010) geht davon aus, dass ressourcenintensive, vorbereitende Aufmerksamkeitsprozesse stets eine *notwendige* Bedingung für erfolgreiches prospektives Erinnern sind. Solche Prozesse können nach Smith, Hunt, McVay, und McConnel (2007) beispielsweise eine kontrollierte Suche nach prospektiven Zielreizen („monitoring“) oder das Memorieren von Intentionen beinhalten.

Die *„multiprocess view“* (MPV; McDaniel & Einstein, 2000, 2007) hingegen postuliert, dass in Abhängigkeit von bestimmten Randbedingungen sowohl kontrollierte als auch automatische Prozesse zur Intentionsinitiierung und Detektion von relevanten Zielereignissen führen. Das Ausmaß erforderlicher ressourcenintensiver Prozesse hängt nach diesem Modell ab von (a) den Eigenschaften der prospektiven Gedächtnisaufgabe (z. B. deren Wichtigkeit), (b) der fortlaufenden Aufgabe (z. B. Komplexität), und (c) individuellen Merkmalen der Personen (z. B. Persönlichkeitseigenschaften; vgl. McDaniel & Einstein 2007, für eine detaillierte Beschreibung spezifischer Randbedingungen).

5. Interferenzeffekte prospektiver Gedächtnisaufgaben

Eine zentrale empirische Methode, kontrollierte Prozesse in prospektiven Gedächtnisaufgaben zu erfassen, besteht in der Analyse der fortlaufenden Aufgabe („ongoing task“). Dabei werden meist die Latenzzeiten oder Fehlerraten in einer Kontrollbedingung (in der nur die fortlaufende Aufgabe bearbeitet wird) mit einer Bedingung verglichen, in der gleichzeitig prospektives Erinnern erforderlich ist (z. B. Cohen et al., 2008; Einstein et al., 2005; Marsh et al., 2003; Smith et al., 2007). Ähnlich wie in Doppelaufgabenparadigmen (vgl. Pashler, 1994) wird angenommen, dass kontrollierte Prozesse der prospektiven Komponente ggf. kognitive Ressourcen benötigen, die dann von der Bearbeitung der fortlaufenden Aufgabe abgezogen werden. Smith (2003) zeigte in einer ersten Arbeit, dass die Latenzzeiten bei einer fortlaufenden lexikalischen Entscheidungsaufgabe reliabel erhöht waren, wenn gleichzeitig prospektives Erinnern erforderlich war. Entscheidend hierbei war, dass dieser *Interferenz-* oder *Kosteneffekt* (Hicks, Marsh, & Cook, 2005; Smith, 2003) in den Latenzzeiten auch gefunden wurde, *bevor* ein prospektives Zielereignis auftrat. Wird nämlich bei einem Zielereignis die intendierte Handlung ausgeführt, so sind alleine aufgrund von Intentionsabruf und Aufgabenwechsel erhöhte Latenzzeiten zu erwarten (Marsh, Hicks, & Watson, 2002).[1] Der prospektive Interferenzeffekt wurde von Smith auf vorbereitende Aufmerksamkeit zurückgeführt, die für die Detektion von Zielereignissen funktional relevant sein soll. In Übereinstimmung mit dieser Annahme wurden wiederholt positive Korrelationen zwischen Interferenzausmaß und prospektiver Gedächtnisleistung nachgewiesen (z. B. Smith & Bayen, 2004; West, Krompinger, & Bowry, 2005).

In Folgearbeiten wurde untersucht und diskutiert, ob Interferenzeffekte prospektiver Gedächtnisaufgaben generell auftreten oder von bestimmten Randbedingungen abhängen (z. B.

[1] Aus diesem Grund werden für die Analyse der Performanz der fortlaufenden Aufgabe meist die prospektiven Zielereignisse und einige nachfolgende Experimentaldurchgänge ausgeschlossen (z. B. Smith & Bayen, 2004).

Einstein et al., 2005; Einstein & McDaniel, 2010; Smith, 2010; vgl. Smith et al., 2007, für eine Übersicht empirischer Arbeiten). So berichten beispielsweise Einstein et al., dass die prospektive Gedächtnisleistung bei ausbleibender Interferenz dennoch hoch sein kann, wenn die relevanten Merkmale der Zielereignisse distinkt sind und auch für die Bearbeitung der fortlaufenden Aufgabe im Fokus der Aufmerksamkeit stehen (sog. *„fokale"* Aufgaben; siehe Kliegel, Jäger, & Phillips, 2008, und Abschnitt 7.2 für weitere Details).

Trotz intensiver Forschung zum prospektiven Interferenzeffekt stützen sich theoretische Ansätze auf das relativ breite Konstrukt des kognitiven Ressourcenverbrauchs (z. B. Einstein et al., 2005; Smith et al., 2007; vgl. Kahneman, 1973). Ferner gingen bisherige Analysen kaum über Unterschiede von Latenzzeitmittelwerten hinaus. Die Prozesse, die zu solchen Unterschieden führen sind bislang hingegen kaum untersucht (McDaniel & Einstein, 2007, S. 225). Ziel der vorliegenden Arbeit war eine modellbasierte Analyse von Interferenzeffekten. Dabei stand weniger die Frage im Fokus, ob (und unter welchen Randbedingungen) prospektives Erinnern mit fortlaufenden Aufgaben interferiert. Vielmehr sollte untersucht werden, welche Prozesse diesen Effekt überhaupt erklären können. Dabei wurden mathematische Modelle verwendet, die nicht nur Latenzzeitmittelwerte, sondern die gesamten Latenzzeitverteilungen und auch die Fehlerraten simultan berücksichtigen. Formale Modellierung bietet dabei den allgemeinen Vorteil, dass sonst verbal formulierte Zusammenhänge zwischen theoretischen Konstrukten und Empirie mathematisch exakt beschrieben werden, die Modellannahmen durch beobachtete Daten testbar sind und mit konkurrierenden Annahmen verglichen werden können (z. B. Klauer, 2002; Lewandowsky & Farrell, 2010).

6. Zur Analyse mit stochastischen Diffusionsmodellen

Kognitive Prozessmodelle wie das Diffusionsmodell (Ratcliff, 1978; für verwandte Modelle vgl. Brown & Heathcote, 2008; Usher & McClelland, 2001) können zum Verständnis von prospektiven Interferenzeffekten beitragen, indem die zugrunde liegenden Ursachen für schnelle oder langsame Entscheidungen und für hohe oder niedrige Fehlerraten beleuchtet werden. Dabei trennt das Diffusionsmodell eine Reihe psychologisch interpretierbarer Komponenten voneinander, die auch für Theorien des prospektiven Gedächtnisses informativ sind: die Geschwindigkeit der Informationsaufnahme, die Zeit für periphere Prozesse wie Enkodierung und motorische Ausführung, die Schwellendistanz, welche festlegt wie viel Information bis zur Entscheidung benötigt wird, sowie Variabilität in diesen Komponenten.

Das Diffusionsmodell eignet sich auch deshalb gut zur Performanzanalyse von fortlaufenden Aufgaben, da Latenzzeiten und Akkuratheit simultan berücksichtigt werden. Dies stellt aufgrund der generellen Austauschbeziehung zwischen Geschwindigkeit und Genauigkeit (*„speed-accuracy-tradeoff“*; vgl. Pachella, 1974; Pew, 1969; Schouten & Bekker, 1967) eine wichtige methodische Verbesserung gegenüber der separaten Analyse dieser abhängigen Variablen dar: so können Versuchspersonen bei gleicher Fähigkeit eine Aufgabe langsamer bearbeiten, um Fehler zu vermeiden, oder umgekehrt schneller antworten und dabei mehr Fehler riskieren. Gerade für die Analyse von fortlaufenden Aufgaben ist es wichtig, solche „trade-offs“ explizit zu erfassen, bevor Rückschlüsse über ressourcenintensive Prozesse beim prospektiven Erinnern getroffen werden können. So könnten die Latenzzeiten erhöht sein, weil Versuchspersonen umsichtiger bzw. „konservativer“ antworten, und nicht weil die fortlaufende Aufgabe schwieriger ist oder für sie weniger Ressourcen zur Verfügung stehen.

6.1 Modellparameter

Das Diffusionsmodell (z. B. Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & Tuerlinckx, 2002) eignet sich für die Analyse von schnellen *binären* Entscheidungen.[2] Eine grundlegende Modellannahme besteht darin, dass entscheidungsrelevante Information in einem stochastischen Wiener-Diffusionsprozess kontinuierlich über die Zeit akkumuliert wird. Ähnlich wie auch in Signalentdeckungsmodellen (vgl. Macmillan & Creelman, 2005) wird dabei eine Dimension der Signalstärke angenommen (die Ordinate in Abbildung 1, die in arbiträren Informationseinheiten Evidenz für bzw. gegen die Entscheidungsalternativen abbildet). In Erweiterung wird aber auch die zeitliche Dimension berücksichtigt (die Abszisse in Abbildung 1) und ein sequentielles „sampling" modelliert (Smith & Ratcliff, 2004).

Diffusionsprozesse starten an einem Punkt z, der auf der Signalstärkeachse zwischen zwei Entscheidungsschwellen liegt. Sobald ein Prozess nach gewisser Zeit eine der Schwellen kreuzt, wird eine Entscheidung getroffen und die entsprechende (motorische) Antwort initiiert. Im Diffusionsmodell wird angenommen, dass Entscheidungsprozesse sowohl zufälligen als auch systematischen Einflüssen unterliegen. Zufälliges Rauschen im Akkumulationsprozess wird über die Diffusionskonstante s modelliert[3] (Abbildung 1 zeigt exemplarisch zwei oszillierende Akkumulationsprozesse). Auch unter identischen Bedingungen kreuzen Prozesse somit nicht determiniert zur selben Zeit dieselbe Entscheidungsschwelle (und generieren potentiell Latenzzeit*verteilungen* und auch Fehler). Daneben beeinflussen die folgenden Modellparameter Dauer und Ergebnis des Akkumulationsprozesses.

[2] Die mittleren Latenzzeiten für diesen Aufgabentyp liegen bei jungen Erwachsenen typischerweise nicht über 1-2 s. Für komplexere, strategische Entscheidungsprozesse eignen sich andere Modellklassen (siehe z. B. Luce & Raiffa, 1957).

[3] Akkumulation innerhalb eines trials mit gegebener Drift ξ variiert normalverteilt um ξ, $N(\xi, s)$. Die Standardabweichung s (die Diffusionskonstante) bestimmt dabei die Amplitude im zufälligen Rauschen und wird als nicht freier Skalierungsparameter in der Literatur meist auf die Werte 0.1 (Ratcliff, 1978; Ratcliff & Rouder, 1998) oder 1 fixiert (Voss, Rothermund, & Voss, 2004). Unterschiedliche Werte von s skalieren einige Modellparameter, ändern aber nicht ihr relatives Verhältnis zueinander.

Die *Driftrate* v stellt eine systematische Einflusskomponente auf den Akkumulationsprozess dar. Der Driftparameter quantifiziert dabei die Geschwindigkeit und Richtung der Informationsaufnahme (d.h., die durchschnittliche Steigung des Prozesses pro Zeiteinheit). Eine positive Driftrate ($v > 0$) impliziert, das durchschnittlich mehr Information zugunsten der oberen Entscheidungsschwelle (A) akkumuliert wird (und umgekehrt bei negativem Drift, $v < 0$).

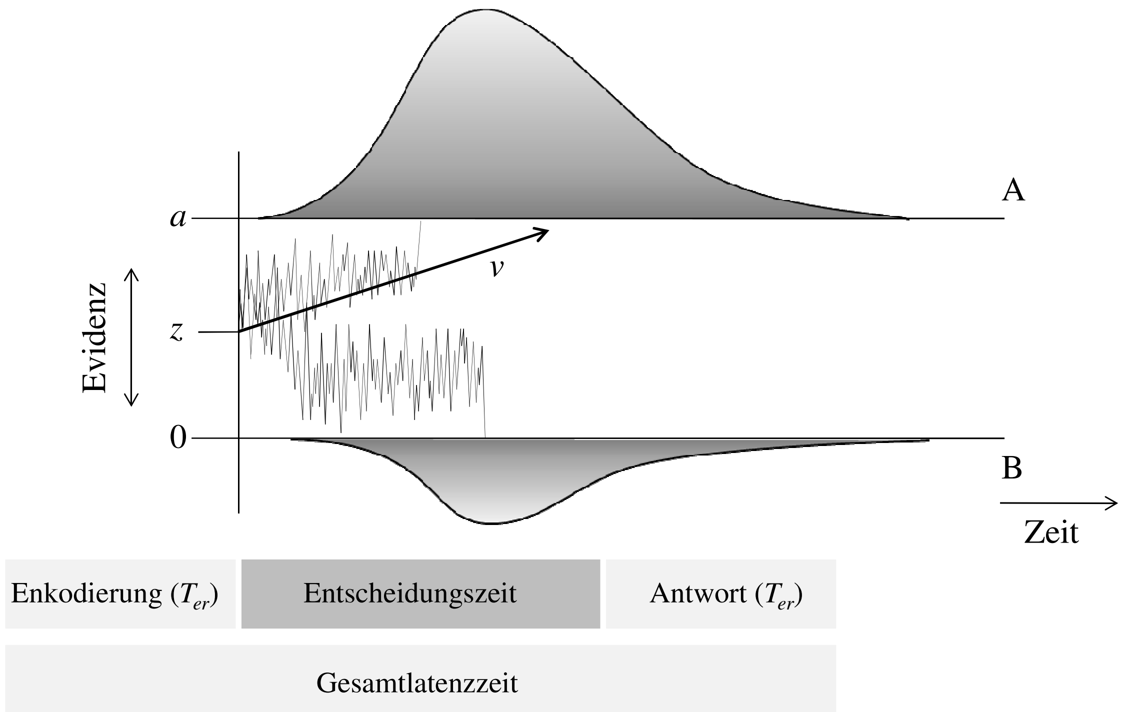


Abbildung 1. Schematische Darstellung des Diffusionsmodells.

Diffusionsprozesse starten am Punkt *z*. Eine Entscheidung („A“ oder „B“) wird getroffen, sobald die obere oder untere Entscheidungsschwelle erstmalig von einem Prozess gekreuzt wird. Die Grafik zeigt für beide Entscheidungen die entsprechenden Latenzzeitverteilungen, die über viele Experimentaldurchgänge hinweg entstehen. Die unterschiedlich großen Flächen dieser Verteilungen sind Folge der positiven Driftrate ($v > 0$), die impliziert dass Diffusionsprozesse mit höherer Wahrscheinlichkeit die obere Schwelle kreuzen. Die beobachtete Gesamtlatenzzeit setzt sich additiv aus der Entscheidungszeit und einer Zeitkonstante T_{er} für die Dauer peripherer Prozesse (Enkodierung, motorische Ausführung) zusammen. Adaptierte Grafik aus „What Can the Diffusion Model Tell us About Prospective Memory?” von S. Horn, U. J. Bayen, und R. E. Smith, 2011, *Canadian Journal of Experimental Psychology*, *65*, S. 70. © 2011 Canadian Psychological Association.

Die Driftrate modelliert die Effizienz der Informationsverarbeitung während der Entscheidungsphase, wobei höhere Driftraten zu schnelleren und akkurateren Entscheidungen führen. In interindividuellen Vergleichen ist sie daher ein Maß für die Fähigkeit des Entscheiders und quantifiziert in intraindividuellen Vergleichen die Aufgabenschwierigkeit (bzw. die von den Stimuli extrahierbare Information; vgl. Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007).

Die Distanz zwischen den beiden Entscheidungsschwellen wird durch den Parameter a quantifiziert (Abb. 1). Eine Vergrößerung der Schwellendistanz a impliziert, dass mehr Information akkumuliert werden muss, bevor eine Entscheidung getroffen werden kann. In diesem Fall werden die Antworten langsamer aber auch akkurater (ein „konservatives“ Entscheidungskriterium). Eine Verringerung der Schwellendistanz bewirkt hingegen, dass weniger Information gesammelt wird, bis eine Entscheidung getroffen wird, was zu schnellen aber weniger akkuraten Entscheidungen führt (ein „liberales“ Kriterium). Somit kann über die Schwellendistanz die Geschwindigkeits-Genauigkeits Gewichtung eines Entscheiders explizit modelliert werden.

Die relative Position des Startpunkts z zwischen den Schwellen erfasst die mögliche a-priori Tendenz zugunsten einer bestimmten Entscheidung („bias“). Bei gleicher Distanz zu beiden Schwellen ($z = a/2$) existiert a-priori keine Bevorzugung einer Entscheidungsalternative. Abweichungen von $a/2$ hingegen implizieren Asymmetrien im Ausmaß der Information, das für die Entscheidungsalternativen jeweils benötigt wird: Startpunktpositionen oberhalb von $a/2$ reflektieren eine Bevorzugung zugunsten der oberen Entscheidungsschwelle, für die dann weniger Information benötigt wird (und umgekehrt für $z < a/2$).

Im Diffusionsmodell wird angenommen, dass sich die beobachtete Latenzzeit additiv aus der Dauer des Entscheidungsprozesses (wie oben beschrieben) und einer Zeitkonstante T_{er} für periphere, nicht entscheidungsbezogene Prozesse zusammensetzt (siehe auch Luce, 1986). Der Parameter T_{er} quantifiziert die Dauer dieser peripheren Prozessen *vor und nach* der Entscheidungsphase, wie beispielsweise perzeptueller Enkodierung und motorischer Antwortausführung. Veränderungen in T_{er} führen im Modell zur Verschiebung der kompletten Latenzzeitverteilung um einen bestimmten Betrag, ohne dabei die Vorhersage der Fehlerraten zu beeinflussen.

Das Modell in der oben beschrieben Fassung (für verwandte Vorläufermodelle, siehe z. B. Laming, 1968; Stone, 1960) nimmt Homogenität der Parameter über verschiedene Experimentaldurchgänge an. Ein wesentlicher Beitrag Ratcliffs (1978; Ratcliff & Rouder, 1998) bestand dabei in der Erweiterung des Modells, um Variabilität in den Parametern über verschiedene Situationen hinweg zu berücksichtigen. So wird in einer erweiterten Modellfassung angenommen, dass sich die Driftraten mit Standardabweichung η normal um den Mittelwert ν verteilen. Der Startpunkt liegt über die Experimentaldurchgänge hinweg gleichverteilt im Intervall s_z um den Mittelwert z. Für die Zeitkonstante wird ebenfalls eine Gleichverteilung mit Mittelwert T_{er} und Intervall s_t angenommen. Mit Variabilität in der Driftrate η können empirische Effekte modelliert werden, bei denen falsche Antworten systematisch langsamer ausfallen als korrekte Antworten, während die Variabilität des Startpunkts (s_z) entgegengesetzte Effekte modellieren kann (Laming, 1968; Ratcliff & Rouder, 1998). Änderungen der Form der Latenzzeitverteilung aufgrund der schnellsten Antworten („leading edge“) können durch Variabilität der Zeitkonstante (s_t) erfasst werden (Ratcliff, Gomez, & McKoon, 2004; Ratcliff & Tuerlinckx, 2002). Hohe Werte der Variabilitätsparameter deuten auf Heterogenität im Stimulusmaterial oder starke Fluktuation in kognitiven Zuständen während der Bearbeitung hin. In den Analysen der vorliegenden Arbeit wurden die Variabilitätsparameter zwar geschätzt, standen aber im Gegensatz zu den zentraleren Parametern nicht im Fokus der Fragestellung. So wird eine gute Modellanpassung bisweilen auch ohne Berücksichtigung von Parametervariabilität erreicht (vgl. Voss, Rothermund, & Voss, 2004). Für die Schätzung der Parameter des Diffusionsmodells und zur Bestimmung der Modellpassung existieren verschiedene Verfahren, die im Folgenden kurz umrissen werden.

6.2 Parameterschätzung und Modellpassung

Unter bestimmten Annahmen lassen sich die die latenten Maße ν, a, und T_{er} über einfache Transformationen direkt aus den Daten errechnen. Wagenmakers, van der Maas, und Grasman (2007) haben in einem vereinfachten „*EZ Diffusionsmodell*“ gezeigt, dass aus nur drei empirischen Werten (Anteil korrekter Antworten, Mittelwert und Varianz der Latenzzeiten) die

Driftrate v, die Schwellendistanz a, und die Zeitkonstante T_{er} errechnet werden kann, *falls* keine Variabilität in den Parametern und keine a-priori Entscheidungstendenz angenommen wird (d.h., $z = a/2$; vgl. Wagenmakers, van der Maas, Dolan, & Grasman, 2008, für zusätzliche Modellierung der Startpunktposition z). Hierbei entsteht eine saturierte Modellvariante (ähnlich wie etwa im klassischen Signalentdeckungsmodell; Macmillan & Creelman, 2005), bei der zusätzlich die zentralen Voraussetzungen dieser Berechnungsmethode überprüft werden sollten (siehe Wagenmakers et al., 2007, für weitere Details). Die „*EZ*" Methode erlaubt gegebenenfalls eine erste Interpretation und eine sehr schnelle Bestimmung der zentralen Modellparameter, die für andere Schätzverfahren oft als Startwert dient (Vandekerckhove & Tuerlinckx, 2007; Voss & Voss, 2007). Allerdings wird dabei die Form der Latenzzeitverteilung ignoriert (vgl. Ratcliff, 2008, für eine kritische Diskussion), und eine Testung der Modellpassung oder Restriktionen bestimmter Parameter über Experimentalbedingungen sind nicht möglich.

Bei einer Analyse mit dem vollständig parametrisierten Diffusionsmodell (z. B. Ratcliff & Tuerlinckx, 2002) werden für eine bestimmte Konstellation der Parameterwerte die vom Modell vorhergesagten Latenzzeitverteilungen errechnet. In einer multidimensionalen, iterativen Suche werden dann die Parameterwerte nach einer vorgegebenen Zielfunktion so optimiert, dass die Abweichung zwischen vorhergesagten (theoretischen) und beobachteten (empirischen) Daten minimal wird. Die Parameterschätzung kann dabei in Abhängigkeit von der Zielfunktion (vgl. Ratcliff & Tuerlinckx, 2002) sehr rechenintensiv sein, wobei die Entwicklung von Programmen für die Modellanalyse die Anwendung für Experimentalpsychologen deutlich erleichtert hat (Vandekerckhove & Tuerlinckx, 2007, 2008; Voss & Voss, 2007; vgl. van Ravenzwaaij & Oberauer, 2009, für vergleichende Simulationsstudien).

In dieser Arbeit wurde für die Schätzung der Modellparameter vorwiegend das Programm *fast-dm* eingesetzt (Voss & Voss, 2007, 2008), welches als Zielfunktion die Teststatistik des Kolmogorov-Smirnov (KS) Tests verwendet (Kolmogorov, 1941). Als Optimierungskriterium wird dabei die maximale vertikale Distanz (die KS-Statistik) zwischen der beobachteten und der vorhergesagten kumulativen Verteilungsfunktion der Latenzzeit minimiert. Weil in diesem Schätzverfahren jeweils nur eine theoretische mit einer beobachteten kumulativen Verteilungsfunktion verglichen wird, ist es notwendig, die Latenzzeiten von korrekten und falschen Entscheidungen in einer gemeinsamen Verteilung zu vereinigen. Zu diesem Zweck werden die Fehlerlatenzen mit dem Wert –1 multipliziert, und die entsprechende Fehlerlatenzzeitverteilung ist auf der Zeitachse am Nullpunkt gespiegelt (Voss et al., 2004).

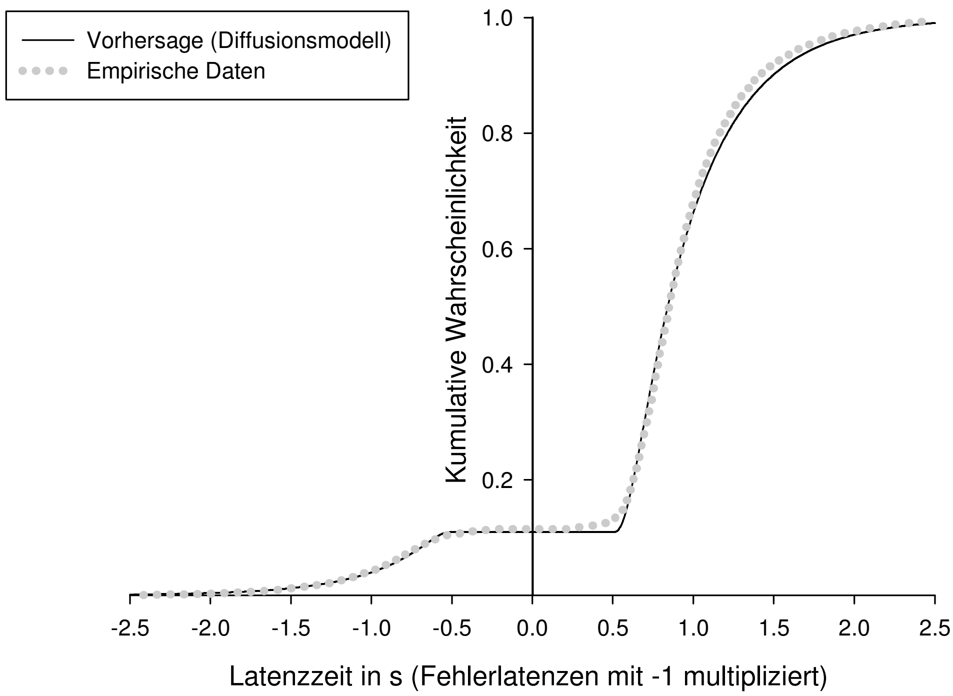


Abbildung 2. Kumulative Verteilungsfunktionen.

Theoretische (vom Modell vorhergesagte) und empirische kumulative Verteilungsfunktion der Latenzzeit (Beispiel aus Studie 2, Experiment 2). Für die Parameterschätzung mit der KS-Methode wurde die maximale vertikale Abweichung zwischen beiden Verteilungsfunktionen minimiert (vgl. Voss et al., 2004). Insgesamt zeigt sich eine gute Passung zwischen empirischen und theoretischen Daten. Die Fehlerrate (Schnittpunkt mit der Ordinate) liegt bei ca. 11%. Negative Werte auf der Abszisse entsprechen den Latenzzeiten von Fehlern (mit –1 multipliziert) und positive Werte sind die Latenzen korrekter Antworten.

Abbildung 2 zeigt exemplarisch eine theoretische und eine empirische Verteilungsfunktion aus Studie 2 (aggregiert über die Daten aller Versuchspersonen, die eine fortlaufende lexikalische Entscheidungsaufgabe bearbeiteten). Der Anteil der Verteilung auf der negative Seite der Abszisse (Zeitachse) entspricht dabei der Fehlerlatenzzeitverteilung und der Anteil auf der positiven Seite der Verteilung korrekter Antworten. Am Schnittpunkt dieser Funktion mit der Ordinate kann der Anteil falscher Entscheidungen abgelesen werden.

Ein Vorteil der KS-Schätzmethode besteht in einer effizienten Nutzung der Information in den Daten, da keine Kategorisierung der Latenzzeitverteilungen erforderlich ist (wie etwa bei Quantil-basierten Schätzverfahren; siehe unten) und die Schätzungen gegenüber Ausreißern in den Latenzzeiten relativ robust sind (Voss & Voss, 2007, 2008). Somit können auch bei Studien, in denen die Anzahl der Experimentaldurchgänge eher gering ist (z. B. $N \approx 100$; vgl. Klauer et al., 2007) bereits Modellanalysen durchgeführt werden. Allerdings erlaubt die KS-Methode keine systematischen Vergleiche zwischen konkurrierenden oder verschachtelten Modellen, bei denen bestimmte Parameter über Bedingungen restringiert werden (z. B., über χ^2-Differenztests oder den Vergleich von Informationskriterien wie AIC oder BIC; vgl. Donkin, Brown, & Heathcote, 2011). Ferner sollten die ausgegebenen p-Werte dieser Schätzmethode nicht als exakte Wahrscheinlichkeiten des KS-Anpassungstests interpretiert werden[4] (Voss, Rothermund, Gast, & Wentura, in press).

Quantil-basierte Parameterschätzung (wie etwa die Minimierung einer χ^2-Statistik als Zielfunktion) beruht hingegen auf einer Kategorisierung der stetigen Latenzzeitverteilungen für Fehler und korrekte Antworten. In vielen Analysen wird dabei eine Unterteilung in sechs Kategorien durch das .10, .30, .50 (Median), .70, und .90 Quantil vorgenommen, die die Form der Verteilungen meist hinreichend approximiert (Ratcliff & McKoon, 2008). Abbildung 3 zeigt dieselben Daten aus Studie 2 als Quantil-Wahrscheinlichkeits-Diagramm („Quantile-Probability-Plot"), welches häufig zur Visualisierung der Modellpassung und der Daten bei Quantil-basierter Schätzung eingesetzt wird (z. B. Ratcliff & McKoon, 2008; Ratcliff &

[4] Die vom Modell vorhergesagten kumulativen Verteilungsfunktionen werden in diesem Schätzverfahren *post-hoc* an die beobachteten Daten angepasst. Dies führt dazu, dass der KS-Tests konservativer wird. Bei Übereinstimmung zwischen empirischer und theoretischer Verteilungsfunktion wird die Nullhypothese der Übereinstimmung somit *seltener* als $N \cdot \alpha$ verworfen (bei N Tests und Signifikanzniveau α; vgl. Voss et al., in press).

Rouder, 1998; Ratcliff et al., 2004). Für die Schätzung wurde hier das Programm *DMAT* (Vandekerckhove & Tuerlinckx, 2007, 2008) verwendet.

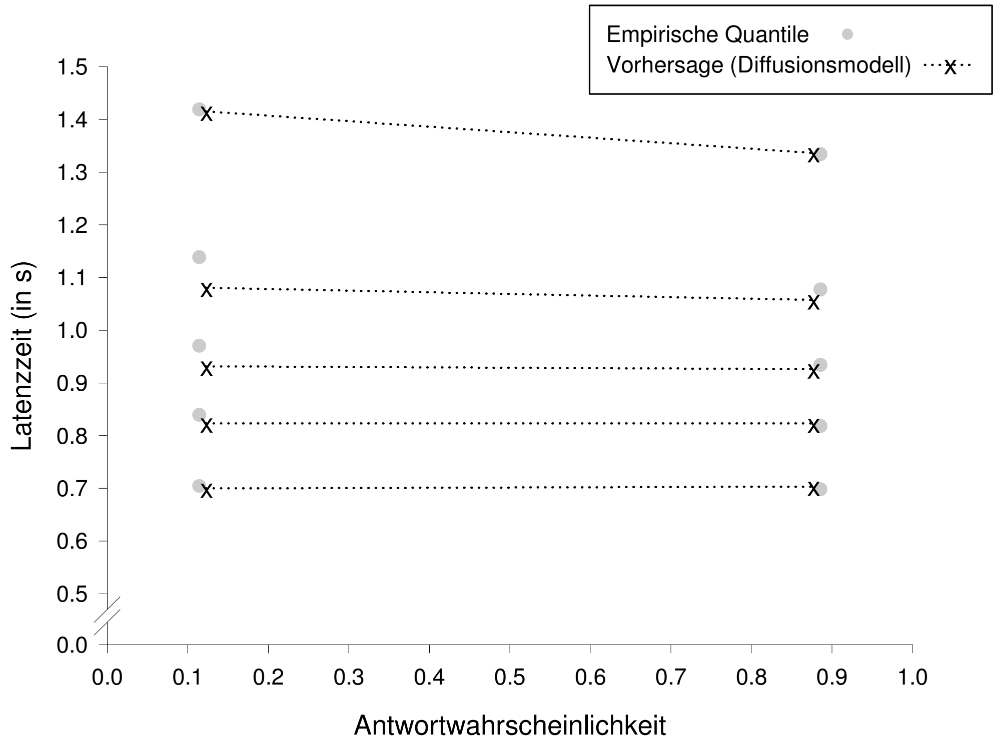


Abbildung 3. Quantil-Wahrscheinlichkeit-Diagramm.

Empirische und theoretische Quantile der Latenzzeitverteilungen für korrekte und falsche Antworten, aggregiert über die Versuchspersonen in dieser Bedingung (Beispiel aus Studie 2, Experiment 2). Fünf Latenzzeit-Quantile (.1, .3., .5, .7, .9) sind auf der Ordinate gegen die Wahrscheinlichkeit falscher Antworten ($p \approx .11$) und korrekter Antworten ($1 - p$) auf der Abszisse abgetragen. Die Latenzzeitverteilung für korrekte Antworten befindet sich somit auf der rechten Seite des Diagramms und die dazu komplementäre Verteilung für die Fehler auf der linken Seite.

Als Optimierungskriterium werden dann diejenigen Parameterwerte gesucht, bei denen über die einzelnen Kategorien hinweg die χ^2-Abweichung zwischen beobachteten und theoretischen Häufigkeiten (bzw. Antworten) minimal wird. Für dieses Verfahren müssen innerhalb jeder Kategorie k genügend Beobachtungen vorliegen (z. B. $n_k > 5$), was bei einfachen Aufgaben viele Experimentaldurchgänge erforderlich macht, da insbesondere die Form der Fehlerlatenzzeitverteilung sonst nicht reliabel erfasst werden kann. So sind mit dem Programm *DMAT* (welches für die Analyse in Abbildung 3 verwendet wurde) mindestens 11 Fehlerlatenzen erforderlich, um die Parameter einer Bedingung schätzen zu können. Ferner beinhaltet die Kategorisierung einer stetigen Verteilung ein subjektives Element (vgl. auch

Klauer et al., 2007).[5] Ein Vorteil Quantil-basierter Schätzverfahren besteht in systematischen Vergleichen zwischen Modellen mit verschiedenen Parameterrestriktionen. So kann beispielsweise dasjenige Diffusionsmodell ermittelt werden, das nach einem gewählten Informationskriterium (z. B. AIC oder BIC) einen optimalen Kompromiss zwischen der Anzahl freier Parameter und der Modellpassung darstellt (Donkin et al., 2011). Tabelle 2 zeigt exemplarisch die Kennwerte verschiedener Modelle (Daten aus Studie 2), bei denen unterschiedliche Parameter zwischen den Experimentalbedingungen frei variieren können. Dabei zeigt sich, dass ein Modell mit nur einem freiem Parameter (T_{er}) die niedrigste Summe in den BIC Werten (über die Stichprobe) aufweist und nach dem BIC Kriterium somit die optimale Menge freier Parameter besitzt.

Tabelle 2. *Parameterrestriktionen, Kennwerte der Modellpassung und Modellselektion*

Modellvariante	Frei variierende Parameter zwischen Experimentalbedingungen	M (χ^2)	K	∑ (BIC)
DM 1	–	77.77	6	299634
DM 2 †	$\boldsymbol{T_{er}}$	**31.11**	**7**	**296494**
DM 3	a	37.37	7	296958
DM 4	T_{er}, a	26.68	8	296633
DM 5	T_{er}, a, v, η, s_z, s_t	16.34	12	297934

Anmerkung. Beispiel für Studie 2 (Exp.2). † = Modelvariante mit niedrigsten BIC Werten über die Stichprobe; M (χ^2) = Mittelwert der approximativ χ^2-verteilten Prüfstatistik der Modellpassung; die Anzahl der Freiheitsgrade des Anpassungstests für ein Modell mit K freien Parametern, M Experimentalbedingungen (Stufen der Manipulation), und 2 × 5 Latenzzeitquantilen (.1, .3, .5, .7, und .9; jeweils für korrekte und falsche Entscheidungen) beträgt $df = M \times (12 - 1) - K$.

Abschließend sei darauf hingewiesen, dass in letzter Zeit für das Diffusionsmodell auch hierarchische Verfahren entwickelt wurden, bei denen die Variabilität zwischen Personen oder Items auf verschiedenen Ebenen geschätzt werden kann (vgl. Vandekerckhove, Tuerlinckx, & Lee, 2011, für eine Bayesianische Implementation mit dem Programm

[5] Erfolgt die Kategorisierung der Latenzzeitverteilung nicht über *a-priori* festgelegte Zeitmarken, sondern basierend auf Quantilen (datengeleitet), dann sollten die p-Werte streng genommen ebenfalls nicht als exakte Wahrscheinlichkeiten des χ^2-Tests interpretiert werden (vgl. Vandekerckhove & Tuerlinckx, 2007).

WinBUGS; Lunn, Thomas, Best, & Spiegelhalter, 2000). Neben der expliziten Modellierung von verschiedenen Variabilitätsquellen kann dies große Analysevorteile bieten, falls aggregierte Daten von vielen Versuchspersonen vorliegen, aber nur mit jeweils wenigen Beobachtungen pro Person.

6.3 Anwendungsbereiche und Parametervalidität

Bereits bei Einführung des Diffusionsmodells in die kognitive Psychologie war es ein erklärtes Ziel, Entscheidungsprozesse in ganz unterschiedlichen experimentellen Paradigmen zu modellieren (Ratcliff, 1978, S. 59). Für die Modellanalyse stehen daher Parameter zur Verfügung, die einerseits hinreichend spezifisch sind, um relevante Entscheidungsprozesse voneinander zu trennen, andererseits Allgemeingültigkeit für verschiedene Aufgaben besitzen sollen. Unter anderem wurde das Diffusionsmodell sehr erfolgreich für die Analyse episodischer Gedächtnisurteile eingesetzt (z. B. Ratcliff, 1978; Ratcliff & Starns, 2009; Spaniol, Madden, & Voss, 2006), für lexikalische Entscheidungsprozesse (Ratcliff et al., 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008), für die Analyse altersbedingter Verlangsamung (z. B. Ratcliff, Thapar, & McKoon, 2006), implizite soziale Kognition (Klauer et al., 2007), und Primingeffekte (Voss et al., in press).

In den meisten Anwendungen ergab sich dabei eine gute Passung des Modells an die empirischen Daten (d.h. die kompletten Latenzzeitverteilungen und Fehlerraten; vgl. Abschnitt 6.2). Mehrere Arbeiten haben zudem gezeigt, dass die Modellparameter die postulierten kognitiven Prozesse valide abbilden und voneinander trennen können (Voss et al., 2004). Auch im Rahmen lexikalischer Entscheidungsaufgaben existieren mehrere Befunde, die für die diskriminante Validität (vgl. Campbell & Fiske, 1959) der Parameter sprechen. So haben Ratcliff et al. (2004) gezeigt, dass lexikalische Variablen, die sich auf die Aufgabenschwierigkeit auswirken (Worthäufigkeit und orthographische Wortähnlichkeit der Nichtwörter), selektiv durch Änderungen in der Driftrate erfasst werden. Instruktionen und Rückmeldungen, schnell oder akkurat zu antworten, beeinflussen systematisch den Schwellenparameter a, während die relative Präsentationshäufigkeit der Stimuli die Entscheidungstendenz und Position des Startpunktes z verändert (Wagenmakers et al., 2008).

Ausgehend von diesen Validierungsergebnissen wurden in den vorliegenden Studien für die fortlaufende Aufgabe lexikalische Entscheidungen verwendet, die sowohl mit dem Diffusionsmodell intensiv untersucht wurden als auch in der prospektiven Gedächtnisforschung oft als Standardaufgabe eingesetzt werden (z. B. Cohen et al., 2008; Marsh et al., 2002, 2003; Smith, 2003). Interferenzeffekte auf lexikalische Entscheidungen durch prospektive Gedächtnisaufgaben wurden allerdings bislang kaum modellbasiert untersucht. Kernanliegen der vorliegenden Studien war es daher, die dieser Verlangsamung zugrunde liegenden Prozesse insbesondere mit Hilfe formaler Modellierung (aber auch mit weiteren Methoden; vgl. Studie 4) zu untersuchen.

7. Zentrale Ergebnisse empirischer Studien

7.1 Studie 1: Modellierung der Daten von Smith (2003)

In dieser ersten Arbeit wurden Daten einer einflussreichen Studie zum prospektiven Interferenzeffekt von Smith (2003; Experiment 1) mit dem Diffusionsmodell reanalysiert. Das Ziel war es, zunächst die Anwendbarkeit und Passung des Modells im prospektiven Gedächtnisparadigma an einem etablierten Datensatz zu prüfen. Im Experiment von Smith bearbeiteten alle Versuchspersonen 504 Durchgänge einer lexikalischen Entscheidungsaufgabe und lernten vorher sechs prospektive Zielwörter, bei denen die intendierte Handlung ausgeführt werden sollte (Taste *F1* drücken). In der prospektiven Gedächtnisgruppe (n = 62) wurden die Versuchspersonen instruiert, sich während der folgenden lexikalischen Entscheidungen an diese prospektive Zusatzaufgabe zu erinnern. In der Kontrollgruppe wurde darauf hingewiesen, dass die prospektive Gedächtnisaufgabe erst in einem späteren Abschnitt des Versuchs relevant wird. Somit waren beide Gruppen hinsichtlich zu lernender Wörter und intendierter Handlung (retrospektive Komponente) parallelisiert, während der erwartete Kontext der Intentionsinitiierung (prospektive Komponente) variierte. Nach Smith sollten daher mögliche Unterschiede in den Latenzen auf Interferenz durch die prospektive Komponente (vorbereitende Aufmerksamkeitsprozesse) zurückzuführen sein. Unsere Analyse der fortlaufenden Aufgabe (unter Ausschluss der prospektiven Zielereignisse; vgl. Abschnitt 5) zeigte, dass die Latenzen der prospektiven Gedächtnisgruppe (M = 925 ms, SD = 163 ms) gegenüber der Kontrollgruppe (M = 747 ms, SD = 111 ms) erhöht waren, $t(87.42) = 6.26$, $p < .001$, und die prospektive Gedächtnisleistung (M = .69, SD = .20; der Anteil der Wörter, bei denen die intendierte Handlung korrekt ausgeführt wurde) positiv mit den Latenzen in der fortlaufenden Aufgabe korreliert war, $r(62) = .34$, $p < .007$.

Tabelle 3. *Diffusionsmodellparameter aus Studie 1*

Parameter	Beschreibung	Kontroll-gruppe		Prospektive Gruppe		$t(93)$ [a]
z	Mittlere Startpunktposition	$a/2$		$a/2$		--
a	Schwellendistanz	1.91	(0.08)	2.22	(0.06)	3.15 **
v	Mittlere Driftrate	3.03	(0.20)	2.36	(0.07)	-3.90 **
T_{er}	Mittlere Zeitkonstante	0.41	(0.01)	0.43	(0.01)	0.94
s_z	Intervall der Startpunkposition	0.73	(0.09)	0.72	(0.04)	-0.12
s_t	Intervall der Zeitkonstante	0.04	(0.01)	0.04	(0.01)	-0.13
η	Variabilität der Driftrate	0.51	(0.13)	0.50	(0.05)	-0.10
KS-Statistik [b]		0.02	(0.001)	0.02	(0.001)	-1.33
p-Wert [c]		.90	(0.02)	.93	(0.02)	0.80

Anmerkung. Die individuellen Parameterschätzungen wurden über die Versuchspersonen einer Gruppe gemittelt. Standardfehler in Klammern. KS = Kolmogorov-Smirnov. [a] Test für unabhängige Stichproben; [b] KS-Distanz zwischen theoretischer und empirischer kumulativer Verteilungsfunktion der Latenzzeit; [c] p-Wert des KS-Anpassungstests; ** $p < .01$. Adaptierte Tabelle aus „What Can the Diffusion Model Tell us About Prospective Memory?” von S. Horn, U. J. Bayen, und R. E. Smith, 2011, *Canadian Journal of Experimental Psychology*, *65*, S. 71. © 2011 Canadian Psychological Association.

Tabelle 3 enthält die wichtigsten Modellierungsergebnisse. Die Parameter des Diffusionsmodells wurden individuell für jede Versuchsperson geschätzt.[6] Insbesondere ergab die Auswertung der KS-Tests und die visuelle Inspektion der Daten (siehe Horn, Bayen, & Smith, 2011, S. 73) eine akzeptable Modellpassung, die eine weitere Anwendung des Diffusionsmodells im prospektiven Gedächtnisparadigma nahelegte. Die Modellparameter zeigten, dass Versuchspersonen in der prospektive Gedächtnisgruppe die fortlaufende Entscheidungsaufgabe generell umsichtiger bearbeiteten und mehr Information anforderten, bevor eine lexikalische Entscheidung erfolgte (größere Schwellendistanz a; Tabelle 3). Auch war die Effizienz der Informationsverarbeitung während der lexikalischen Entscheidungen in dieser Gruppe reduziert. Allerdings sind die Parameterschätzungen mit gewisser Vorsicht zu interpretieren, da die Fehlerraten in diesem Datensatz generell niedrig waren. Darauf aufbauend wurden in den Folgestudien die fortlaufenden lexikalische Entscheidungsaufgaben für robustere Modellanalysen modifiziert.

7.2 Studie 2: Determinanten der Interferenz durch prospektive Gedächtnisaufgaben

In den drei Experimenten dieser Studie wurden Eigenschaften prospektiver Gedächtnisaufgaben modellbasiert untersucht, die das Ausmaß der Interferenz beeinflussen. Die Versuchsanordnung wurde zunächst im Hinblick auf spätere Modellanalysen angepasst. So wurde zur Stabilisierung der Latenzzeiten ein anfänglicher Trainingsblock eingefügt, in dem die Versuchspersonen bei zu langsamen oder falschen Antworten Rückmeldungen erhielten. Die Anzahl der Experimentaldurchgänge in der fortlaufenden Aufgabe wurde erhöht (N = 1000) und lag damit deutlich über bisherigen Anwendungen im prospektiven Gedächtnisparadigma

[6] Für die Modellanalyse wurden die Antworten über die beiden Stimulustypen (Wörter, Nichtwörter) der lexikalischen Entscheidungsaufgabe aggregiert. Die Antworten wurden somit als *korrekt* (obere Entscheidungsschwelle) oder *falsch* (untere Entscheidungsschwelle) klassifiziert. Mit einer solchen Kodierung kann die Startpunktposition fixiert werden ($z = a/2$), falls beide Stimulustypen gleich häufig auftreten (siehe Horn et al., 2011, und Klauer et al., 2007, für weitere Details).

(z. B. Marsh et al., 2002, 2003). Ferner wurden für die lexikalische Entscheidungsaufgabe schwierigere Stimuli verwendet (niedrige Worthäufigkeiten; Pseudowörter mit hoher orthographischer Ähnlichkeit zu den Wörtern), die die Variabilität in den Fehlerraten erhöhen. In bisherigen Studien wurde der prospektive Interferenzeffekt meist in den Latenzzeiten der lexikalischen Entscheidungen gefunden, während gleichzeitig Deckeneffekte in der Genauigkeit auftraten (Marsh et al., 2002). Der Einfluss prospektiven Erinnerns auf die Austauschbeziehung zwischen Geschwindigkeit und Genauigkeit könnte dadurch verdeckt worden sein.

In Experiment 1 bearbeiteten Versuchspersonen in der prospektiven Gedächtnisgruppe (*PG*) in einer ersten Phase die fortlaufende Aufgabe in einem Kontrollblock, bevor in einer zweiten Phase die prospektive Gedächtnisaufgabe eingeführt wurde. Dies ermöglichte es, Interferenz *innerhalb* der Versuchspersonen zu analysieren und individuelle Veränderungen in den Modellparametern mit der prospektiven Gedächtnisleistung in Verbindung zu bringen. Ferner wurde in allen drei Experimenten eine Kontrollgruppe (*KG*) eingesetzt, in der nur die lexikalische Entscheidungsaufgabe bearbeitet wurde, um potentielle Ermüdungs- oder Übungseffekte abschätzen zu können. In Experiment 1 wurde der prospektive Interferenzeffekt mit einer sogenannten *nonfokalen* Aufgabe (vgl. Kliegel et al., 2008; McDaniel & Einstein, 2007) untersucht, bei der die Versuchspersonen bei bestimmtem Anfangsbuchstaben (*G*, *H*, *M*) ihre lexikalischen Entscheidungen unterbrechen sollten, um die intendierte Handlung auszuführen. In theoretischen Arbeiten zum prospektiven Gedächtnis (Marsh et al., 2003; McDaniel & Einstein, 2007; Smith et al., 2007) wird übereinstimmend postuliert, dass unter solchen Bedingungen ressourcenintensive Aufmerksamkeit für die Intentionsinitiierung erforderlich wird. Im Gegensatz zu den Ziel*wörtern* im Experiment von Smith (2003) wird angenommen, dass die relevanten Merkmale der Zielereignisse während der fortlaufenden lexikalischen Entscheidungen nicht im Fokus der Bearbeitung stehen, und für die Intentionsinitiierung zusätzliche Aufmerksamkeit auf die Anfangsbuchstaben gerichtet werden muss (Scullin, McDaniel, Shelton, & Lee, 2010). Die wichtigsten Modellierungsergebnisse der fortlaufenden lexikalische Entscheidungsaufgabe sind im oberen Abschnitt von Abbildung 4 zusammengefasst. Die Diffusionsmodellparameter wurden dabei separat für jede Versuchsperson geschätzt.

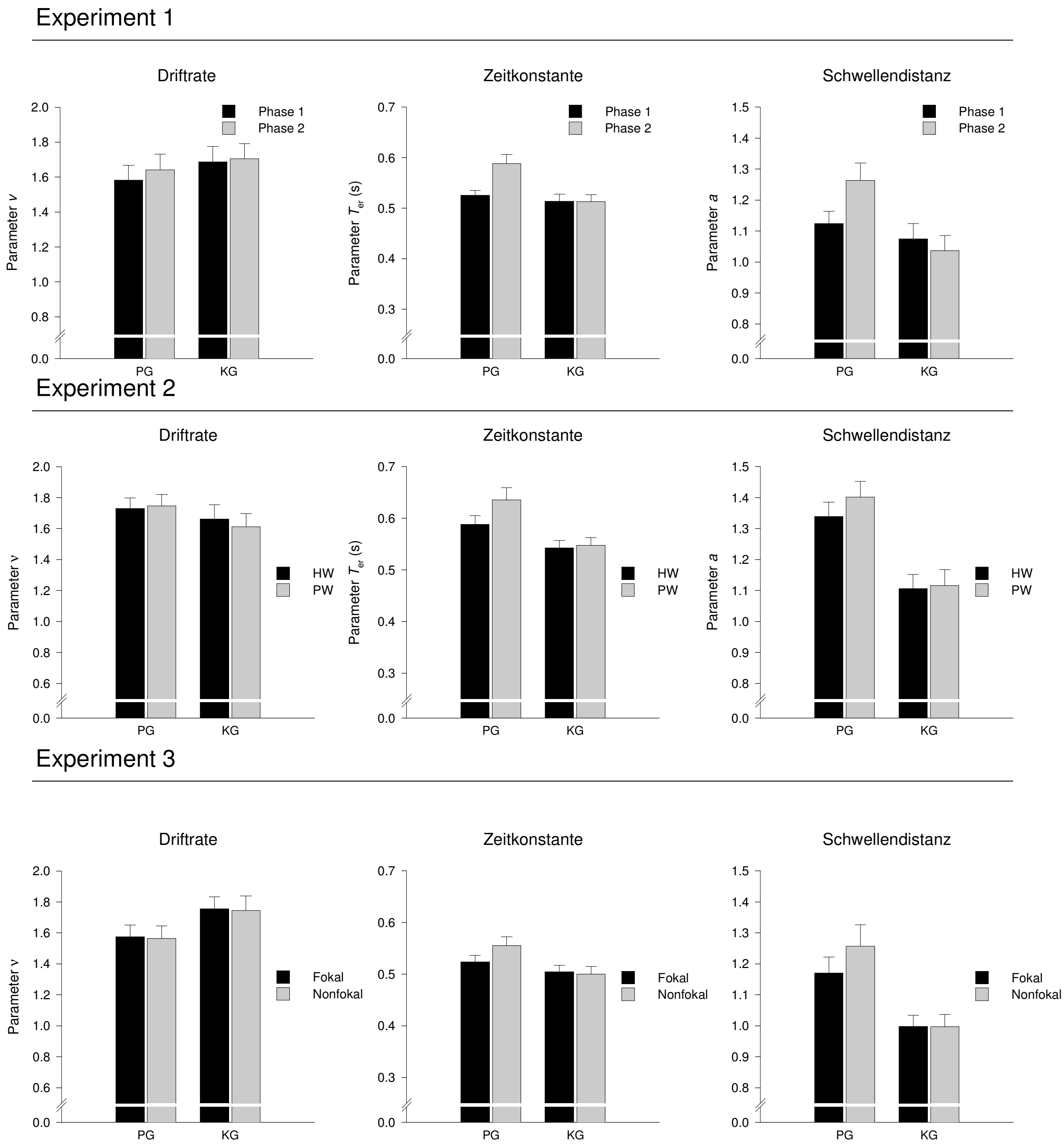


Abbildung 4. Mittelwerte der Driftrate *v*, Zeitkonstante T_{er}, und Schwellendistanz *a* aus Studie 2.

Schätzungen der Diffusionsmodellparameter für die Experimente 1 bis 3 als Funktion der Versuchsbedingung und Experimentalgruppe (*PG* = prospektive Gedächtnisgruppe; *KG* = Kontrollgruppe). Die Fehlerbalken kennzeichnen die Standardfehler der Mittelwerte. *HW* = Wichtigkeit der fortlaufenden Aufgabe betont; *PW* = Wichtigkeit der prospektiven Gedächtnisaufgabe betont. Diffusionskonstante *s* = 1.

Wie bereits in Studie 1 induzierte die prospektive Gedächtnisaufgabe einen deutlichen Anstieg der Schwellendistanz (Parameter *a*): Versuchspersonen in der *PG* Gruppe bearbeiteten die lexikalische Entscheidungsaufgabe generell langsamer, aber auch akkurater. Diese Ergebnisse deuten darauf hin, dass ein Teil des prospektiven Interferenzeffektes auf eine Verschiebung im Geschwindigkeits-Genauigkeits-Kriterium zurückzuführen ist. Die Geschwindigkeit der Informationsverarbeitung (Driftrate *v*) wurde durch eine zusätzliche prospektive Aufgabe hingegen nicht beeinträchtigt, während die Zeit für periphere Prozesse *vor und nach* der Entscheidungsphase erhöht war (Zeitkonstante T_{er}).

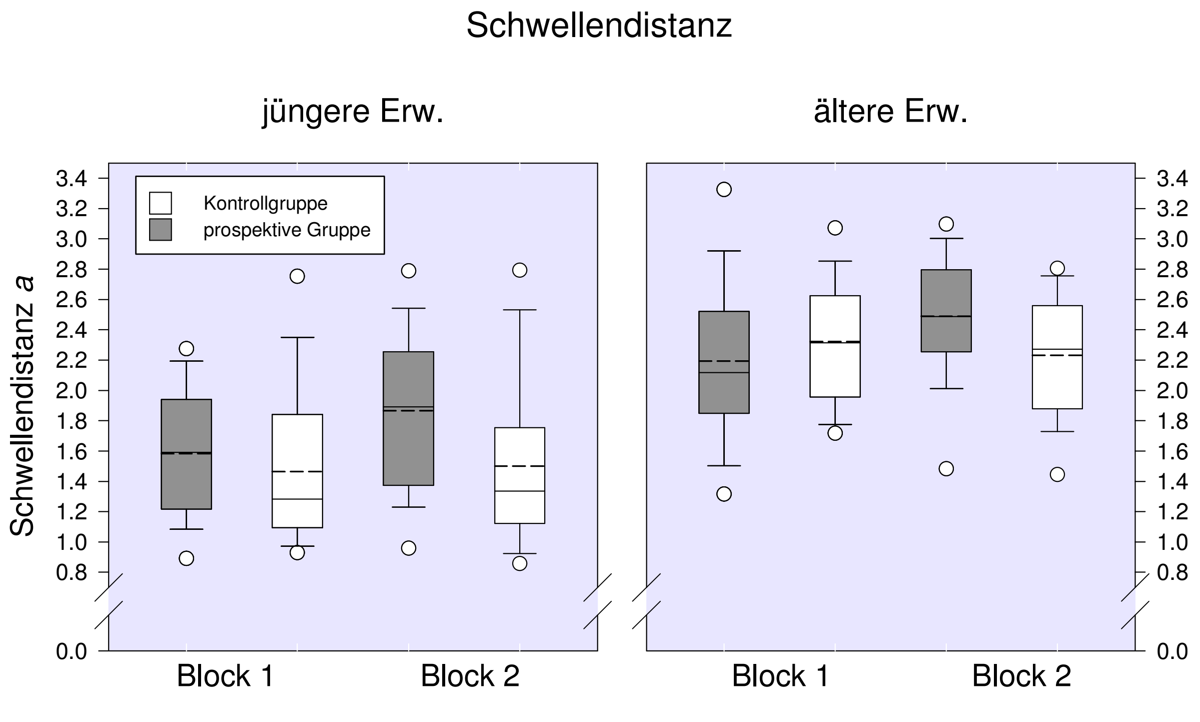
In den beiden Folgeexperimenten dieser Studie wurden innerhalb der *PG* Gruppe die Eigenschaften der prospektiven Gedächtnisaufgabe systematisch variiert. Aus mehreren Arbeiten ist bekannt, dass die wahrgenommene relative Wichtigkeit einer intendierten Handlung das Ausmaß prospektiver Interferenz massiv beeinflussen kann (Kliegel, Martin, McDaniel, & Einstein, 2001, 2004). Hierfür wurde bislang unspezifisch eine Intensivierung vorbereitender Aufmerksamkeit (z. B. Smith & Bayen, 2004) verantwortlich gemacht. Ziel von Experiment 2 war es, diesen Wichtigkeitseffekt genauer modellbasiert zu untersuchen. In einem Block des Experiments wurde über Instruktionen die Wichtigkeit der fortlaufenden Aufgabe betont (d.h., Genauigkeit und Schnelligkeit der lexikalischen Entscheidungen), während in einem anderen Block die Wichtigkeit der prospektiven Aufgabe betont wurde (d.h., möglichst kein prospektives Zielereignis zu verpassen). Die Auswertung der Modellparameter zeigte, dass die Schwellendistanz *a* und die Zeitkonstante T_{er} erhöht waren (Abbildung 4), wenn die prospektive Gedächtnisaufgabe wichtiger war. In Replikation von Experiment 1 waren es auch genau diese beiden Komponenten, die im Vergleich zur Kontrollgruppe dem prospektiven Interferenzeffekt zugrunde lagen. In Experiment 3 wurde systematisch die *Fokalität* (McDaniel & Einstein, 2007) der Aufgabe manipuliert. In einem Block des Experiments waren die relevanten Zielereignisse bestimmte Anfangsbuchstaben (wie bereits in Experiment 1 und 2; *nonfokale* Bedingung), während in einem anderen Block die Zielereignisse durch bestimmte Wörter definiert wurden (vgl. Smith, 2003; *fokale* Bedingung). Mehrere Studien fanden bislang unter fokalen Bedingungen eine verminderte (oder sogar eliminierte; Einstein et al., 2005) Interferenz bei gleichzeitig relativ hoher prospektiver Gedächtnisleistung (für eine kritische Diskussion, siehe Einstein & McDaniel, 2010; Smith, 2010). Auch fallen Altersunterschiede in der prospektiven Gedächtnisleistung bei nonfokalen Aufgaben in der Regel größer aus (Kliegel et al., 2008). Als Erklärung wird in theoretischen Arbeiten angeführt, dass unter fokalen Bedingungen die Merkmale der prospektiven Zielereignisse auch für die Bearbeitung der fortlaufenden Aufgabe relevant sind, während unter nonfokalen Bedingungen zusätzliche Aufmerksamkeit (bzw. exekutive Kontrolle) aufgewendet werden muss (McDaniel & Einstein, 2007). Der Fokalitätseffekt in den Latenzzeiten von Experiment 3 ließ sich auf Veränderungen in der Schwellendistanz a und in der Zeitkonstante T_{er} zurückführen: Versuchspersonen bearbeiteten die fortlaufenden lexikalischen Entscheidungen in der fokalen im Vergleich zu nonfokalen Bedingung schneller und riskierten dabei mehr Fehler; zusätzlich waren hier periphere Prozesse vor und nach der eigentlichen Entscheidungsphase verkürzt. In der nonfokalen Bedingung wurde im Vergleich mit der Kontrollgruppe wieder der prospektive Interferenzeffekt in den Parametern a und T_{er} repliziert, während in der fokalen Bedingung alleine Parameter a erhöht war. Dieses qualitativ unterschiedliche Muster in den Prozesskomponenten deutet darauf hin, dass bei fokalen Aufgaben die Verzögerungen vor/nach der eigentlichen Entscheidungsphase weniger ins Gewicht fallen, in Übereinstimmung mit der meist niedrigeren Interferenz bei diesem Aufgabentyp (vgl. Kliegel et al., 2008). Beachtlicherweise fanden sich in der vorliegenden Studie keinerlei Effekte in den Driftraten.

7.3 Studie 3: Modellierung von Altersunterschieden

Potentielle Altersunterschiede in der Interferenz durch prospektive Gedächtnisaufgaben wurden bislang in nur wenigen Studien untersucht. Während der prospektive Interferenzeffekt in den Latenzzeiten bei älteren Erwachsenen manchmal größer als bei jüngeren Erwachsenen ausfiel (Gao, Cheung, Chan, Chu, & Lee, 2012; McDaniel, Einstein, & Rendell, 2008), wurde auch ähnlich große Interferenz in beiden Altersgruppen gefunden (Einstein, McDaniel, & Scullin, 2011), tendenziell sogar mit größeren Effekten bei jüngeren Erwachsenen (Smith & Bayen, 2006). Eine mögliche Erklärung für dieses heterogene Befundmuster liegt in der großen interindividuellen Variabilität, mit der insbesondere ältere Erwachsene ihre limitierten kognitiven Ressourcen zwischen der prospektiven Aufgabe und der fortlaufenden Aufgabe aufteilen (vgl. Hicks et al., 2005). Ein weiteres methodisches Problem bei der Interpretation der Altersunterschiede in den genannten Studien resultiert aus der möglichen Abhängigkeit von einem Skalierungsfaktor (z. B. aufgrund genereller Verlangsamung in der Verarbeitungsgeschwindigkeit; Salthouse, 1996). So ist eine stärkere Interferenz bei älteren Erwachsenen aufgrund einer prospektiven Gedächtnisaufgabe (d.h., eine Interaktion *Alter* $\times$ *Bedingung*) nicht interpretierbar, wenn diese Interaktionen bei anderer Skalierung der abhängigen Variable (z. B. durch monotone Transformation der Latenzzeiten) eliminiert wäre (Wagenmakers, Krypotos, Criss, & Iverson, 2012; vgl. Verhaeghen, 2000, für eine kritische Diskussion von altersbedingten Interferenzunterschieden bei Stroop-Aufgaben). Eine Möglichkeit, dieses Problem zu adressieren, besteht in der direkten Schätzung der interessierenden latenten Variablen. Im Diffusionsmodell ist die Beziehung zwischen diesen Variablen und den manifesten Daten explizit definiert (vgl. Wagenmakers et al., 2012) und die Latenzzeiten werden nicht als eine prozessreine Variable behandelt.

In Studie 3 wurden 43 ältere Erwachsene (Alter: M = 70 Jahre; SD = 5 Jahre) und 46 jüngere Erwachsene (M = 23 Jahre; SD = 3.5 Jahre) verglichen. Eine Kontrollgruppe bearbeitete nur die lexikalische Entscheidungsaufgabe in zwei Blöcken, während in einer prospektiven Gruppe im zweiten Block eine zusätzliche prospektive Gedächtnisaufgabe eingeführt wurde. Der Versuchsplan enthielt somit die Zwischensubjektfaktoren *Alter* (jüngere vs. ältere Erwachsene), *Experimentalgruppe* (Kontrollgruppe vs. prospektive Gruppe), und den Innersubjektfaktor *Block* (1, 2). Generell waren die lexikalischen Entscheidungen älterer Erwachsener langsamer, $F(1, 85) = 41.89$, $\eta_p^2 = .33$, $p < .001$, aber auch akkurater, $F(1, 85) = 26.55$, $\eta_p^2 = .24$, $p < .001$, als von jüngeren Erwachsenen. Die Modellanalyse führte dieses Ergebnis auf ein größere Schwellendistanz a (d.h., konservativeres Entscheidungsverhalten) und eine längere Zeitkonstante T_{er} (verlangsamte Enkodierung und motorische Ausführung) bei älteren Erwachsenen zurück, in Übereinstimmung mit einer ganzen Reihe von Altersstudien (z. B. Ratcliff et al., 2006). Ferner war die Interferenz durch eine prospektive Gedächtnisaufgabe bei älteren Erwachsenen etwas größer als bei den jüngeren Erwachsenen [$F(1, 41) = 5.05$, $\eta_p^2 = .11$, $p = .03$, für die Interaktion *Alter* × *Block* in den Latenzzeiten in der *PG*-Gruppe].

Abbildung 5 zeigt die wichtigsten Ergebnisse der Diffusionsmodellanalyse. In Übereinstimmung mit den vorausgehenden Analysen ließ sich der prospektive Interferenzeffekt bei älteren und jüngeren Erwachsenen wieder auf einen Anstieg in der Schwellendistanz a und der Zeitkonstante T_{er} zurückführen: eine prospektive Gedächtnisaufgabe führt demnach zu konservativerem Entscheidungsverhalten in der fortlaufenden Aufgabe und zur Verzögerung peripherer Prozesse vor und nach der eigentlichen Entscheidungsphase.



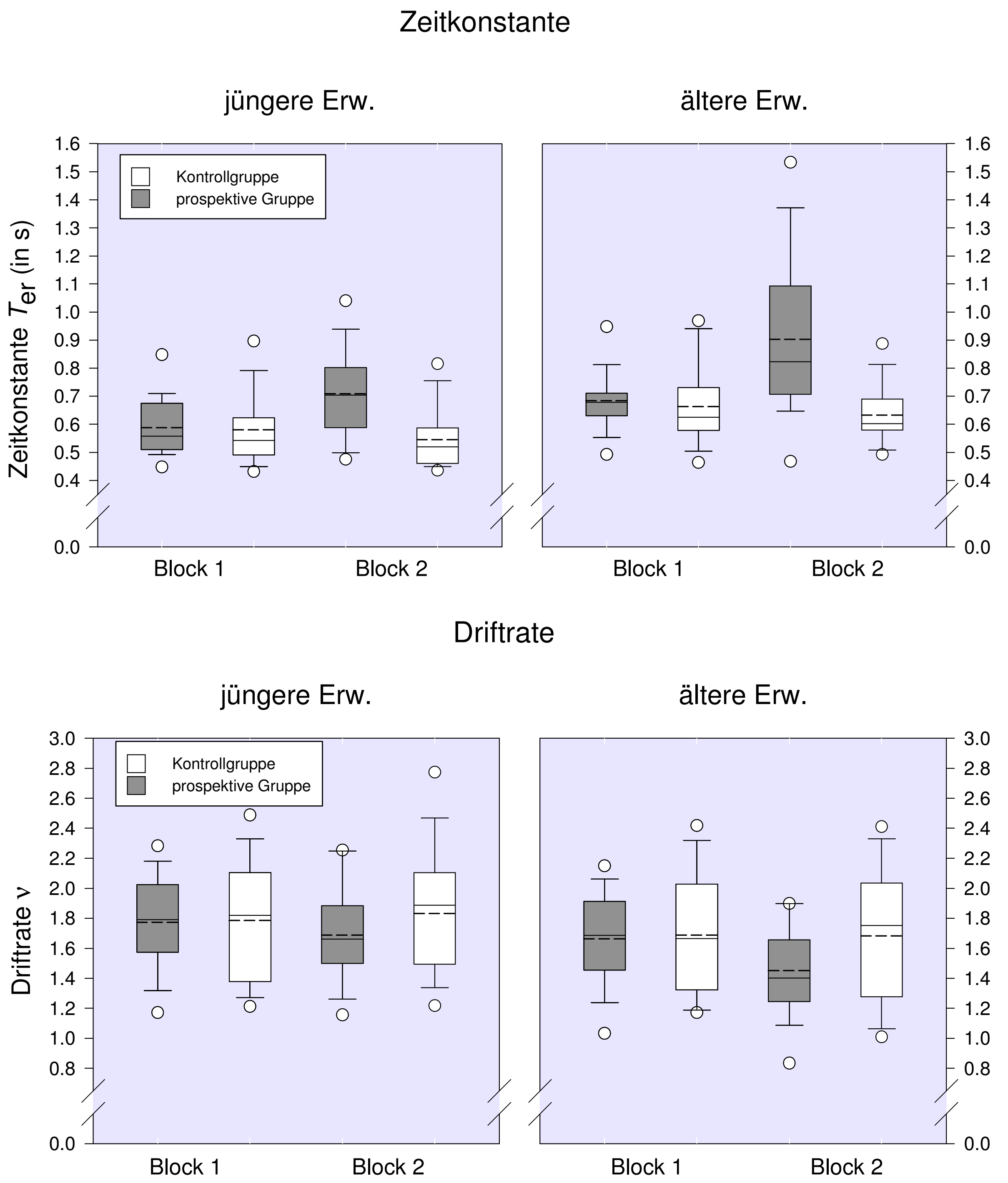


Abbildung 5. Verteilungen der Schwellendistanz a, Zeitkonstante T_{er}, und Driftrate v, aus Studie 3.

Diffusionsmodellparameter als Funktion der Altersgruppe, Experimentalgruppe, und Block. Die prospektive Gedächtnisaufgabe wurde nach Block 1 eingeführt. Die Boxplot Diagramme zeigen die Mediane (durchgezogene Linien) und Mittelwerte (gestrichelte Linien). Die Fehlerbalken über bzw. unter den Boxen markieren das 10. bzw. 90. Perzentil, und Punkte markieren das 5. bzw. 95. Perzentil der Verteilung. Diffusionskonstante s = 1.

Zusätzlich fanden wir bei den älteren Erwachsenen eine Reduktion in der Geschwindigkeit der Informationsaufnahme (Driftrate *v*), die möglicherweise auf limitierte kognitive Ressourcen in dieser Altersgruppe hinweist (vgl. Salthouse, 1988). Beachtlich war, dass sich in den modellbasierten Maßen bei gleicher Teststärke keine Altersunterschiede in der Interferenz durch die prospektiven Gedächtnisaufgabe fanden (d.h., ein Ausbleiben von *Altersgruppe* × *Block* Interaktionen, im Gegensatz zu der beobachteten Interaktion in den Latenzzeiten). Dieses Ergebnis kann auf die nichtlineare Beziehung zwischen den interessierenden latenten Prozessen und den beobachteten Latenzzeiten zurückgeführt werden (Wagenmakers et al., 2012). Zusammenfassend sprechen die Modellierungsergebnisse dieser Altersstudie somit für qualitativ und quantitativ sehr ähnliche Auswirkungen einer prospektiven Aufgabe bei älteren und jüngeren Erwachsenen (vgl. auch Verhaeghen, 2000, für eine Diskussion gleichförmiger Interferenzeffekte zwischen Altersgruppen).

7.4 Studie 4: Häufigkeitseffekte beim prospektiven Erinnern

In der Literatur zum prospektiven Gedächtnis wurden generelle Unterschiede in den kognitiven Prozessen zwischen solchen Aufgaben angenommen, bei denen die Intention konstant im Arbeitsgedächtnis repräsentiert ist (sogenannte „Vigilanzaufgaben" nach Graf & Uttl, 2001) und solchen, bei denen die Intention den Fokus der Aufmerksamkeit verlässt und wieder selbstinitiiert abgerufen werden muss (prospektive Gedächtnisaufgaben im engeren Sinn; „PM proper"; Graf & Uttl, 2001, S. 438). Um diese beiden Situationen im Labor systematisch zu vergleichen, wurde in der vorliegenden Studie die Auftretenswahrscheinlichkeit der prospektiven Zielereignisse (3% vs. 20%) manipuliert. Dabei wurde angenommen, dass bei hoher Auftretenswahrscheinlichkeit eine intensivierte Suche nach relevanten Ereignissen („monitoring"; Guynn, 2003) zu erhöhten Kosten in der fortlaufenden Aufgabe führt, aber auch zu einer konsistenteren Repräsentation der Intention im Arbeitsgedächtnis (vgl. Baddeley, 2003). Bislang wurde kaum Forschung zu dieser Frage durchgeführt (siehe aber Brandimonte et al., 2001), obwohl in vielen neurowissenschaftlichen Untersuchungen zum prospektiven Gedächtnis (z. B. Burgess, Quayle, & Frith, 2001) die Häufigkeit der Zielreize zwecks reliabler Messung deutlich über den typischen Häufigkeiten in Verhaltensexperimenten liegt (z. B., Marsh et al., 2002). Träfen die von Graf und Uttl (2001) postulierten Unterschiede zu, so könnten die Ergebnisse vieler neurowissenschaftlicher Arbeiten nur für sogenannte Vigilanzaufgaben Geltung haben. In der vorliegenden Studie wurden mögliche qualitative und quantitative Unterschiede zwischen diesen beiden Aufgabentypen mit ereigniskorrelierten Potentialen (EKPs) untersucht. Dabei wurde auf modellbasierte Maße verzichtet, so dass die Studie nicht zum Kern dieser Dissertation zählt. Weil aber die Analyse von Interferenzeffekten auch hier von Interesse war, werden die wichtigsten Ergebnisse im Folgenden kurz dargestellt.

Alle 16 Versuchspersonen dieser Studie bearbeiteten zunächst eine Kontrollbedingung, in der nur lexikalische Entscheidungen erforderlich waren („*control*“), gefolgt von einem Block mit geringer Auftretenswahrscheinlichkeit prospektiver Zielreize („*PM rare*“), und einem Block mit hoher Auftretenswahrscheinlichkeit („*PM frequent*“). Jeder Block beinhaltete 500 Experimentaldurchgänge. Die Reihenfolge der prospektiven Blöcke wurde ausbalanciert. Wie in den Vorgängerstudien sollten sich Versuchspersonen als prospektive Gedächtnisaufgabe daran erinnern, bei bestimmten Anfangsbuchstaben (*G*, *H*, *M*, als Zielereignisse) eine intendierte Handlung auszuführen und die lexikalischen Entscheidungen zu unterbrechen.

Die Analyse der Latenzzeiten der fortlaufenden Aufgabe zeigte deutliche Interferenz durch die prospektive Gedächtnisaufgabe [$F(1, 15) = 74.26$, $\eta_p^2 = .83$, $p < .001$, für den Vergleich der beiden prospektiven Blöcke mit der Kontrollbedingung]. Insbesondere waren auch die Latenzzeiten bei hoher (M = 847ms; SD = 130ms) im Vergleich zu geringer (M = 795ms; SD = 129ms) Auftretenswahrscheinlichkeit der Zielreize erhöht, $F(1, 15) = 9.78$, $\eta_p^2 = .40$, $p < .01$. Damit korrespondierend war die prospektive Gedächtnisleistung (Treffer-Rate) in der *frequent*-Bedingung ($M = 0.89$, $SD = 0.09$) auch deutlich höher als in der *rare*-Bedingung ($M = 0.56$, $SD = 0.19$), $F(1,15) = 9.78$, $p < .01$, $\eta_p^2 = .40$. Insgesamt spricht dieses Muster für eine intensivierte Suche nach relevanten Zielereignissen („monitoring"; Guynn, 2003), wenn diese häufig auftreten, was in Folge zu größeren Kosten in der fortlaufenden Aufgabe führt.

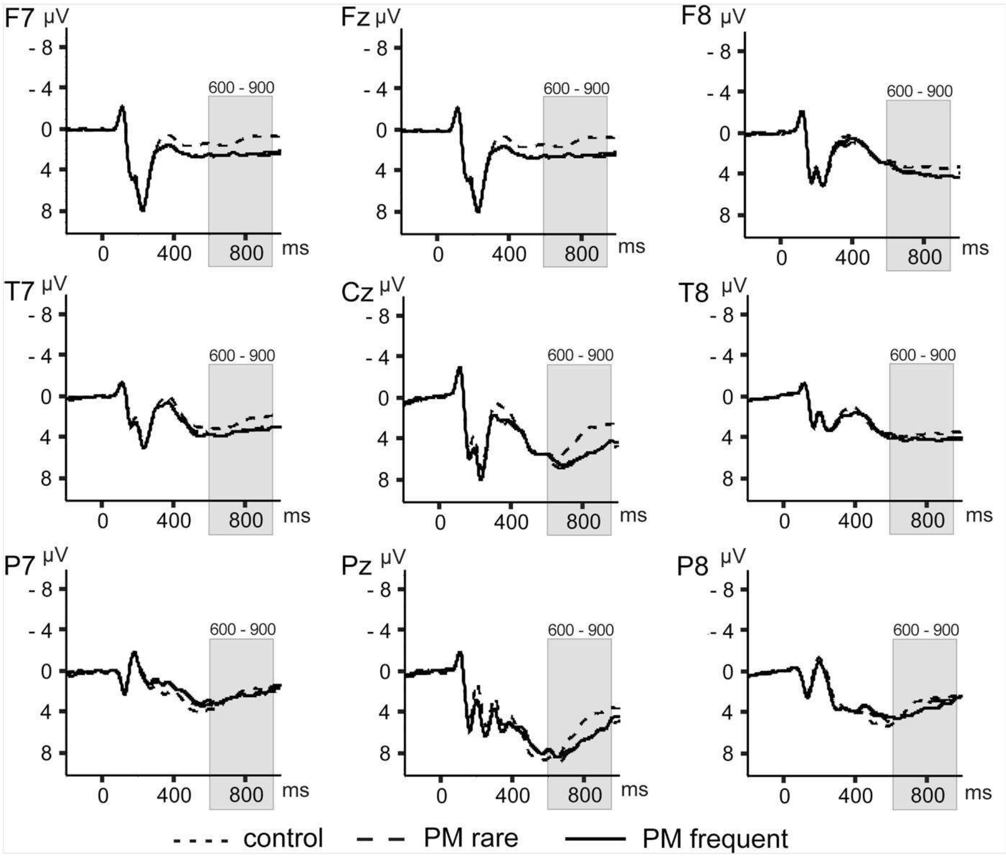


Abbildung 6. EKP-Analyse der fortlaufenden Aufgabe (lexikalische Entscheidungsaufgabe).

Spannungsverlauf (in µV; *y*-Achse) an neun Elektroden als Funktion der Zeit nach Stimuluspräsentation (*x*-Achse; 0 ms: Erscheinen der Buchstabenkette) für den Kontrollblock (gepunktete Linien), die prospektive Bedingung mit selten auftretenden Zielereignissen (*PM rare*; gestrichelte Linien), und häufig auftretenden Zielereignissen (*PM frequent*; durchgezogene Linien). Reliable Unterschiede zur Kontrollbedingung wurden für beide prospektive Bedingungen im Zeitfenster 600-900ms an den Elektroden F7, Fz, T7, and Cz gefunden, und zusätzlich an Pz für die „*PM rare*" Bedingung. Adaptierte Grafik aus „Does frequency matter? ERP and behavioral correlates of monitoring for rare and frequent prospective memory targets." von D. Czernochowski, S. Horn, und U. J. Bayen, 2012, *Neuropsychologia*, *50*, S. 71. © 2012 Elsevier Ltd.

Die Analyse der EKP Daten (siehe Abbildung 6) ergab deutliche Amplitudenunterschiede zwischen der Kontrollbedingung (gepunktete Linien) und den beiden prospektiven Bedingungen, in denen eine relativ späte, anhaltende Positivierung (ca. 600-900 ms nach Präsentationsbeginn der Stimuli) auftrat. Dieser Effekt war fronto-zentral am stärksten ausgeprägt (Elektroden *F7*, *Fz*, *Cz*) und wurde mit einem kontrollierten, kontinuierlichen „Suchmodus“ zur Detektion prospektiver Zielereignisse in Verbindung gebracht (siehe Guynn, 2003; Reynolds, West, & Braver, 2009). Bemerkenswert war, dass in den EKP-Analysen keinerlei Einfluss der Auftretenswahrscheinlichkeit der Zielereignisse festgestellt wurde (vgl. die überlappenden Linien der beiden prospektiven Bedingungen in Abb. 6), im Gegensatz zu den großen Unterschieden in den Latenzzeiten und der prospektiven Gedächtnisleistung. Aus den vorausgehenden Analysen ist bekannt, dass ein wichtiger Anteil der Latenzzeitveränderungen im prospektiven Gedächtnisparadigma durch Verschiebungen im Geschwindigkeits-Genauigkeits-Kriterium (Schwellenparameter *a*) erklärt werden kann. Mehrere Arbeiten haben gezeigt, dass subkortikale Strukturen, insbesondere die Basalganglien, in solche Kriteriumsverschiebungen involviert sind (z. B. Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; Forstmann et al., 2010). Eine mögliche Erklärung wäre folglich, dass die Latenzzeitveränderungen durch Aktivität in Gehirnstrukturen vermittelt wurden, für die EKP-Analysen wenig sensitiv sind.

7.5 Zusammenfassende Diskussion

In dieser Arbeit wurde die Interferenz durch ereignisbasierte prospektive Gedächtnisaufgaben mit dem Diffusionsmodell analysiert (z. B. Ratcliff, 1978; Ratcliff & Tuerlinckx, 2002), um zugrunde liegende kognitive Prozesse zu identifizieren und voneinander zu trennen. Bislang wurden solche Interferenzeffekte eher unspezifisch als Indikator für ressourcenintensive, vorbereitende Aufmerksamkeit betrachtet (z. B. Smith et al., 2007). Ein methodischer Vorteil der Diffusionsmodellanalyse bestand dabei in der simultanen Berücksichtigung der Genauigkeit und Geschwindigkeit (d.h., der gesamten Latenzzeitverteilungen).

Die modellbasierten Auswertungen ergaben über die Experimente hinweg ein relativ homogenes Muster: die Verlangsamung durch eine prospektive Aufgabe ließ sich fast durchweg auf (a) eine vorsichtigere Bearbeitung der fortlaufenden Aufgabe (Erhöhung des Schwellen-parameters *a*) und (b) Verzögerungen in der nicht-entscheidungsbezogenen Zeit für periphere Prozesse (Parameter T_{er}) zurückführen. Effekte auf die Geschwindigkeit der Informationsverarbeitung (Driftrate *v*) während der fortlaufenden Aufgabe wurden in nur wenigen Fällen beobachtet (Studie 1; Studie 3, ältere Erwachsene). Der ausbleibende Effekt auf die Driftraten spricht auf den ersten Blick gegen die Annahme ressourcenintensiver Prozesse, die die lexikalischen Entscheidungen beeinträchtigen. Dies ist umso überraschender, da wir den Interferenzeffekt unter Bedingungen untersuchten, in denen etablierte Theorien zum prospektiven Gedächtnis übereinstimmend ressourcenintensive Kontrollprozesse annehmen (Einstein et al., 2005; Marsh et al., 2003; Smith et al., 2007), die mit der fortlaufenden Aufgabe interferieren.

Wir gehen davon aus, dass sich strategische Prozesse eher auf die Zeitkonstante T_{er} (und nicht etwa die Driftrate) auswirken könnten, weil die Überprüfung einer Buchstabenkette nach einem relevanten Zielmerkmal (z. B. einem bestimmten Anfangsbuchstaben) sequentiell erfolgt (d.h., vor oder nach der eigentlichen lexikalischen Entscheidung). Damit übereinstimmend berichten Scullin et al. (2010), dass Versuchspersonen für die Detektion prospektiver Zielereignisse relativ konsistente Strategien verfolgen: ungefähr die Hälfte der Versuchspersonen suchte zunächst die Buchstabenkette nach Existenz eines relevanten Zielmerkmals ab (gefolgt von einer lexikalischen Entscheidung), während die andere Hälfte zunächst eine lexikalische Entscheidung traf (gefolgt von einer weiteren Suche). Ein solches Vorgehen würde die eigentliche Entscheidungsphase nicht verlangsamen, sich im Diffusionsmodell aber auf die Zeitkonstante T_{er} auswirken und die Antwortzeit insgesamt erhöhen.

Das konsistenteste Modellierungsergebnis in allen Experimenten war die vorsichtigere Bearbeitung der fortlaufenden Aufgabe, wenn zusätzlich prospektives Erinnern erforderlich war (d.h., eine Veränderung im Geschwindigkeits-Genauigkeits-Kriterium *a*). Einstein und McDaniel (2010; vgl. auch Hicks et al., 2005) gehen davon aus, dass Versuchspersonen metakognitive Annahmen bilden (z. B. durch Hinweise zu Beginn eines Experiments), die die folgende Zuteilung von Ressourcen und damit das Ausmaß der Interferenz wesentlich beeinflussen. Aufgrund der vorliegenden Daten kann angenommen werden, dass Versuchspersonen den nachfolgenden Aufgabenkontext generell als komplexer oder schwieriger betrachten, sobald Instruktionen für eine prospektive Gedächtnisaufgabe eingeführt werden, was in Folge zu umsichtigerem Entscheidungsverhalten führt. Übereinstimmend damit wurde postuliert, dass fortlaufende Aufgaben generell in einem anderen „Modus“ bearbeitet werden („prospective retrieval mode“; Guynn, 2003), der nach Einführung einer prospektiven Gedächtnisaufgabe konstant beibehalten wird, um potentielle prospektive Hinweisreize zu entdecken (Guynn, 2008, S. 57).

Eine interessante Implikation der beobachteten Veränderungen im Geschwindigkeits-Genauigkeits-Kriterium ist, dass prospektive Gedächtnisaufgaben – zumindest im Labor – eine *akkuratere* Bearbeitung der fortlaufenden Aktivitäten induzieren. Somit kann nicht von einem reinen *Kosten*effekt (vgl. Smith, 2003) gesprochen werden, da ein Teil der Verlangsamung im Austausch zu einer genaueren Bearbeitung führt. Dieser kontraintuitive Effekt wurde in bisherigen Arbeiten nicht diskutiert, möglicherweise, weil der größte Anteil der Variabilität oft in den Latenzzeiten lag, während die Fehlerraten sehr niedrig waren (z. B. Marsh et al., 2003; Smith et al., 2007). Im Diffusionsmodell hingegen können bereits kleinere Unterschiede in den Fehlerraten größere Veränderungen im Geschwindigkeits-Genauigkeits-Parameter a implizieren, wenn die Aufgabe generell leicht ist (z. B., Fehlerraten < 5%; Ratcliff, Spieler, & McKoon, 2000). Ein Prozessmodell, das diese nichtlineare Beziehung zwischen Geschwindigkeit und Genauigkeit explizit definiert, kann somit zu neue Einsichten führen, warum prospektives Erinnern andere Aktivitäten verlangsamt.

8. Literatur

Baddeley, A. D. (2003). Working Memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4,* 829–839.

Bargh, J., & Chartrand, T. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462–479.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86.

Bisiacchi, P. S., Schiff, S., Ciccola, A., & Kliegel, M. (2009). The role of dual-task and task-switch in prospective memory: Behavioural data and neural correlates. *Neuropsychologia, 47,* 1362–1373.

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends in Neurosciences, 33,* 10–16.

Brandimonte, M. A., Ferrante, D., Feresin, C., & Delbello, R. (2001). Dissociating prospective memory from vigilance processes. *Psicologica, 22*, 97–113.

Bredenkamp, J. & Erdfelder, E. (1996). Methoden der Gedächtnispsychologie. In: Albert, D. & Stapf, K. H. (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich C: Theorie und Forschung*, Serie II: Kognition, Band 4: Gedächtnis (S. 1–94). Göttingen: Hogrefe.

Breneiser, J. E., & McDaniel, M. A. (2006). Discrepancy processes in prospective memory retrieval. *Psychonomic Bulletin & Review, 13,* 837–841.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178.

Burgess, P., Quayle, A., & Frith, C. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia, 39,* 545–555.

Campbell, D., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cohen, A-L., Jaudas, A., & Gollwitzer, P. M. (2008). Number of cues influence the cost of remembering to remember. *Memory & Cognition, 36,* 149–156.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409–422). Amsterdam: Elsevier.

Crovitz, H. F. & Daniel, W. F. (1984). Measurements of everyday memory: Toward the prevention of forgetting. *Bulletin of the Psychonomic Society, 22*, 413–414.

Czernochowski, D., Horn, S. S., & Bayen, U. J. (2012). Does frequency matter? ERP and behavioral correlates of monitoring for rare and frequent prospective memory targets. *Neuropsychologia, 50,* 67–76.

Dembitzer, A. & Lai, E. J. (2003). Images in clinical medicine. Retained surgical instrument. *The New England Journal of Medicine, 348,* 228.

Donkin, C., Brown, S. D., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology, 55,* 140–151.

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 717–726.

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 115–141). Mahwah, NJ: Erlbaum.

Einstein, G. O., & McDaniel, M. A. (2010). Prospective Memory and what costs do not reveal about retrieval processes: A commentary on Smith, Hunt, McVay, and McConnell

(2007). *Journal of Experimental Psychology: Learning Memory & Cognition, 36,* 1082–1088.

Einstein, G. O., McDaniel, M. A., & Scullin, M. K. (2011). Prospective memory and aging: Understanding the variability. In M. Naveh-Benjamin & N. Ohta (Eds.), *Memory and Aging: Current Issues and Future Directions.* Philadelphia, PA: Psychology Press.

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., et al. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General, 134,* 327–342.

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models. A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology, 217,* 108–124.

Forstmann, B. U., Anwander, A., Schafer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., et al. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 15916–15920.

Freud, S. (1901/2009). *Zur Psychopathologie des Alltagslebens*. Frankfurt am Main: Fischer.

Gao, J. L., Cheung, R. T. F., Chan, Y. S., Chu, L. W., & Lee, T. M. C. (2012). Increased prospective memory interference in normal and pathological aging: different roles of motor and verbal processing speed. *Aging, Neuropsychology, and Cognition.* doi:10.1080/13825585.2012.672948

Graf, P., & Uttl, B. (2001). Prospective memory: A new focus for research. *Consciousness and Cognition: An International Journal, 10,* 437–450.

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology, 38,* 245–256.

Guynn, M. J. (2008). Theory of strategic monitoring: Instantiating a retrieval mode and periodic target checking. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 53–76). Mahwah, NJ: Erlbaum.

Harris, J. E. (1984). Remembering to do things: A forgotten topic. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory, actions, and absentmindedness* (pp. 71–92). London: Academic Press.

Hasher, L., & Zacks, R. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hicks, J. L., Marsh, R. L., & Cook, G. I. (2005). Task interference in time-based, event-based, and dual intention prospective memory conditions. *Journal of Memory and Language, 53*, 430–444.

Horn, S. S., Bayen, U. J., & Smith, R. E. (2011). What can the diffusion model tell us about prospective memory? *Canadian Journal of Experimental Psychology, 65,* 69–75.

Horn, S. S., Bayen, U. J., Smith, R. E., & Boywitt, C. D. (2011). The multinomial model of prospective memory: Validity of ongoing-task parameters. *Experimental Psychology, 58,* 247–255.

Horn, S. S., Smith, R. E., Bayen, U. J., & Voss, A. (2008). Analyzing the cost of prospective memory with the diffusion model. *International Journal of Psychology, 43,* 787.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). New York: Academic Press.

Klauer, K. C. (2002). Theory of model testing and selection. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social and Behavioral Sciences* (pp. 9927–9931). Oxford, England: Pergamon.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93,* 353–368.

Kliegel, M., Jäger, T., & Phillips, L. H. (2008). Adult age differences in event-based prospective memory: a meta-analysis on the role of focal versus nonfocal cues. *Psychology and Aging, 23,* 203–208.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2001). Varying the importance of a prospective memory task: Differential effects across time- and event-based prospective memory. *Memory, 9,* 1–11.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2002). Complex prospective memory and executive control of working memory: A process model. *Psychologische Beiträge, 44,* 303–318.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2004). Importance effects on performance in event-based prospective memory tasks. *Memory, 12,* 553–561.

Kliegl, R. (2007). Towards a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, & Reichle (2007). *Journal of Experimental Psychology: General, 138,* 530–537.

Kolmogorov, A. (1941). Confidence limits for an unknown distribution function. *Annals of Mathematical Statistics, 12,* 461–463.

Laming, D. R. J. (1968). *Information theory of choice reaction time.* New York: Wiley.

Lewandowsky, S., & Farrell, S. (2010). *Computational Modeling in Cognition: Principles and Practice.* Thousand Oaks, CA: Sage.

Lewin, K. (1926). Vorsatz, Wille und Bedürfnis. [Intention, will, and need]. *Psychologische Forschung, 7,* 330–385.

Loftus, E. (1971). Memory for intentions: The effect of presence of a cue and interpolated activity. *Psychonomic Science, 23,* 315–316.

Luce, R. D. (1986). *Response times.* New York: Wiley.

Luce, R. D., & Raiffa, H. (1957). *Games and Decisions.* New York: Wiley.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS – a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing, 10,* 325–337.

Macmillan, N., & Creelman, C. (2005). *Detection theory: A user's guide.* Mawah, NJ: Erlbaum.

Marsh, R. L., & Hicks, J. L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 336–349.

Marsh, R. L., Hicks, J. L., Cook, G. I., Hansen, J., & Pallos, A. (2003). Interference to ongoing activities covaries with the characteristics of an event-based intention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 861–870.

Marsh, R. L., Hicks, J. L., & Watson, V. (2002). The dynamics of intention retrieval and coordination of action in event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 652–659.

Martin B. A., Brown N. L., & Hicks J. L. (2011). Ongoing task delays affect prospective memory more powerfully than filler task delays. *Canadian Journal of Experimental Psychology, 65,* 48–56.

Maylor, E. (1990). Age and prospective memory. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 42,* 471–493.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14,* S127–S144.

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field.* Thousand Oaks, CA: Sage.

McDaniel, M. A., Einstein, G. O., & Rendell, P. G. (2008). The Puzzle of inconsistent age-related declines in prospective memory: A multiprocess explanation. In M. Kliegel, M. A. McDaniel, and G. O. Einstein (Eds.), *Prospective Memory: Cognitive, Neuroscience, Developmental, and Applied Perspectives* (pp.141–160). Mahwah, NJ: Erlbaum.

McDaniel, M. A., Guynn, M. J., Einstein, G. O., & Breneiser, J. (2004). Cue-focused and reflexive-associative processes in prospective memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 605–614.

Meacham, J. A. & Leiman, B. (1975, September). *Remembering to perform future actions.* Paper presented at the meeting of the American Psychological Association, Chicago, IL. Also in U. Neisser (Ed.) (1982), *Memory observed: Remembering in natural contexts* (pp. 327–336). San Francisco: Freeman.

Miller, G., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behavior.* New York: Holt.

Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 269–310). Cambridge, MA: MIT Press.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Pachella, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Hillsdale, NJ: Lawrence Erlbaum.

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

Patton, G. W. & Meit, M. (1993). Effect of aging on prospective and incidental memory. *Experimental Aging Research, 19*, 165–176.

Pew, R. W. (1969). The speed-accuracy operating characteristic. *Acta Psychologica, 30,* 16–26.

Pongratz, L.J. (1998). Die Kontroverse zwischen Wilhelm Wundt und Karl Bühler. Analyse einer Wende der Psychologie. In W. Baumgartner, F.-P. Burkhard & F. Wiedman (Eds.), *Zur Entwicklung und Bedeutung der Würzburger Schulen: Band 7. Brentano-Studien* (pp. 255–266). Dettelbach, Germany: Röll.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (2008). The EZ diffusion method: Too EZ? *Psychonomic Bulletin & Review, 15,* 1218–1228.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two choice decision tasks. *Neural Computation, 20,* 873–922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9,* 347–356.

Ratcliff, R., Spieler, D., & McKoon, G. (2000). Explicitly modeling the effects of aging on response time. *Psychonomic Bulletin & Review, 7,* 1–25.

Ratcliff, R., & Starns, J.J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review, 116,* 59–83.

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two choice decisions. *Psychonomic Bulletin & Review, 13,* 626–635.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9,* 438–481.

Reynolds, J. R., West, R., & Braver, T. (2009). Distinct neural circuits support transient and sustained processes in prospective memory and working memory. *Cerebral Cortex, 19,* 1208–1221.

Salthouse, T. A. (1988). Resource-reduction interpretations of cognitive aging. *Developmental Review, 8,* 238–272.

Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Schouten, J. F., & Bekker, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica, 27,* 143–153.

Scullin, M. K., McDaniel, M. A., Shelton, J. T., & Lee, J. H. (2010). Focal/nonfocal cue effects in prospective memory: Monitoring difficulty or different retrieval processes? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 736–749.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114,* 727–741.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience*, 27, 161–168.

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 347–361.

Smith, R. E. (2010). What costs do reveal and moving beyond the cost debate: Reply to Einstein and McDaniel (2010). *Journal of Experimental Psychology: Learning Memory & Cognition, 36,* 1089–1095.

Smith, R. E., & Bayen, U. J. (2004). A multinomial model of event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 756–777.

Smith, R. E., & Bayen, U. J. (2006). The source of adult age differences in event-based prospective memory: A multinomial modeling approach. Journal *of Experimental Psychology: Learning, Memory, and Cognition, 32,* 623–635.

Smith, R. E., Horn, S. S., & Bayen, U. J. (2012). Prospective memory in young and older adults: The effects of ongoing task load. *Aging, Neuropsychology, and Cognition, 19,* 495–514.

Smith, R. E., Hunt, R. R., McVay, J. C., & McConnell, M. D. (2007). The cost of event-based prospective memory: Salient target events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 734–746.

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 101–117.

Squire, L. R. (2004). Memory Systems of the Brain: A brief history and current perspective. *Neurobiology of Learning & Memory, 82,* 171–177.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika, 2*5, 251–260.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review, 14,* 1011–1026.

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods, 40,* 61–72.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods, 16,* 44–62.

van Ravenzwaaij, D., & Oberauer, K. (2009). How to Use the Diffusion Model: Monte–Carlo Simulations Evaluating Parameter Recovery of Three Methods: EZ, Fast-DM, and DMAT. *Journal of Mathematical Psychology, 53,* 463–473.

Verhaeghen, P. (2000). The parallels in beauty's brow: Time-accuracy functions and their implications for cognitive aging theories. In T. J. Perfect and E. A. Maylor (Eds.), *Models of cognitive aging* (pp.50–86). Oxford: University Press.

Voss, A., Rothermund, K., Gast, A., & Wentura, D. (in press). Cognitive Processes in Categorical and Associative Priming: A Diffusion Model Analysis. *Journal of Experimental Psychology: General.*

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32,* 1206–1220.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39,* 767–775.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52,* 1–9.

Wagenmakers, E. J., Krypotos, A. M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition, 40,* 145–160.

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58,* 140–159.

Wagenmakers, E. J., van der Maas, H. L. J., & Grasman, R. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14,* 3–22.

Wagenmakers, E. J., van Der Maas, H. L. J., Dolan, C. V., & Grasman, R. (2008). EZ does it! extensions of the EZ-diffusion model. *Psychonomic Bulletin & Review, 15,* 1229–1235.

West, R., Krompinger, J., & Bowry, R. (2005). Disruptions of preparatory attention contribute to failures of prospective memory. *Psychonomic Bulletin & Review, 12,* 502–507.

Whittlesea, B., & Williams, L. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 547–565.

Zeigarnik, B. (1927). Das Behalten erledigter und unerledigter Handlungen. [Remembering of finished and unfinished activities]. *Psychologische Forschung, 9,* 1–85.

Anhang

Studie 1:

Horn, S. S., Bayen, U. J., & Smith, R. E. (2011). What can the diffusion model tell us about prospective memory? *Canadian Journal of Experimental Psychology, 65,* 69–75.

Studie 2:

Horn, S. S., & Bayen, U. J., & Smith, R. E. (2011). Interference From Remembering to Remember: A Diffusion Model Account. *Manuscript submitted for publication.*

Studie 3:

Horn, S. S., Bayen, U. J., & Smith, R. E. (2012). Adult Age Differences in Interference From a Prospective-Memory Task: A Diffusion-Model Analysis. *Manuscript submitted for publication.*

Studie 4:

Czernochowski, D., Horn, S. S., & Bayen, U. J. (2012). Does frequency matter? ERP and behavioral correlates of monitoring for rare and frequent prospective memory targets. *Neuropsychologia, 50,* 67–76.

Erklärung:

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, 15. 10. 2012

Sebastian Horn

Canadian Journal of Experimental Psychology
2011, Vol. 65, No. 1, 69–75

1196-1961/11/$12.00 DOI: 10.1037/a0022808

What Can the Diffusion Model Tell Us About Prospective Memory?

Sebastian S. Horn and Ute J. Bayen
Heinrich-Heine-Universität Düsseldorf

Rebekah E. Smith
The University of Texas at San Antonio

Cognitive process models, such as Ratcliff's (1978) diffusion model, are useful tools for examining cost or interference effects in event-based prospective memory (PM). The diffusion model includes several parameters that provide insight into how and why ongoing-task performance may be affected by a PM task and is ideally suited to analyse performance because both reaction time and accuracy are taken into account. Separate analyses of these measures can easily yield misleading interpretations in cases of speed–accuracy trade-offs. The diffusion model allows us to measure possible criterion shifts and is thus an important methodological improvement over standard analyses. Performance in an ongoing lexical-decision task was analysed with the diffusion model. The results suggest that criterion shifts play an important role when a PM task is added, but do not fully explain the cost effect on reaction time.

Keywords: prospective memory, diffusion model, monitoring and attentional resources, response time models

Event-based prospective memory (PM) tasks involve remembering to perform intended actions after a delay, when a specific target event occurs. Such tasks often occur in the midst of other activities that must be interrupted to perform the intended action. To capture this aspect of real world PM, the PM task is embedded in an ongoing task in the typical laboratory paradigm (Einstein & McDaniel, 1990). For instance, participants may be busily engaged in an ongoing lexical-decision task, and at the same time must remember to interrupt their decisions to carry out another action (i.e., press a certain key on a computer keyboard) when a particular target occurs (e.g., the word *tiger* appears on the screen).

Much theoretical and experimental work on PM has focused on the processes involved in retrieving such intentions (e.g., Einstein & McDaniel, 1996; Smith, Hunt, McVay, & McConnell, 2007; West, 2007). Particularly, researchers have examined whether successful PM always requires resource-demanding preparatory attentional processes (Smith, 2003, 2008, 2010), or whether spontaneous retrieval of the intention occurs under specific circumstances (McDaniel & Einstein, 2000, 2007). The empirical approach toward addressing this question rests on the analysis of ongoing-task performance in the presence versus absence of a PM task. The *cost* or *interference effect* of PM refers to the finding that reaction time (RT) on non-PM-target trials in the ongoing task can be increased by the need to remember the PM task, and can covary with PM performance (Smith, 2003). It is assumed that RT can increase as PM absorbs attentional resources that would otherwise be devoted to the ongoing task. In the last decade, numerous studies have examined cost effects (see Smith et al., 2007, for an overview) and their relationship with characteristics of the PM targets, such as salience and focality (Einstein et al., 2005), individuals' resource allocation (Marsh, Hicks, & Cook, 2005), and potential boundary conditions to demonstrations of cost effects (Cohen, Jaudas, & Gollwitzer, 2008; but see Smith, 2010). However, to date surprisingly little is known about the specific processes that lead to the slowing when cost effects occur. *Why* and *how* does processing change in the ongoing task with an additional requirement to remember an intention?

In this article, we argue that cognitive process models, such as the diffusion model (e.g., Ratcliff, 1978), are useful tools for addressing these questions through the measurement of latent variables assumed to underlie performance in ongoing tasks. We will first describe the diffusion model in more detail and point out the importance of considering speed–accuracy trade-offs in task performance. We will then present a model-based reanalysis of data from Smith (2003, Experiment 1) to demonstrate how additional insight into cost effects can be gained.

The Diffusion Model

In cognitive psychology, the diffusion model has been successfully applied to a variety of paradigms in which individuals make simple and fast two-choice decisions, with mean latencies not much over 1 to 1.5 s. Previous applications of the model have included, among others, lexical decisions (Ratcliff, Gomez, & McKoon, 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008), animacy categorisation (Spaniol, Madden, & Voss, 2006), recognition memory (Criss, 2010; Ratcliff, 1978), practice effects (Dutilh, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2009), and the identification of factors underlying age-related slowing (e.g., Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2006). See Ratcliff

Sebastian S. Horn and Ute J. Bayen, Institut für Experimentelle Psychologie, Heinrich-Heine-Universität Düsseldorf; Rebekah E. Smith, Department of Psychology, The University of Texas at San Antonio.

Parts of this project were presented at the 29th International Congress of Psychology, Berlin, Germany, July 2008, and at the 3rd International Conference on Prospective Memory, Vancouver, Canada, July 2010. The research was supported in part by Grant AG034965 from the National Institute on Aging (RES).

Correspondence concerning this article should be addressed to Sebastian S. Horn, Institut für Experimentelle Psychologie, Abteilung Mathematische und Kognitive Psychologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, Geb. 23.02, 40225 Düsseldorf, Germany. E-mail: s.horn@uni-duesseldorf.de

and McKoon (2008) and Wagenmakers (2009), for reviews of the model, Gardiner (2004) and Smith (2000) for mathematical foundations, and Ratcliff and Tuerlinckx (2002), and Voss and Voss (2008) for different methods of parameter estimation. In general, the diffusion model provided a close fit to the observed RT distributions and response accuracy in most applications. Many ongoing tasks routinely used in PM research could provide appropriate data; the current experiment used a lexical-decision task, which has been studied in detail with the diffusion model (Ratcliff, Gomez, et al., 2004; Wagenmakers et al., 2008), and which is a frequently used ongoing task in PM research (e.g., Marsh et al., 2005).

A basic assumption of Ratcliff's diffusion model is that two-choice decisions are based on continuously accumulating information, starting from a value z on a decision-related strength-of-evidence axis (y-axis in Figure 1). A diffusion process moves from this starting point z over time (x-axis in Figure 1) until one of two thresholds, associated with a decision ("A" vs. "B"), is reached. The average speed of information uptake (i.e., the ratio of accumulated evidence per time unit) is a systematic influence (drift rate parameter v) that drives the process to one of the thresholds. A positive slope of v (as shown in Figure 1) indicates that relatively more evidence is collected for the upper ("A") than the lower ("B") threshold, whereas negative slopes of v imply the opposite. Drift rate determines processing efficiency in the actual decision phase, with high absolute values predicting both fast and accurate decisions. However, the accumulation process within a trial is also affected by random noise[1], implying that trials with the same drift rate do not always terminate at the same time or threshold (thereby producing RT distributions and errors, respectively; see the different process tracks in Figure 1). Parameter a represents the distance between the two decision thresholds (the value of the lower threshold is set to 0) and quantifies the amount of evidence that is

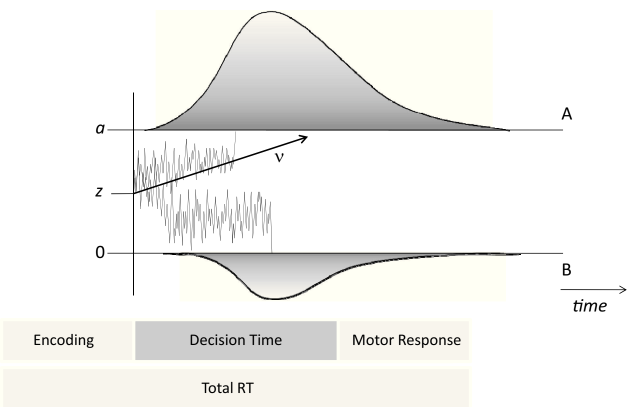


Figure 1. Illustration of the diffusion model. The vertical axis is the decision axis, and the horizontal axis is the time axis. Diffusion processes start at point z and move over time until the upper threshold (positioned at a) or the lower threshold (positioned at zero) is reached. Reaction time distributions for decisions associated with the upper and the lower thresholds are shown. The two sample process tracks illustrate that different thresholds can be reached with the same (positive) drift rate because of random influence. Total RT is the sum of the decision time and a nondecision component that represents the duration of processes such as perceptual encoding and motor response execution.

required until a decision is made. This parameter reflects the decision maker's speed-accuracy criterion. Smaller values of parameter a predict faster RTs and more errors (a liberal speed-accuracy setting), whereas higher values predict slower RTs and fewer errors (a conservative speed-accuracy setting). The relation of the starting point z to the thresholds is an indicator of response bias. If a decision maker does not favour one response over the other, the starting point position is equidistant from both decision thresholds, $z = a/2$, which holds for many experiments (Wagenmakers, van der Maas, & Grasman, 2007). Values larger or smaller than $a/2$ indicate bias in favour of the upper or lower threshold, respectively. The diffusion model assumes that observed RT can be split into a decision phase (described above) and a nondecision component, T_{er}, which includes the processes *before* and *after* the actual decision (e.g., stimulus encoding and motor response execution), and which is added to the diffusion exit time.

The complete version of the diffusion model allows for variation of parameters across trials of an experiment. Drift rate is assumed to vary normally around mean v with standard deviation η to capture trial-by-trial fluctuations in stimulus features or alertness. The starting point z and the nondecision component T_{er} are assumed to be uniformly distributed with width s_z and s_t, respectively. With intertrial variability of drift rate η, the model predicts slower RTs for error responses than for correct responses, and variability of the starting point (s_z) predicts the reverse pattern (see Ratcliff & Rouder, 1998). All parameters of the diffusion model are summarised in Table 1 with a description.

Speed-Accuracy Tradeoffs

In research on PM, analyses of ongoing-task performance are typically conducted to assess the resource demands of PM (McDaniel & Einstein, 2007), and inference mainly concerns mean RT of correct responses or mean accuracy. It is well-known, however, that both measures are in a tradeoff relationship (e.g., Pachella, 1974; Schouten & Bekker, 1967). That is, participants can increase their accuracy at the expense of slower responding, or vice versa. Such speed–accuracy trade-offs imply that RT can be slower not because a task is more difficult or more resource-demanding, but because participants adopt a different criterion when they weigh the importance of speed versus accuracy. Consider the example of ongoing-task performance in the upper half of Table 2 (cf. Wagenmakers et al., 2007). Participants are faster in Condition A than in Condition B, but they also commit more errors. Thus, it is possible that both conditions are equally difficult (or resource-demanding), but participants in Condition B sacrifice speed for higher accuracy. Of course, it is also possible that fewer processing resources are available in Condition B than in A, or vice versa. With observed mean RT and accuracy, we cannot disentangle these possibilities. Finally, a comparison of Conditions B and C reveals that participants in the latter respond more slowly, whereas accuracy is identical. Many researchers would assume that Condition C is more difficult (or resource-demanding) than Condition B, as speed–accuracy trade-offs alone cannot provide an explanation. However, it

[1] Within-trial Gaussian noise with variance s^2 is a scaling parameter in the model. This diffusion coefficient is set to 1 for all present model fits. Different values for s^2 would simply rescale the absolute values of the other parameters, but would not change their relations.

Table 1
Diffusion Model Parameters

Parameter	Description	No-PM condition	PM condition	$t(93)$[a]
z	Mean starting point	$a/2$	$a/2$	—
a	Threshold separation	1.91 (0.08)	2.22 (0.06)	3.15**
v	Mean drift rate	3.03 (0.20)	2.36 (0.07)	−3.90**
T_{er}	Mean nondecision time	0.41 (0.01)	0.43 (0.01)	0.94
s_z	Range of starting point	0.73 (0.09)	0.72 (0.04)	−0.12
s_t	Range of nondecision time	0.04 (0.01)	0.04 (0.01)	−0.13
η	Drift rate variability	0.51 (0.13)	0.50 (0.05)	−0.10
KS-distance[b]		0.02 (0.001)	0.02 (0.001)	−1.33
p-value[c]		.90 (0.02)	.93 (0.02)	0.80

Note. Best-fitting parameter values are averaged across participants in the No-PM condition and in the PM condition. Standard errors are in parentheses. KS = Kolmogorov-Smirnov.
[a] PM condition versus No-PM condition (independent-samples test). [b] KS-distance T between predicted and empirical cumulative distribution functions of response time. [c] Probability values of the KS-tests larger than .05 indicate no significant deviations of predicted from empirical distributions.
** $p < .01$.

is not trivial to go beyond ordinal conclusions and *quantify* which of the underlying processes are responsible for a slowing in Condition C.

Together, mean RT and accuracy provide important information about task performance, but they should not be considered in isolation; this problem is obviously relevant when PM researchers compare ongoing-task performance in the presence versus absence of PM, or when they compare the impact of different PM conditions (e.g., focal/nonfocal PM tasks, number of PM target cues, etc.). In most previous studies that examined the cost of PM, the dependent variable of interest was *either* RT *or* accuracy in the ongoing task; most of the variance in performance typically appeared in one of these measures, but not in both (e.g., in lexical decisions and many other ongoing tasks, differences in mean accuracy are uncommon if aggregated over participants; see Marsh et al., 2005, Experiment 3, or Smith & Bayen, 2006, Experiment 1, for rare exceptions). In addition, analyses of RTs and accuracy were always performed separately, lacking a combined index of performance. It is therefore not surprising that speed–accuracy trade-offs in the PM paradigm have not been considered thus far although they could play a role, indeed.

Table 2
Performance and Corresponding Parameter Values in Three Hypothetical Ongoing-Task Conditions

	Condition		
	A	B	C
Mean performance			
RT	591	690	740
Proportion correct	.881	.953	.953
Diffusion model parameters			
Drift rate v	2.0	2.0	2.0
Threshold distance a	1.0	1.5	1.5
Nondecision time T_{er}	0.4	0.35	0.4

Note. The conditions differ in nondecision time and threshold distance, but not in the rate of information uptake (simulated data). For this example, $z = a/2$, and the variability parameters η, s_z, s_t were fixed to 0. RT = reaction time in milliseconds.

The diffusion model (e.g., Ratcliff, 1978) can address this issue, because it allows us to separate criterion shifts from other processing components involved in two-choice (ongoing) tasks (see Brown & Heathcote, 2008, or Usher & McClelland, 2001, for alternative models). Consider the bottom half of Table 2 with the corresponding diffusion model parameters for the above example. The assessment at the level of latent parameters is theoretically more informative and reveals that differences in nondecision time (T_{er}) and speed-accuracy criterion (a), but not in the rate of information uptake (v), explain the differences in observed performance. That is, participants are slower in Condition C than in Condition B, because they have a longer nondecision time (slower stimulus encoding or response execution), but their processing speed during the actual ongoing-task decisions is not affected. Meaningful psychological interpretations of this type could not be derived from standard analyses; we therefore argue that the diffusion model is a useful tool for the PM paradigm and provide an application in the following section (cf. Horn, Smith, Bayen, & Voss, 2008).

Experiment

Method

Design and materials. We analysed the data from an experiment by Smith (2003) with the diffusion model. The objective of this experiment was to examine the impact of PM on ongoing-task performance. Each participant studied six PM target words, which subsequently occurred twice in 504 trials of an ongoing lexical-decision task. In the PM condition ($n = 62$), participants were told to remember to press the *F1* key when they saw a target word during the experiment. In the No-PM condition ($n = 33$), participants received the same instructions, but were additionally told that they did not have to remember this intention until after the completion of the lexical decision phase (i.e., their intention was not linked to the context of this phase).

Stimuli presented during the lexical-decision task were letter strings that included 126 medium-frequency words (Kučera & Francis, 1967; 6 targets, 120 filler words), and 126 nonwords, created by moving the first syllable of each word to the end. The strings appeared in random order and were repeated in a different random order in the second half of the experiment. See Smith (2003, Experiment 1) for further details concerning the procedure.

Results

The results are reported in two ways. In the first section, we reanalyzed the behavioural data of all non-PM-target trials (accuracy and trimmed RTs). Smith's (2003) original study focused on a selected number of control items in the ongoing task, but we used more trials because a diffusion model analysis requires sufficient observations for robust parameter estimation. The subsequent sections include the corresponding modelling results.

For all analyses of ongoing-task data, we considered non-PM-target trials exclusively, and RTs smaller than 300 ms or greater than 3,000 ms were eliminated (1.27% of all nontarget trials), because they can lead to degenerate parameter estimates (e.g., Ratcliff & Tuerlinckx, 2002).[2]

Ongoing-task performance. Smith (2003, Experiment 1) reported a cost on ongoing-task performance as a consequence of the embedded PM task. The present analysis replicated this finding, indicating that participants' RTs in the PM Condition ($M = 925$, $SD = 163$) were significantly slower than in the No-PM Condition ($M = 747$, $SD = 111$), $t(93) = 6.26$, $p < .001$ (unequal-variances test, $\varepsilon = .94$). The accuracy of lexical decisions did not differ between the PM condition ($M = .972$, $SD = .019$) and the No-PM condition ($M = .977$, $SD = .015$), $t(93) = 1.12$, $p = .27$. Aggregated over participants, accuracy was not far from ceiling, and most of the variance in ongoing-task performance between the conditions appeared in RT. However, we found indications for speed–accuracy trade-offs at the interindividual level in the course of the present reanalysis. There was a positive relationship between participants' speed and their accuracy in the ongoing lexical-decision task; this was the case for the PM condition, $r(62) = .27$, $p = .032$, and the No-PM condition, $r(33) = .55$, $p < .001$. Thus, slower individuals tended to be more accurate within both groups, and it is important to extract such effects with an explicit model *before* conclusions about actual performance are derived.

Prospective memory. PM performance ($M = .69$, $SD = .20$) was positively correlated with ongoing-task RT in the PM condition, $r(62) = .34$, $p = .007$. This finding has been interpreted as competition for limited resources devoted to lexical decisions versus the PM task (Smith, 2003). Furthermore, there was a positive relationship between ongoing-task accuracy and PM performance, $r(62) = .59$, $p < .001$.

Model fit. The present diffusion model analysis rests on the Kolmogorov–Smirnov statistic (KS; Kolmogorov, 1941) for parameter estimation. This approach provides a robust and fine-grained alternative to other estimation procedures (e.g., maximum likelihood, chi square; see Ratcliff & Tuerlinckx, 2002, for an overview), because the whole RT distributions are used as input, and binning into RT quantiles is not required. The test statistic T represents the maximum vertical distance between the predicted and the empirical cumulative distribution function of RT, and the parameter values are determined in such a way that T is minimised (cf. Voss & Voss, 2007, 2008).

Goodness of fit was evaluated on the individual and on the aggregate level. First, we computed KS-tests for each model. None of 95 model tests detected significant deviations from empirical data (the averaged T statistics and p values are in Table 1). To visualize individual model fit, empirical and predicted RT quantiles (.3, .5, and .7) of correct responses and the proportion of errors are plotted in Figure 2 for each participant. A diagonal line (with a slope of $+1$) would represent perfect correspondence between empirical and predicted values in each panel. As can be seen, the correspondence is generally high and the residuals do not indicate systematic biases in the model predictions. The diffusion model adequately reproduces the effects on individual RTs and accuracy in the PM paradigm.

Second, the complete empirical and the complete predicted RT distributions can be visualized to assess model fit. In Figure 3, the RTs from all participants in a condition are combined in a single cumulative distribution function, including correct and error responses. For this purpose, error responses are multiplied by -1, and the corresponding distribution of error RTs is mirrored at the zero point of the time axis (Voss, Rothermund, & Voss, 2004). Thus, the portion of the distribution function on the negative side of the x-axis represents the error latencies and the portion on the positive side of the x-axis represents the latencies for correct responses. The distribution functions intercept the y-axis at the error rate.

As seen in Figure 3, the error rates in the present lexical-decision task are very low, and there is a marked difference in the whole RT distribution between the PM and the No-PM condition. In summary, Figure 3 indicates an excellent fit at the aggregate level, as there are almost no deviations between the predicted and empirical RT distributions.

Diffusion model analysis. We fit the diffusion model separately to the data from each participant. The resulting mean parameter estimates for the PM and for the No-PM condition are listed in Table 1. Because some participants made few errors in the lexical-decision task and lexical variables were not the focus of the present analysis, we aggregated responses over stimulus types for the sake of robust parameter estimation.[3] That is, responses were coded as *correct* (upper threshold) versus *false* (lower threshold) to obtain a single measure of drift rate, thereby aggregating over the two string types (words, nonwords). With this coding, the mean starting point z was fixed to $a/2$ (e.g., Ratcliff, 2002), because biases (e.g., a bias to respond "nonword") cancel out if both string types occur equally often. That is, participants favour correct and false responses equally often and the expected value of the starting

[2] As suggested by a reviewer, we systematically examined further cutoff-criteria for the lower tail (200, 225, 250, 275, 325, 350 ms) and the upper tail (2000, 2250, 2500, 2750, 3250, 3500 ms) of the RT distributions and fit the diffusion models to the so-trimmed data. The pattern of significant and nonsignificant effects was consistently the same as that reported in the present results section.

[3] To examine whether biases in favor of a stimulus type were present, we compared the relative speed of correct versus error responses (cf. Wagenmakers et al., 2007). An a priori bias for a particular stimulus type (e.g., nonwords) would imply faster correct than error responses for this stimulus type, and the reverse pattern of RT for the other type (e.g., words). To obtain reliable estimates of mean error-RT, participants with more than three errors on each stimulus type were included in these analyses. No significant disordinal interactions emerged in the 2 (response correctness) $\times$ 2 (stimulus type) ANOVAs, suggesting that biases in favor of a particular stimulus type were neither present in the PM nor the No-PM condition (largest $F < 2.91$, smallest $p > .10$).

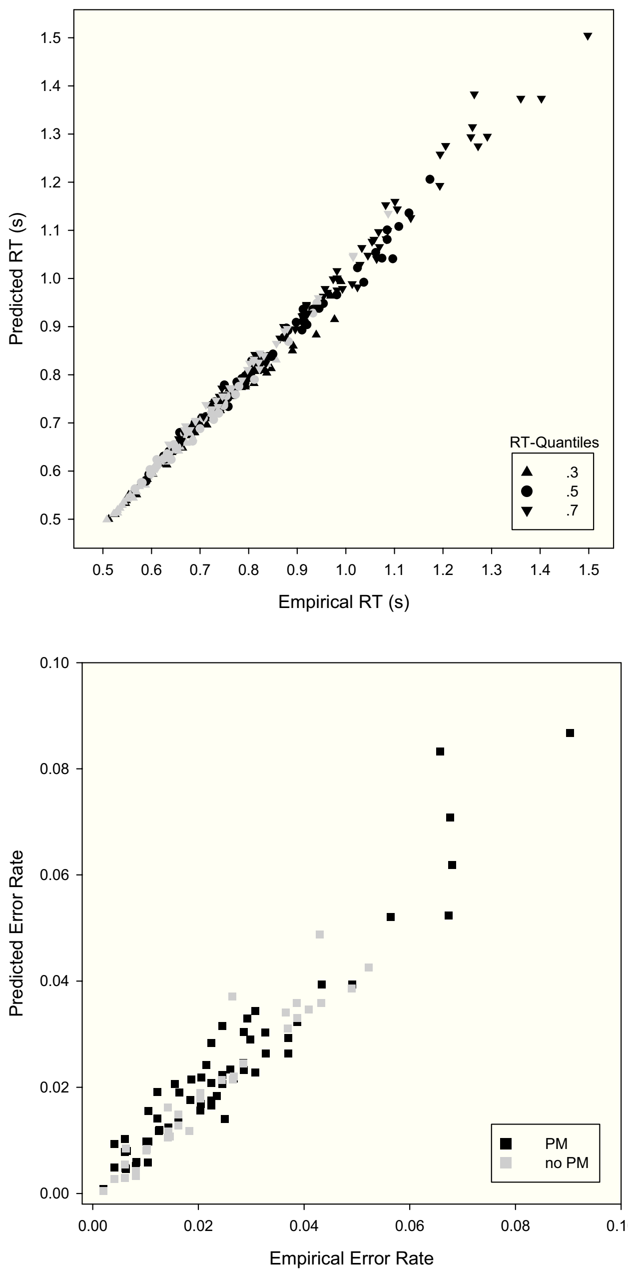


Figure 2. Individual fits of the diffusion model. The upper panel shows empirical (observed) and predicted RT quantiles (.3, .5, and .7) for each participant in the PM condition (black) and in the No-PM condition (grey). The lower panel shows the correspondence of the proportions of error responses. Diagonal lines with a slope of +1 would indicate perfect fit.

point across trials equals $a/2$. Biases can contribute to starting point variability s_z, however (see Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007, for this approach).

The statistical comparison of the diffusion model parameters that yielded the best model fit for participants in the PM and the No-PM conditions revealed a significantly higher threshold parameter a (speed-accuracy setting) when there was an additional PM task (see Table 1). On average, individuals in this group responded more cautiously and required relatively more information before they terminated the decision processes in the ongoing task. However, the model-based analyses also indicated that criterion shifts alone do not fully explain the cost effect of PM on RT. For the present ongoing task, the rate of information accumulation v (collapsed across words and nonwords) was lower for the condition with the embedded PM task. Holding an intention may have interfered with the efficiency of processing the letter strings, as captured by this parameter. Finally, nondecision time T_{er} and the variability parameters s_t, s_z, and η had comparatively little effect on the cost of PM in this study. In general, the variability parameters have smaller impact on overall mean RT or accuracy, and are more important for the relative speed of correct and false responses, which was not the focus of this study.

Discussion

The diffusion model analysis demonstrates that the cost of PM on an ongoing task (as reported by Smith, 2003) can be split into separate processing components that are theoretically informative for PM researchers. More specifically, the addition of a PM task induces more cautious speed-accuracy settings, thereby increasing the latencies of the ongoing lexical decisions. It is typically assumed that such settings are stable across an experimental block, (e.g., Ratcliff, 1978; Wagenmakers et al., 2008; the diffusion model does not assume variability in threshold distance a), and are formed at the outset (e.g., when participants receive instructions at the beginning of an ongoing task). Differences in criterion settings may therefore reflect participants' metacognitive beliefs about the upcoming task, which is perceived as more complex or more demanding with additional PM instructions. This is in line with research on strategic allocation policies and their impact on cost effects (Hicks, Marsh, & Cook, 2005; Marsh et al., 2005). The

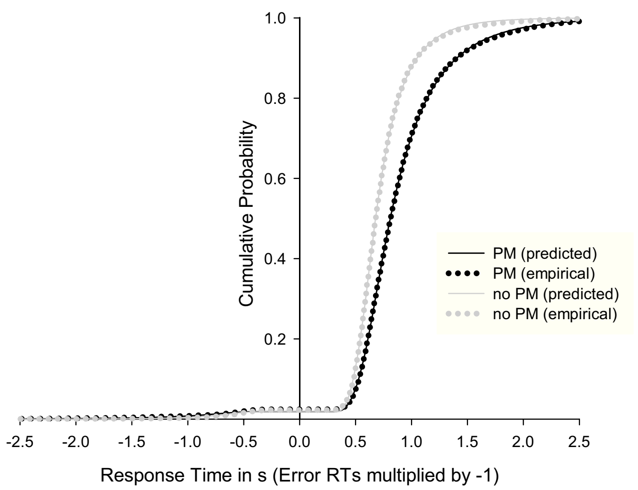


Figure 3. Distributions of response time. The graphs show the empirical (observed) and the model's predicted cumulative distribution functions of response time, aggregated over individuals in the PM condition (black) and No-PM condition (grey), respectively. Negative values on the horizontal axis are latencies of error responses (multiplied by −1), and positive values are latencies of correct responses.

current results indicate that the model's *a* parameter could be a useful candidate for quantifying the effects of metacognitive beliefs. Shifts of the boundaries are also consistent with Smith's (2003, 2010) proposal that embedding a PM task in an ongoing task changes the fundamental nature of the ongoing-task context. The demonstration of a criterion shift in the current study is an important first step in the application of the diffusion model for understanding the processes that contribute to cost effects.

The present results also suggest that the speed of information uptake *v* is decreased in the PM condition, likely because PM absorbs resources that would otherwise be devoted to the ongoing task, thereby slowing processing efficiency (cf. Smith, 2008, 2010). This interference effect occurs during the actual decision process in the ongoing task, when information about the stimulus is accumulated. The current analysis provides a foundation for future research examining whether other components can contribute to cost effects—dependent on characteristics of the PM task—and how changes in the model parameters relate to PM performance. For instance, our continuing work with the diffusion model indicates that nondecision time (and not drift rates) can account for a slowing if participants adopt a more controlled strategy in PM tasks (e.g., "monitoring" for targets in nonfocal PM tasks or even more vigilance-like tasks; cf. Graf & Uttl, 2001). Such effects are expected from other studies, as individuals have been shown to sequentially check for the presence of targets before or after their actual ongoing-task decisions (Scullin, McDaniel, Shelton, & Lee, 2010), thereby producing a slowing. Moreover, motor response coordination and task-set switching, which would map onto nondecision time (cf. Klauer et al., 2007), may become more prominent in such situations.

The observed parameter differences between the conditions may be due to the retrospective component (remembering the intention content; i.e., the targets and associated actions), the prospective component (remembering *that* an intention needs to be realised), or both (cf. Einstein & McDaniel, 1990). In the present study, all participants initially learned the PM target words to criterion and recall in a posttest-questionnaire was similar in both groups and generally high (ca. 85%; cf. Smith, 2003). Thus, substantial RT-variance between the conditions is likely attributable to the impact of the prospective component, and not an effect of differences in retrospective-memory load. In light of similarities between PM- and dual-task paradigms, note that participants did not continuously respond to a frequently occurring secondary task and their PM performance was far from ceiling. These characteristics may support the view that part of the slowing reflects intention-retrieval processes, and not a load from any secondary task (see Graf & Uttl, 2001, for a conceptual framework that distinguishes "PM proper" from "vigilance" tasks).

In conclusion, criterion shifts (parameter *a*) play an important role when a PM task is added to an ongoing task, but they do not fully explain the increase in RT, as other processing components also contribute to the cost of PM. The model-based approach is more informative than standard cost analyses, because parameters quantify the processes that are of theoretical interest at the latent, individual level. This approach can therefore fruitfully contribute to the debate about what costs can or cannot reveal about PM retrieval (Einstein & McDaniel, 2010; Smith, 2010).

Résumé

Les modèles de traitement cognitif, comme le modèle de diffusion de Ratcliff (1978), sont des outils utiles pour examiner les effets de coût ou d'interférence en mèmoire prospective (MP) basèe sur les èvènements. Le modèle de diffusion inclut plusieurs paramètres permettant de se pencher sur le comment et le pourquoi de l'influence de la tâche de MP sur une tâche en cours et est idéal pour analyser la performance car les temps de réaction et la précision sont tous deux considérés. Des analyses séparées de ces mesures peuvent facilement mener à des interprétations erronées dans les cas de compromis vitesse-précision. Le modèle de diffusion nous permet de mesurer les déplacements possibles du critère et constitue donc une importante amélioration méthodologique par rapport aux analyses traditionnelles. La performance dans une tâche de décision lexicale en cours a été analysée avec le modèle de diffusion. Les résultats suggèrent que les déplacements du critère jouent un rôle important lorsqu'une tâche de MP est ajoutée mais n'expliquent pas complètement le coût sur le temps de réaction.

Mots-clés : mémoire prospective, modèle de diffusion, surveillance et ressources attentionnelles, modèles de temps de réponses

References

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178. doi:10.1016/j.cogpsych.2007.12.002

Cohen, A.-L., Jaudas, A., & Gollwitzer, P. M. (2008). Number of cues influence the cost of remembering to remember. *Memory & Cognition, 36,* 149–156.

Criss, A. H. (2010). Differentiation and response bias in episodic memory: Evidence from reaction time distributions. *Journal of Experimental Psychology: Learning Memory & Cognition, 36,* 484–499. doi:10.1037/a0018435

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E. J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review, 16,* 1026–1036. doi:10.3758/16.6.1026

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 717–726.

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 115–141). Mahwah, NJ: Erlbaum.

Einstein, G. O., & McDaniel, M. A. (2010). Prospective Memory and what costs do not reveal about retrieval processes: A commentary on Smith, Hunt, McVay, and McConnell (2007). *Journal of Experimental Psychology: Learning Memory & Cognition, 36,* 1082–1088. doi:10.1037/a0019184

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General, 134,* 327–342. doi:10.1037/0096-3445.134.3.327

Gardiner, C. W. (2004). *Handbook of stochastic methods.* Berlin, Germany: Springer.

Graf, P., & Uttl, B. (2001). Prospective memory: A new focus for research. *Consciousness and Cognition, 10,* 437–450. doi:10.1006/ccog.2001.0504

Hicks, J. L., Marsh, R. L., & Cook, G. I. (2005). Task interference in time-based, event-based, and dual intention prospective memory conditions. *Journal of Memory and Language, 53,* 430–444. doi:10.1016/j.jml.2005.04.001

Horn, S. S., Smith, R. E., Bayen, U. J., & Voss, A. (2008). Analyzing the cost of prospective memory with the diffusion model. *International Journal of Psychology, 43,* 787.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93,* 353–368. doi:10.1037/0022–3514.93.3.353

Kolmogorov, A. (1941). Confidence limits for an unknown distribution function. *Annals of Mathematical Statistics, 12,* 461–463.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. *Journal of Experimental Psychology: Learning Memory & Cognition, 31,* 68–75. doi:10.1037/0278–7393.31.1.68

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14,* S127–S144.

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field.* Thousand Oaks, CA: Sage.

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In H. B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Potomac, MD: Erlbaum.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review, 9,* 278–291.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182. doi:10.1037/0033-295X.111.1.159

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9,* 347–356.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19,* 278–289. doi:10.1037/0882–7974.19.2.278

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin & Review, 13,* 626–635.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9,* 438–481.

Schouten, J. F., & Bekker, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica, 27,* 143–153.

Scullin, M. K., McDaniel, M. A., Shelton, J. T., & Lee, J. H. (2010). Focal/Nonfocal cue effects in prospective memory: Monitoring difficulty or different retrieval processes? *Journal of Experimental Psychology: Learning Memory & Cognition, 36,* 736–749. doi:10.1037/a0018971

Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology, 44,* 408–463. doi:10.1006/jmps.1999.1260

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 347–361. doi:10.1037/0278–7393.29.3.347

Smith, R. E. (2008). Connecting the past and the future: Attention, memory, and delayed intentions. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 27–50). Mahwah, NJ: Erlbaum.

Smith, R. E. (2010). What costs do reveal and moving beyond the cost debate: Reply to Einstein and McDaniel (2010). *Journal of Experimental Psychology: Learning Memory & Cognition, 36,* 1089–1095. doi:10.1037/a0019183

Smith, R. E., & Bayen, U. J. (2006). The source of adult age differences in event-based prospective memory: A multinomial modeling approach. *Journal of Experimental Psychology: Learning Memory & Cognition, 32,* 623–635. doi:10.1037/0278–7393.32.3.623

Smith, R. E., Hunt, R. R., McVay, J. C., & McConnell, M. D. (2007). The cost of event-based prospective memory: Salient target events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 734–746. doi:10.1037/0278–7393.33.4.734

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning Memory & Cognition, 32,* 101–117. doi:10.1037/0278–7393.1.101

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592. doi:10.1037//0033-295x.108.3.550

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32,* 1206–1220.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39,* 767–775.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52,* 1–9. doi:10.1016/j.jmp.2007.09.005

Wagenmakers, E. J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21,* 641–671. doi:10.1080/09541440802205067

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58,* 140–159. doi:10.1016/j.jml.2007.04.006

Wagenmakers, E. J., van der Maas, H. L. J., & Grasman, R. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14,* 3–22.

West, R. (2007). The influence of strategic monitoring on the neural correlates of prospective memory. *Memory & Cognition, 35,* 1034–1046.

Received September 3, 2010
Accepted January 4, 2011 ■

Interference From Remembering to Remember: A Diffusion Model Account

Sebastian S. Horn and Ute J. Bayen

Heinrich-Heine-Universität Düsseldorf

Rebekah E. Smith

The University of Texas at San Antonio

Author Note

Sebastian S. Horn and Ute J. Bayen, Institut für Experimentelle Psychologie, Heinrich-Heine-Universität Düsseldorf; Rebekah E. Smith, Department of Psychology, The University of Texas at San Antonio.

Parts of this research were presented at the 43rd Annual Conference of the Society for Mathematical Psychology, Portland, Oregon, August 2010, and at the 3rd International Conference on Prospective Memory, Vancouver, Canada, July 2010. We thank André Haese, Viviane Küppers, Maike Lex, and Julia Sasse for assistance with data collection. The research was supported in part by Grant AG034965 from the National Institute on Aging (RES).

Correspondence concerning this article should be addressed to Sebastian Horn, Institut für Experimentelle Psychologie, Heinrich-Heine-Universität, Universitätsstr. 1, 40225 Düsseldorf, Germany. E-mail: S.Horn@uni-duesseldorf.de

Abstract

Research on event-based prospective memory (PM) has increasingly focused on the extent to which PM requires attentional resources by examining ongoing-task performance. Performing a PM task can affect the latencies or accuracy of an ongoing task. However, surprisingly little is known about the specific reasons for this interference effect. To date, conclusions regarding this effect have primarily relied on mean latencies or on accuracy in isolation. The authors introduce a diffusion model account (Ratcliff, 1978) of interference effects in the paradigm of PM. Across three experiments, performing a PM task increased the latencies in an ongoing lexical-decision task. This effect was exclusively explained by increases of two separate components of the diffusion model, boundary separation and nondecision time. Primarily the latter component explained variance in PM performance (Experiment 1). The same two components were consistently affected by instructional manipulations of task importance (Experiment 2) and by focal/nonfocal PM targets (Experiment 3). In sum, these results suggest that speed-accuracy settings and nondecision processes provide an important contribution to the interference effect. Mechanisms how PM could affect these components are described.

Keywords: prospective memory, diffusion model, monitoring, attentional resources, intentions, lexical decision

Interference From Remembering to Remember: A Diffusion Model Account

Prospective memory (PM) refers to remembering to perform intended actions after a delay, which is of great importance in everyday life (Crovitz & Daniel, 1984). In some instances, intentions must be remembered at a particular time (e.g., take the cake out of the oven after 35 min; time-based PM), whereas in others, intentions become relevant when a target event occurs (e.g., remember to stop at the post office on your way home; event-based PM). A major challenge is that PM tasks often occur in the context of other ongoing activities that must be interrupted to perform the intended action (Shallice & Burgess, 1991). Paralleling naturalistic settings, the PM laboratory paradigm therefore generally involves embedding the PM task in an ongoing task (Einstein & McDaniel, 1990). For instance, participants may be busily engaged in an ongoing lexical decision task (LDT) in which they classify letter strings either as words or as nonwords. When an event-based PM task is added to this ongoing task, participants must remember to carry out another action when a particular target event occurs. For example, participants may be asked to press the *F1* key when the word *tiger* appears.[1] Successful completion of such PM tasks comprises two components (Einstein & McDaniel, 1990, 1996). The *prospective* component refers to the processes needed to realize *that* something needs to be done (i.e., remembering to remember; Harris, 1984). The *retrospective* component involves recognition of relevant targets and remembering *what* the intended action was. Research on event-based PM has largely focused on the prospective component, which has been conceptualized in different ways.

In one view, environmental cues may trigger spontaneous noticing of a target event or spontaneous retrieval of the intention. For instance, a target event could elicit an experience of familiarity (Einstein & McDaniel, 1996) or discrepancy (Breneiser & McDaniel, 2006; cf.

Whittlesea & Williams, 2000) that causes the target to be noticed as significant, thereby stimulating a controlled search in memory; moreover, well-encoded associations between intended actions and targets might promote reflexive retrieval of the intention when a target is encountered (McDaniel, Guynn, Einstein, & Breneiser, 2004; cf. Moscovitch, 1994). In an alternative view, the prospective component typically imposes high demands on self-initiated retrieval and executive control (Craik, 1986; Marsh & Hicks, 1998). Preparatory attentional and memory processes theory (PAM; Smith, 2003, 2010) takes the strong stance that resources from a limited pool of capacity must always be allocated for successful PM. Individuals must engage in preparatory attentional processes during the performance interval in order for the intention to be retrieved. These processes can involve strategic monitoring, rehearsal of intended actions, or preparatory processes on the periphery of awareness. In the latter case, this involves a subtle shift in processing that allows an individual to be prepared to make a response that is different from the ongoing response, but this preparation need not be the focus of attention (Smith, 2008; Smith, Hunt, McVay, & McConnel, 2007). Finally, the multiprocess view (MPV; McDaniel & Einstein, 2000, 2007) suggests that qualitatively different processes can predominate in PM retrieval, depending on the situation. That is, the extent to which resource demanding processes contribute to successful PM depends on the characteristics of the PM task, the ongoing task, and the individual (see Kliegel, Martin, McDaniel, & Einstein, 2001, and McDaniel & Einstein, 2000, for the factors assumed to promote spontaneous retrieval).

The empirical approach towards testing these conflicting theories relies on the comparison of ongoing-task performance in the absence versus presence of a PM task. The core assumption is that a nonautomatic prospective component absorbs attentional resources that would otherwise be devoted to performing the ongoing task. Smith (2003) showed that response time (RT) in the ongoing task was increased by the simultaneous need to remember a PM task. This *cost-* or *interference-effect* of PM was measured on non-PM-target trials, suggesting that additional resources were absorbed, not only in the presence of the target event (e.g., for executing the intended action), but also prior to the target occurrence. Recently, numerous studies have measured interference to ongoing-task performance to assess the resource demands of PM (see Smith et al., 2007, for an overview). This includes examinations of the functional relationship between task interference and PM (McNerney & West, 2007; Smith & Bayen, 2004; West, Krompinger, & Bowry, 2005), and investigating whether or not interference is always demonstrated (Cohen, Jaudas, & Gollwitzer, 2008; Einstein et al., 2005; Smith et al., 2007). Despite extensive debate on these questions, surprisingly little is known about the underlying processes that contribute to the cost of PM and thus to the nature of this task interference. As summarized by McDaniel and Einstein (2007) “the point here is that we do not yet have good understanding of what is producing the slowing . . .” (p. 225).

Addressing this open question, the present work examines the processes contributing to the PM interference effect with Ratcliff’s (1978) diffusion model. Cognitive process models, such as the diffusion model, can help us better understand interference effects, because they provide substantive insight into the reasons *why* RTs are fast or slow, and *why* few or many errors are made. Particularly, the diffusion model quantifies and disentangles several process components in an ongoing task that are informative for theories of PM: the speed of information accumulation, the boundaries determining the amount of information that must accumulate before a decision is made, the nondecision time needed for encoding and response execution, and variability in these components.

The model is ideally suited to analyzing ongoing-task performance because it simultaneously accounts for both RT and accuracy. A combined metric of performance is an important improvement over separate analyses of these dependent measures, which can easily yield misleading interpretations in cases of speed-accuracy tradeoffs (e.g., Pew, 1969; Schouten & Bekker, 1967). That is, participants may decrease their speed for the sake of higher accuracy, or vice versa. Such tradeoffs imply that RTs could be slower not because a task is more difficult or more resource-demanding, but because participants set a more conservative decision criterion. It is important to extract such tradeoffs before conclusions about actual ongoing-task performance are derived. Moreover, the degree of data utilization with the diffusion model is excellent because the complete RT distributions for correct and error responses are used as input. In standard analyses of ongoing-task performance, 50% or more of the data are typically excluded (in the LDT, PM researchers have mainly considered RTs of correct responses on words; but see Cohen et al., 2008, for a discussion). By contrast, fitting the diffusion model relies on information from both stimulus types (e.g., words and nonwords) and response types (correct vs. error), from the whole RT distributions, their shape, and the error rates. In sum, a diffusion-model analysis contributes to our understanding of the nature of task interference by decomposing this effect into substantive components and comes with strong methodological improvements over extant analyses.

The Diffusion Model

The diffusion model is a sequential-sampling model for fast two-choice decisions with mean latencies not much over 1 s (e.g., Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & Tuerlinckx, 2002; see Brown & Heathcote, 2008, or Usher & McClelland, 2001, for related approaches). A core assumption of the diffusion model is that two-choice decisions are based on noisy information that accumulates continuously over time. Like signal-detection models, the diffusion model includes a decision-related strength-of-evidence axis. This is the vertical axis in Figure 1, whereas the horizontal axis represents the time domain. Diffusion processes start from a point z that lies between two decision boundaries on the strength-of-evidence axis, and move over time until one of these boundaries is reached. In this case, a decision is made and the corresponding response is initiated.

The drift rate parameter v quantifies the direction of information accumulation and its mean speed (i.e., the amount of accumulated information per time unit). The drift rate is a systematic influence that drives a diffusion process to one of the boundaries. A positive drift ($v > 0$) indicates that more evidence is collected in favor of the upper boundary (A) and against the lower boundary (B), whereas a negative drift ($v < 0$) implies the opposite. Drift rate determines processing efficiency in the actual decision phase, with high absolute values predicting both fast and accurate decisions. Thus, drift rate can be interpreted as a measure of the decision maker's ability (in interindividual comparisons), or as a measure of task difficulty (in intraindividual comparisons). There is also a nonsystematic, stochastic influence on the accumulation process within a trial[2] (intratrial variability s), as illustrated by the two sample tracks oscillating between the boundaries in Figure 1. This implies that trials with the same mean drift rate v do not necessarily terminate at the same time (thereby generating RT distributions) or at the same boundary (thereby generating errors).

Parameter a quantifies the distance between the boundaries and hence the amount of information that is required to come to a decision. This parameter reflects a decision maker's speed-accuracy criterion. That is, larger boundary separation indicates that more information must accumulate until a decision is made. In this case, decisions are slower and more accurate (a conservative speed-accuracy setting). Smaller boundary separation indicates that less information must accumulate, and decisions are faster but less accurate (a liberal speed-accuracy setting). The relation of parameter z to the boundaries quantifies decision bias. If the position of the starting point z is equidistant from both boundaries ($z = a/2$), there is no a priori bias to favor one decision over the other, which approximately holds in many applications (Wagenmakers, van der Maas, & Grasman, 2007). Deviations from $a/2$ reflect asymmetries in the amount of information required for the decisions. That is, values larger or smaller than $a/2$ indicate a bias in favor of the upper or the lower boundary, respectively.

In the diffusion model, observed total RT is the sum of the decision duration and a nondecision component, RT = $T_{decision} + T_{er}$ (cf. Luce, 1986). The nondecision component T_{er} quantifies the duration of processes *before and after* the actual decision phase, such as stimulus encoding and motor response execution. Changes in T_{er} shift the complete RT distribution, but do not affect decision accuracy.

Finally, the diffusion model allows for variability of parameters across trials of an experiment. Drift rate is assumed to vary normally around mean v with standard deviation η, which reflects trial-by-trial fluctuations of equivalent items. The starting point is uniformly distributed around mean z with range s_z. The nondecision component is uniformly distributed around mean T_{er} with range s_t. With intertrial variability in drift rate η, the model can account for error responses that are systematically slower than correct responses, and variability of the starting point s_z accounts for the reverse pattern (Laming, 1968; Ratcliff & Rouder, 1998). The variability in the nondecision component s_t can account for shifts in the leading edge of the RT distribution (Ratcliff, Gomez, & McKoon, 2004; Ratcliff & Tuerlinckx, 2002).

The diffusion model has been successfully applied in many areas of cognitive science, including memory retrieval (Criss, 2010; Ratcliff, 1978; Spaniol, Madden, & Voss, 2006), lexical decision (Ratcliff, Gomez, et al., 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008), implicit social cognition (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007), assessment of clinical disorders (White, Ratcliff, Vasey, & McKoon, 2010), and the identification of components underlying age-related slowing (e.g., Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2006). In general, the diffusion model provided a close fit to the observed RT distributions and accuracy in these studies. Support for the validity of the model parameters comes from both behavioral (Voss, Rothermund, & Voss, 2004) and neuroscientific studies (Ratcliff, Philiastides, & Sajda 2009). For reviews, see Ratcliff and McKoon (2008) or Wagenmakers (2009).

Overview of the Experiments

In the current research, we analyzed the PM interference effect with the diffusion model. Theoretical assumptions about the processes contributing to this effect can thereby be tested in a more direct manner than is possible with traditional methods of analysis. In Experiment 1, we examined the individual differences in the underlying processes of the ongoing task in the absence versus presence of a PM task. This allowed us to assess the relationship between the model parameters and observed PM performance. The subsequent experiments included manipulations that affect the size of the interference effect (e.g., Einstein et al., 2005; Smith & Bayen, 2004), that is, the relative importance of the PM task (Experiment 2), and the degree to which relevant PM target features are in the focus of attention during ongoing-task processing (Experiment 3). The main goal here was to determine how ongoing-task decision processes are affected by characteristics of a PM task that moderate the magnitude of interference.

Various ongoing tasks could provide appropriate data for a diffusion model analysis. For our studies, we chose lexical decisions, which are frequently used ongoing tasks in PM research

(Cohen et al., 2008; Marsh, Hicks, & Watson, 2002; Smith et al., 2007), and which have been studied in detail with the diffusion model. That is, Ratcliff, Gomez, et al. (2004) showed that changes in drift rate alone can explain the impact of lexical information quality (i.e., word frequency and orthographic wordlikeness) on all aspects of the behavioral data. Moreover, emphasizing speed or accuracy in the LDT affected parameter a, whereas proportion manipulations of the stimuli induced decision bias that mapped on the relative position of the starting point z (Wagenmakers et al., 2008). For the interpretation and validity of the model parameters in the PM paradigm, we can build on this foundation.

In the current experiments, we fit the diffusion model separately to the data from each participant (cf. Voss et al., 2004) in order to examine individual differences and to conduct inferential analyses on the level of latent parameters. In order to explore the feasibility of diffusion model analyses in the PM paradigm, we performed reanalyses of data from an experiment published by Smith (2003, Exp. 1). This preliminary application suggested that the model may be a promising tool for the analysis of PM data (Horn, Bayen, & Smith, in press). However, since the experiment on which we performed these reanalyses was not specifically designed for a diffusion model analysis of task interference, the reanalysis was subject to a number of limitations (e.g., very low error-rates, and no baseline control block). In the present work, we addressed these and further issues by using more trials, more difficult stimuli, and a baseline phase for each participant in order to relate the change in model parameters to PM performance (Experiment 1). That is, the LDTs included more trials than in previous applications in the PM paradigm (cf. Marsh et al., 2002; Smith, 2003; Smith et al., 2007) to obtain stable parameter estimates for each individual. Moreover, we used a relatively difficult stimulus composition in the LDT (i.e., low-frequency and very-low-frequency words) to increase variance in accuracy, which makes the design more sensitive to possible speed-accuracy tradeoffs. In previous studies that examined the PM interference effect with the LDT, most of the variance in ongoing-task performance appeared in the RTs, whereas accuracy was near ceiling (e.g., Marsh et al., 2002; Smith et. al., 2007).[3] As a result, possible tradeoffs may have been masked, and effects on accuracy were rarely observed when aggregated across participants. Therefore, we used the well-studied LDT as ongoing task, but with a relatively difficult stimulus composition.

Experiment 1: The Interference Effect of Prospective Memory

The main objective was to identify the processes underlying ongoing-task decisions that are affected by an embedded PM task. To examine the nature of this interference effect, we selected a *nonfocal* task in which interference to ongoing-task performance is generally expected to occur (McDaniel & Einstein, 2000, 2007). An ongoing task that is nonfocal to the PM task does not involve processing the relevant features that define the PM target events. The impact of focality is discussed in more detail in Experiment 3. Specifically, we embedded strings with initial letters *G*, *H*, or *M* as target events in the LDT (2.4 % of the trials) and asked participants to remember to carry out the intended action (press the *F1* key) on these trials. In an LDT, initial letters would be considered nonfocal targets (cf. Scullin, McDaniel, Shelton, & Lee, 2010). Prominent theories of event-based PM agree that in this situation, successful PM requires attentional resources that would otherwise be devoted to the ongoing task (Marsh, Hicks, Cook, Hansen, & Pallos, 2003; McDaniel & Einstein, 2000, 2007; Smith, 2003, 2010). We therefore expected to replicate the interference effect of PM, which should appear as increased RTs on *nontarget* trials in the ongoing task. Moreover, assuming a functional relationship with PM, the amount of slowing in the ongoing task should correlate with PM performance, when other factors are held equal (Smith, 2003; Smith & Bayen, 2004).

Despite converging predictions about behavioral measures for task interference (e.g., Einstein et. al. 2005; Smith et al., 2007), little is known about the underlying processes that contribute to this effect. Possibly, part of the PM interference effect reflects participants' metacognitive beliefs about the upcoming task segment (Einstein & McDaniel, 2008, 2010; Smith, 2008). Marsh, Hicks, and Cook (2005, 2006) suggested that individuals calibrate their attentional allocation policies in the PM paradigm, for instance by using contextual cues at the beginning of an experimental phase. In this view, the interference effect can be explained by strategic policies that are usually formed at the outset on an entire task-set (see also Hicks, Marsh, & Cook, 2005). These strategic policies could result in criterion-changes (i.e., speed-accuracy settings) to some extent. Support for this hypothesis comes from the Marsh et al. (2005) study, in which participants' allocation policies were manipulated through the use of auditory importance signals that induced different degrees of effort toward an ongoing task. Their assumption was that PM may benefit from lower effort toward the ongoing task because PM target detection can receive relatively more attention in this case. Notably, lower effort towards the ongoing task resulted in significantly *slower* latencies as well as *higher* accuracy in the ongoing task. In fact, a significant speed-accuracy tradeoff emerged in one of these experiments, even despite high accuracy in the ongoing LDT (.95 to .98). These data suggest that manipulations of allocation policies may influence participants' speed-accuracy settings, as well. Consistent with this view, our initial analysis suggested that increased boundary separation (i.e., more conservative speed-accuracy settings) contributes to the interference effect (Horn et al., in press). We assume that participants may have perceived the overall task-context as more complex or difficult when they received additional instructions for the PM task, which led them to shift their criterion and respond more cautiously. In Ratcliff's diffusion model, boundary shifts are possible whenever individuals can know which condition is being tested (e.g., with explicit instructions at the outset of a phase; cf. Ratcliff, 1978, p. 69). For the current experiments, we therefore anticipated that participants adopt more cautious speed-accuracy settings (i.e., an increase in parameter *a*) when they receive instructions for a PM task that is subsequently embedded in the ongoing LDT.

Importantly, the PM task is expected to impact other parameters, in addition to boundary-changes, for the following reason. Changes in parameter *a* reflect criterion shifts on the speed-accuracy operating characteristic, in conceptual analogy to operating characteristics of signal-detection models (cf. Pew, 1969; Starns & Ratcliff, 2010); we know that criteria may be adjusted by anticipated difficulties in order to compensate for poorer information processing (e.g., Ratcliff & Van Dongen, 2009), but they do not directly reflect more difficult ongoing-task decisions per se. That is, parameter *a* does not capture the additional processing requirements that may be associated with a PM task, but instead reflects a change in the way the ongoing task is approached overall (Marsh et al., 2005; cf. Smith, 2008, 2010). As noted above, the current experiment uses a PM task that is expected to encourage strategic monitoring (McDaniel & Einstein, 2000) and this implies changes in at least one additional component (nondecision time T_{er} or drift rate *v*), because PM absorbs attentional resources from the ongoing task. As an analogy, consider that decreased processing resources associated with aging appear to underlie age-related differences in a variety of cognitive tasks (e.g., Stoltzfus, Hasher, & Zacks, 1996). By applying the diffusion model, researchers have found that age-related differences in speed and accuracy in a variety of tasks (including the LDT), were due to changes of the boundaries (parameter *a*) and nondecision time (parameter T_{er}); effects of aging on drift rates were found in only few paradigms (Ratcliff, Thapar, et al. 2004; Ratcliff, Thapar, & McKoon, 2006). If embedding a PM task in the ongoing task decreases resource availability for ongoing-task processing, this may also affect the nondecision component, possibly drift rate, or both. Thus, the diffusion model provides multiple but separable explanations for the interference effect of PM. Our main goal was to use the diffusion model to determine the best explanation for PM interference effects. In addition, by applying the model to individual participant data from two different phases (ongoing task only vs. ongoing task + PM), we intended to measure how variance in PM performance is related to intraindividual variation in the different processing components of the diffusion model.

Method

Participants and design. Forty-seven students (39 female) participated for monetary compensation or for course credits in psychology-classes at the University of Düsseldorf. Their age ranged from 19 to 38 years (M = 24). The participants in all three experiments were native speakers of German and were tested in sessions involving one to three participants. Each participant was randomly assigned to one of two groups. In the PM group, 22 participants first performed a phase of the LDT alone, followed by a second phase with the embedded event-based PM task. In the No-PM group, 25 participants performed the LDT alone in both phases and did not receive PM instructions. The No-PM control group was included to assess the extent of fatigue or practice on LDT performance in the absence of a PM task. Stimulus type was manipulated within participants.

Materials. A pool consisting of 539 low-frequency words (4 or 5 occurrences per million, M = 4.42, SD = 0.49), and 541 very-low-frequency words (0 or 1 occurrence per million, M = 0.33, SD = 0.47) was selected from the CELEX lexical database for German lemmas (Baayen, Piepenbrock, & Gulikers, 1995). Word length ranged from 8 to 12 letters (M = 9.83, SD

= 1.30). From each word, a pseudoword was generated by randomly replacing one interior vowel or one umlaut with another vowel or umlaut, respectively (cf. Perea, Rosa, & Gomez, 2005). The words were screened by four student assistants, and words that any of these students did not know were eliminated. The pseudowords were also screened to avoid the possibility of a newly generated word. From this larger pool, we created matched stimulus sets. Each of the sets included 50% words and 50% pseudowords. Each item occurred as a word in one stimulus set and as the corresponding pseudoword in the matched stimulus set. In this manner, we created two stimulus sets for the ongoing LDT (filler sets) and two stimulus sets for the PM task (target sets). One of each was randomly selected with equal probability during an experiment. Thus, no participant was ever presented with both a word and the corresponding pseudoword within a session, in order to avoid repetition priming. We counterbalanced the sets across participants and conditions That is, the four possible combinations of filler set and target set occurred approximately equally often across groups.

The two filler sets included 988 items each and were matched for word frequency and length. Each of the filler sets contained 247 low-frequency words, 247 very-low-frequency words, and 247 + 247 pseudowords, created from the words of the respective other set. During the experiment, items were randomly drawn from one of the filler sets without replacement, with the restriction that an equal number of words and pseudowords, as well as an equal number of low-frequency words and very-low-frequency words, occurred in each phase of the experiment.

For the PM task, two sets of 12 target items each were selected and matched for word frequency and length. The target sets exclusively contained items with the initial letters *G*, *H*, and *M*. None of the filler items started with any of these letters. Each of the three initial letters occurred four times in each target set: in one low-frequency word, in one very-low-frequency word, and in 1 + 1 pseudowords, created from the words of the respective other set. For the PM tasks in all experiments, we used items starting with the letters *G*, *H*, or *M* as targets because their initial-letter frequency in German texts is approximately equal (ranks 9, 14, and 8, respectively, among 29 letters; Schönpflug, 1969). Moreover, these letters are phonologically and orthographically relatively distinct (unlike, for example, the letters *G* and *C*). During the second phase of the experiment, items were randomly drawn without replacement from one of the target sets when a target trial occurred. The target items occurred in both the PM and No-PM group, but filler items and target items were indistinguishable from a participant's perspective in the No-PM group, as there was no PM instruction.

Procedure. Presentation of instructions and materials, as well as response collection, was computer directed in all experiments. At the start of each experiment, participants read the LDT instructions, which emphasized both speed and accuracy. Participants then completed 48 LDT practice trials (24 words and 24 pseudowords), presented in random order. Participants received feedback on both speed and accuracy in the practice trials. In case of an erroneous response or if response latency exceeded 700 ms, the feedback message *incorrect* or *too slow*, respectively, appeared on the screen in red for 800 ms. At the end of the practice phase, participants saw a summary of their LDT performance (mean accuracy and latency) and had the opportunity to ask questions.

The experiment then continued with two phases of 500 trials each, without any further performance feedback. Each LDT trial started with a fixation cross (12 × 12 mm) that remained in the center of the screen for 400, 450, 500, 550, or 600 ms. These display times were randomly selected with equal probability to prevent anticipatory responding. Immediately after disappearance of the fixation cross, a letter string was displayed. The strings appeared in black uppercase letters in sans serif font in the center of a white screen with a height of 7 mm and a width of 42 - 82 mm. A random half of the participants were instructed to press the *F* key with their left index finger when they believed the letter string to be a German word and to press the *J* key with their right index finger when they did not believe the letter string to be a German word; the other half of the participants received instructions for the reversed key-response mapping. Key-response mappings were counterbalanced to control for potential handedness bias in the LDT responses. After response entry, the string disappeared, and a white screen was displayed for 200 ms, followed by the fixation cross for the next trial.

At the end of the first phase, participants had a short break. They were then informed that they should memorize a few letters before the second phase. In the PM group, participants read that they should remember to press the *F1* key instead of the *F* or *J* key if a string started with one of these letters in the subsequent phase of the LDT. The instructions included one word and one pseudoword with the initial sample letter *A* as an example, which participants did not have to remember afterwards. Participants were also informed that they could still press the *F1* key immediately after their response to the ongoing LDT, if necessary. In the No-PM group, participants read that they should remember the letters as well as the alphanumeric code "F1", but no instruction to perform another action was given. In both groups, participants were informed that their memory for these letters would be tested at the end of the session, and the letters *G*, *H*, and *M,* were then presented sequentially in random order, for 5 s each. Participants had to recall the letters by typing them into the keyboard. If errors occurred, the letters were again displayed until perfect recall. The experiment continued with the LDT instructions for the second phase, again emphasizing both speed and accuracy, and there was no further mention of the PM task.

The 12 PM target events occurred in trial ranges 38-42, 78-82, 118-122, 158-162, 198-202, 238-242, 278-282, 318-322, 358-362, 398-402, 438-442, and 478-482. We used ranges of five trials, on which a target event could occur with equal probability, to avoid small distances between targets (mean target distance was 40 trials, or approximately 1 min), and to avoid deterministic, evenly spaced target events. Both the number of target events and the target distance are consistent with previous PM research (e.g., Marsh et al., 2002; Smith, 2003).

Following the second phase of LDT trials, participants in the PM group recalled which key they were supposed to press in response to a target item. Participants in both groups then completed an old-new recognition-memory test in which the three target letters and three distracter letters were presented in random order. Finally, participants reported how they weighted speed versus accuracy in the LDT, completed a demographic questionnaire, and a multiple choice vocabulary test (MWT-B; Lehrl, 1995).

Results and Discussion

Prospective memory. In all three experiments, PM was measured as the proportion of target events to which participants made a correct PM response. A press of the *F1* key any time after occurrence of a PM target string and within the subsequent two LDT trials was counted as a correct PM response. We included these shortly delayed PM responses because the ongoing task was highly speeded and computer-paced. Across the three experiments, 65.76% of all correct PM responses occurred on the target trial, 33.79% occurred on the first trial following a target, and the remaining PM responses occurred on the second trial following a target.

PM performance was moderately high (M = .45, SE = .06) in the PM group, and may have been lower than in other studies employing the LDT (e.g., Marsh et al., 2003) because participants were busily engaged in a more demanding and speeded ongoing task.

Retrospective memory. Recognition memory (hit rate minus false alarm rate) for the letters did not differ between groups ($t < 1$) and was near ceiling in the postexperimental test ($M = .96$, $SE = .02$). These findings imply that both the variability and the demands of the retrospective component were low and PM failures were likely due to failures in retrieving the intended PM action when a target event occurred (i.e., due to variability in the prospective component). Finally, when queried at the end of the experiment, 20 out of 22 participants in the PM group correctly recalled the PM action (press the *F1* key). The two participants who failed to recall the action clearly showed PM performance above zero and likely misunderstood the postexperimental recall question.

Ongoing-task performance. For all analyses of ongoing-task data, we excluded the PM target trials as well as the two trials immediately following each target trial, and any trials on which participants made PM false alarms. This is a standard approach to avoid possible contamination by switch costs or by response coordination on target trials (e.g., Cohen et al., 2008; Marsh et al., 2002). Moreover, responses shorter than 200 ms or longer than 3000 ms were eliminated, as were responses that were extreme outliers in an individual's RT distribution as defined by Tukey's criterion,[4] which led to the exclusion of 0.83% of the nontarget trials in Experiment 1.

Table 1 shows mean latencies and accuracy in the ongoing LDT, aggregated over stimulus types and participants in each group. We report ANOVAs on these data, with *group* (PM vs. No-PM) as between-participants factor and *phase* (Phase 1 vs. Phase 2) as within-participants factor. As seen in Table 1, ongoing-task RTs increased selectively with the embedded PM task in Phase 2. In consequence, the main effect of phase, $F(1, 45) = 19.80$, $p < .01$, $\eta_p^2 = .31$, was qualified by a phase × group interaction, $F(1, 45) = 36.01$, $p < .01$, $\eta_p^2 = .45$.

As expected, mean RT did not differ between the PM group and the No-PM group in Phase 1 of the experiment, in which all participants worked on the LDT alone as a baseline. By contrast, a large group difference in RT emerged in Phase 2, when a PM task was embedded for the PM group (see Table 1 for *t* tests).

To examine the interference to ongoing-task performance within participants, individual difference scores were calculated by subtracting Phase 1-latencies from Phase 2-latencies (cf. Cohen et al., 2008; Smith & Bayen, 2004). Replicating the interference effect of PM, participants in the PM group showed pronounced slowing when a PM task was added to the ongoing task, $t(21) = 5.31$, $p < .01$, $d = 1.13$, whereas participants in the No-PM group responded even faster in Phase 2, possibly reflecting practice or a slight criterion change, $t(24) = 2.13$, $p = .04$, $d = 0.43$. Therefore, any RT slowing in the PM group is unlikely due to fatigue. The extent of this slowing correlated positively with PM performance, $r(22) = .57$, $p < .01$. This finding is in line with research advocating a functional relationship between the interference effect and PM (e.g., Smith, 2003; Smith & Bayen, 2004).

As in previous studies (e.g., Marsh et al., 2002), embedding a PM task in the LDT did not affect the mean proportion of correct responses in the LDT; the ANOVA revealed no significant effects on this measure when data were aggregated across participants, all Fs(1, 45) < 3.23, ps > .07. However, separate *t* tests on d' scores indicated an *increase* in accuracy within the PM group when the PM task was embedded, $t(21) = 2.80$, $p = .01$, $d = 0.60$, but no such change within the No-PM group, $t < 1$. These results suggest that variability in error rates deserves more attention.

Speed-accuracy tradeoffs. To shed more light on the interplay between speed and accuracy when a PM task is added, we plotted each individual's change between the two phases in mean accuracy against the change in mean RT. As seen in Figure 2, there is a clear tendency for participants responding more slowly in Phase 2 to respond more accurately, and vice versa. This positive relationship was found for both the PM group, $r(22) = .56$, $p < .01$, and the No-PM group, $r(25) = .53$, $p < .01$, possibly reflecting speed-accuracy tradeoffs between the phases that should be considered before interpreting any changes in ongoing-task performance. Thus, it remains unclear to what extent the slowing in the PM group can be attributed to a criterion change, and whether or not there are additional processes that contribute to this slowing in the second phase with an embedded PM task. The scatterplot therefore emphasizes a clear advantage of a diffusion model analysis. That is, speed and accuracy need not be considered in isolation, and criterion settings can be measured and separated from possible changes in other processes.

Diffusion model analysis. For each participant, diffusion models with six free parameters were fitted separately to the data from the two experimental phases, yielding 47×2 submodels. Each submodel was based on 500 responses minus the number of excluded responses, if any, thereby following the criteria described above. Because lexical variables were not the focus of this study, we aggregated the responses over the three item types (low-frequency words, very-low-frequency words, pseudowords) to obtain an overall-measure of drift rate v.[5] Responses in the LDT were coded as *correct* (upper boundary) versus *false* (lower boundary) and, in consequence, the mean starting point z was fixed at $a/2$ (e.g., Ratcliff et al., 2006). Any decision biases are expected to cancel each other out across trials in the present analysis (as the model aggregates over stimuli that are presented in equal proportions), but can contribute to starting-point variability (see Klauer et al., 2007). That is, biases towards an item-type may even be present, but they would not shift the mean starting-point value because correct and false responses are favored equally often under this coding. The current parameter estimates are based on the Kolmogorov-Smirnov statistic (KS; Kolmogorov, 1941). This estimation method uses the exact shape of the entire RT distribution and is relatively robust against outliers (Voss & Voss, 2008; for alternative estimation methods, see Ratcliff & Tuerlinckx, 2002). The test statistic *T* represents the maximum vertical distance between the empirical and the predicted cumulative distribution function of RT. In a multidimensional search, a program determines the parameter values in such a way that *T* is minimized (Voss & Voss, 2007, 2008).

Figure 3 shows the results of the diffusion-model analysis for boundary separation *a*, nondecision time T_{er}, and drift rate *v*, averaged over participants in a group. As seen in the upper panels (Experiment 1), the interference effect of PM is clearly reflected in a marked increase of boundary separation and nondecision time, but not in drift rates. Focusing on each process component, we conducted ANOVAs with the factors phase and group.

For boundary separation *a*, main effects of phase, $F(1, 45) = 6.92$, $p = .01$, $\eta_p^2 = .13$, and group, $F(1, 45) = 4.23$, $p < .05$, $\eta_p^2 = .09$, were qualified by an interaction of both factors, $F(1, 45) = 21.26$, $p < .01$, $\eta_p^2 = .32$. That is, speed-accuracy settings did not differ between groups in Phase 1 ($t < 1$), but there was a large effect in Phase 2, $t(45) = 3.04$, $p < .01$, $d = 0.89$. As expected, these differences indicate an increase in boundary separation (more cautious speed-accuracy settings) when the PM task was embedded in the LDT, $t(21) = 3.99$, $p < .01$, $d = 0.61$. By contrast, there was a trend towards even more liberal speed-accuracy settings in the second phase for the No-PM group, $t(24) = 2.01$, $p = .06$, reflecting the slight speed-up in RTs between the phases.

For the nondecision component T_{er}, a similar pattern emerged. That is, main effects of phase, $F(1, 45) = 23.51$, $p < .01$, $\eta_p^2 = .34$, and group, $F(1, 45) = 5.21$, $p = .03$, $\eta_p^2 = .10$, were qualified by an interaction of both factors, $F(1, 45) = 23.81$, $p < .01$, $\eta_p^2 = .35$. No differences in nondecision time emerged between the groups in Phase 1 ($t < 1$), but there was a large difference in Phase 2, $t(45) = 3.41$, $p < .01$, $d = 0.99$. The interaction is explained by the strong increase of participants' nondecision time (i.e., a slowdown before or after the actual ongoing-task decision) when the PM task was added in Phase 2, $t(21) = 5.03$, $p < .01$, $d = 1.07$, whereas nondecision time remained stable between the phases in the No-PM group, $t < 1$.

Notably, there were no effects on drift rate in this experiment (all Fs < 1, all ts < 1.06). Because the power of tests for the different process components was equal and sufficiently high ($1 - \beta \approx .80$ for an interaction effect size $f = .15$, Cohen, 1988)[6], and because the magnitude of drift rate variability η was comparatively low (cf. Wagenmakers et al., 2007, for typical value-ranges), we conclude that the interference effect of PM could not be explained by differences in drift rates.

Table 2 includes the variability parameters, for which we conducted the same analyses. The variability parameters are important for the shape of the RT distributions and for the relative speed of correct versus false responses (Ratcliff & Rouder, 1998). However, they have relatively little impact on mean latencies and accuracy and are therefore of secondary interest in accounting for the interference effect of PM. There were no significant effects on the starting point range s_z (divided by the maximum range a; Fs $(1, 45) < 2.23$, ps > .14), or on variability in drift rate η (Fs < 1). For the range of the nondecision component s_t, a similar pattern emerged as for nondecision time T_{er} itself, with a significant phase × group interaction that could be explained by more variable nondecision time with an embedded PM task, $F(1, 45) = 8.50$, $p < .01$, $\eta_p^2 = .16$.

To summarize, changes in boundary separation and changes in nondecision time contributed jointly to the interference effect of PM, whereas no differences occurred in the drift rates. Thus, one portion of the PM interference effect was indeed due to participants' adopting more cautious speed-accuracy settings with an embedded PM task, and was mirrored in positive correlations between the change in speed and in accuracy discussed above. Importantly, the results also indicated a strong increase in nondecision time T_{er} with an embedded PM task, and strategic monitoring or other preparatory processes could therefore contribute to the interference effect and be reflected in this parameter. One way to examine the role of nondecision time is to regress PM performance on the change in model parameters when the PM task is added (cf. Klauer et al., 2007). If strategic monitoring is reflected in nondecision time T_{er}, it should have a functional role for PM, as current theories of PM would predict (e.g., McDaniel & Einstein, 2007; Smith, 2003).

Figure 4 shows the results of a regression analysis (PM group) with change in boundary separation Δa, nondecision time ΔT_{er}, and drift Δv, as predictors for PM. Notably, a larger amount of variance of PM was explained by the diffusion model parameters (R^2 = .45) than by observed RT (R^2 = .33). Although there were pronounced increases in both boundary separation and nondecision time when the PM task was added, only the change in the latter component predicted PM. Neither Δa nor Δv contributed significantly to the regression equation, although there was a trend for changes in drift rate (associated with processing efficiency) to account for variance in PM, with decreased processing efficiency associated with increased PM performance. Together, these findings support the idea of a functional relationship between nondecision time T_{er} and PM, which may reflect the contribution of strategic monitoring. Although shifts of the boundaries were a source of variance contributing to the PM interference effect as well, they did not account for PM performance. This may be due to the fact that speed-accuracy settings are influenced by a variety of additional factors, such as personality traits, mood, age, and the like (see Klauer et al., 2007; Starns & Ratcliff, 2010).

Model fit. Goodness of fit between model predictions and empirical data was evaluated in three steps. First, we computed separate KS tests for the models of each participant. None of 47 KS tests indicated significant deviations of predicted from empirical cumulative RT distributions at the .05 alpha level, with an average *p* value of .85. Second, we examined the model's capability to predict each individual's mean RT and accuracy. Across participants, groups, and phases ($n = 94$), predicted RT and accuracy correlated almost perfectly with observed RT ($r = .99$) and accuracy ($r = .95$), with a mean absolute deviation of 8 ms in latencies and 0.3 % in accuracy, respectively. Thus, the model reproduced the PM interference effect well for each individual. Third, the complete predicted and empirical RT distributions can be visualized to assess model fit. In Figure 5, the RTs from all participants in a group are combined in a single cumulative distribution function, including correct and error responses (see top panels for Experiment 1). For this purpose, RTs of error responses are multiplied by –1, and the corresponding distribution is mirrored at the zero point of the time axis (Voss et al., 2004). Thus, the part of the distribution on the negative side of the *x*-axis represents error RTs and the part on the positive side of the *x*-axis represents RTs of correct responses. The functions intercept the *y*-axis at the error rate. As seen in Figure 5, changes in the complete RT distribution (for both correct and error responses) emerged when the PM group received the additional PM task in Phase 2, and about 14% of the LDT responses were erroneous in this group. Overall, the plots indicate an excellent fit at the aggregate level, with only minor deviations between predicted and empirical RT distributions.

Experiment 2: Importance Effects

Several factors have been identified that influence the size of the PM interference effect (e.g., Einstein et al., 2005). In Experiment 2, we focused on a crucial factor that influences the interplay between ongoing-task performance and PM, namely the perceived importance of the

PM task. A primary objective was to apply the diffusion model to identify those underlying processes that are affected by task importance and that lead to variations in the interference effect.

In previous research, emphasizing the relative importance of the PM task increased ongoing-task RTs, whereas emphasizing the relative importance of the ongoing task had the reverse effect (Einstein et al., 2005; Kliegel et. al., 2001, 2004). It has been suggested that instructional emphasis towards or away from a PM task affects participants' strategies for performing the ongoing task (McDaniel & Einstein, 2007). Specifically, emphasizing the importance of the PM task can intensify resource-demanding preparatory processes directed at the PM task (Smith & Bayen, 2004). We therefore expected corresponding changes in the nondecision component that may reflect strategic monitoring. As indicated by other manipulations of allocation policies (Marsh et al., 2005), inducing lower effort toward the ongoing task by emphasizing the importance of the PM task could also lead to criterion-shifts as a side effect (i.e., increased boundary separation). Finally, increased importance of the PM task should have beneficial effects on PM performance with a PM task that is nonfocal to the ongoing task (see Einstein et al., 2005, and Kliegel et al. 2004, for a discussion of when importance manipulations would be expected to improve PM).

Method

Participants and design. Seventy-five participants (55 female) volunteered for monetary compensation or course credit. Their age ranged from 19 to 39 years (*M* = 23). Each participant was randomly assigned to one of two groups and completed two blocks of the LDT. In the PM group, 38 participants worked on two blocks with an embedded PM task: one block with instructions emphasizing the importance of the PM task (PMI), and another block with instructions emphasizing the importance of the ongoing task (OI). The order of the two blocks was counterbalanced. In the No-PM group, 37 participants performed the LDT in both blocks, but did not receive instructions for a PM task.

Materials. The stimuli were taken from the same pool described in Experiment 1. Two sets of 960 filler items each were chosen for the ongoing LDT and were matched on word frequency and length, respectively. Each set contained 240 low-frequency words, 240 very-low-frequency words, and 240 + 240 pseudowords, created from the words of the respective other set. During the experiment, items were randomly drawn from one of the filler sets without replacement.

For the PM task, two sets of 24 target items each were selected and matched for word frequency and length, respectively. The target sets exclusively contained items with the initial letters *G*, *H*, and *M*; none of the filler items started with any of these letters. Each of the three initial letters occurred eight times in each target set: in two low-frequency words, in two very-low-frequency words, and in 2 + 2 pseudowords, created from the words of the respective other set. The four possible combinations of filler set and target set were counterbalanced across groups and participants.

Procedure. Procedures were the same as in Experiment 1 with the exceptions noted here. To give participants opportunities to rest, each block of 492 trials was divided by two 1-min breaks into three subblocks of 164 trials (160 fillers, 4 targets). Items were randomly selected from the target and filler sets under the following restriction for both target and filler items: each subblock included an equal number of words and pseudowords as well as an equal number of low-frequency words and very-low-frequency words.

Before each of the two blocks, all participants were informed that they would see several letters, that their memory for these letters would be tested at the end of the session, and then encoded the letters *G*, *H*, and *M*.

In the PM group, participants were asked to remember to press the *F1* key if an item started with one of these letters. Before the PMI block, participants read the following instruction: “It is most important that you do not miss a single letter-string beginning with one of the previously presented letters! Please try to remember to press the *F1* key if you see one of these letter-strings. This is the part of the task that is most important for us!” (transl. from German). Before the OI block, participants read: “It most important that you continue deciding as fast and as accurately as possible whether the letter-string is a word or a nonword! Please try to complete this decision task as well as possible. This is the part of the task that is most important for us!” (transl.). In each of the two blocks, 12 PM target events occurred within the trial ranges 36-40, 77-81, 118-122, 159-163, 200-204, 241-245, 282-286, 323-327, 364-368, 405-409, 446-450, 487-491. Items were randomly drawn from one of the target sets without replacement when a target trial occurred. The target items were inserted in both blocks for both the PM and No-PM groups. For easy reference, we also report PMI and OI blocks for the No-PM group in the subsequent comparisons, although the instructions in these two control-blocks were identical.

Results and Discussion

Prospective memory. As expected, PM performance was higher in the PMI block (M = .46, SE = .05) than in the OI block (M = .31, SE = .04), $t(37) = 2.74$, $p < .01$, $d = 0.62$.

Retrospective memory. Recognition memory for the letters was near ceiling in the postexperimental test (M = .96, SE = .02) and did not differ between the PM and No-PM group, $t(51.44) = 1.07$, $p = .29$, suggesting that retrospective memory demands were constantly low.

Moreover, 100% of the participants in the PM group correctly recalled the PM action (press the *F1* key).

Ongoing-task performance. Trimmed mean latencies and accuracy in the ongoing LDT are summarized in the middle of Table 1. For the latencies, an ANOVA with the factors *group* (PM vs. No-PM) and *importance* (OI vs. PMI) indicated significant main effects of group, $F(1, 73) = 10.81$, $p < .01$, $\eta_p^2 = .13$, and importance, $F(1, 73) = 16.27$, $p < .01$, $\eta_p^2 = .18$, which were qualified by an interaction of both factors, $F(1, 73) = 8.35$, $p < .01$, $\eta_p^2 = .10$. Compared with the No-PM control group, the latencies of the PM group were higher in both the OI block and the PMI block. Importantly, participants in the PM group were significantly slower in the PMI block than in the OI block, $t(37) = 4.21$, $p < .01$, $d = 0.68$. Consistent with previous work, emphasizing the importance of the PM task increased the RTs in the ongoing task.

The proportion of correct responses in the LDT was significantly higher in the PM group than in the No-PM group, $F(1, 73) = 14.09$, $p < .01$, $\eta_p^2 = .16$, and this was the only significant effect to emerge in this ANOVA. Notably, a paired *t* test revealed that accuracy was significantly higher in the PMI block than in the OI block in the PM group, $t(37) = 2.36$, $p = .02$, $d = 0.38$. Together, these findings point strongly to the necessity of considering differences in speed-accuracy settings: RT as well as accuracy was higher (a) when a PM task was added and (b) when the importance of this PM task was emphasized. It is therefore possible that the same processes contributing to the PM interference effect in general are also responsible for a slowdown when the importance of a PM task is emphasized.

Diffusion model results. We fit diffusion models separately to the ongoing-task data from each participant and block, yielding 75×2 submodels. Each submodel was based on 456 nontarget trials minus the excluded responses (1.02% of the nontarget trials of Experiment 2 were excluded, following the same criteria described for Experiment 1). The results of the diffusion-model analysis for boundary separation a, nondecision time T_{er}, and drift rate v, averaged over participants in a group, are shown in the middle panels of Figure 3. Replicating the previous experiment, the PM interference effect mapped onto boundary separation a and nondecision time T_{er}. Importantly, these same two parameters were also affected by the importance manipulation, whereas drift rates were indistinguishable as a function of group and importance. For a more detailed inferential analysis, we calculated 2 × 2 split-plot ANOVAs on each model parameter with *importance* and *group* as factors.

Boundary separation a was larger with instructions emphasizing the importance of the PM task and with an embedded PM task, as indicated by main effects of importance, $F(1, 73) = 6.16$, $p = .02$, $\eta_p^2 = .08$, and group, $F(1, 73) = 14.77$, $p < .01$, $\eta_p^2 = .17$. Paired t tests indicated that participants within the PM group adopted more conservative speed-accuracy settings when instructions emphasized the importance of the PM task, $t(37) = 2.91$, $p < .01$, $d = 0.47$, whereas no differences emerged in the No-PM group, $t < 1$, which received identical instructions in both blocks. Regarding the PM interference effect, the PM group adopted more cautious speed-accuracy settings than the no-PM group in both the OI block, $t(73) = 3.58$, $p < .01$, $d = 0.83$, and the PMI block, $t(73) = 3.92$, $p < .01$, $d = 0.91$.

The pattern of the nondecision component T_{er} was similar to that of boundary separation, suggesting that both parameters contributed jointly to the increases in ongoing-task latencies. Nondecision time was longer with instructions emphasizing the importance of the PM task and longer with an embedded PM task, as indicated by main effects of importance, $F(1, 73) = 9.45$, $p < .01$, $\eta_p^2 = .12$, and group, $F(1, 73) = 7.84$, $p < .01$, $\eta_p^2 = .10$, which were qualified by an interaction of both factors, $F(1, 73) = 6.19$, $p = .02$, $\eta_p^2 = .08$. Separate t tests indicated that emphasizing the importance of the PM task led to an increase in nondecision time of approximately 50 ms within the PM group, $t(37) = 3.58$, $p < .01$, $d = 0.58$, but the two LDT blocks did not differ in the control No-PM group, $t < 1$. Finally, nondecison time in the PM group was higher than in the No-PM group in both the PMI block, $t(61.53) = 3.19$, $p < .01$, $d = 0.73$, and the OI block $t(73) = 1.99$, $p < .05$, $d = 0.46$.

No effects on drift rates emerged (all Fs < 1; all ts < 1.18, smallest $p = .25$). As in the previous experiment, this was the case despite sufficient power to detect smaller effects (i.e., $1 - \beta \approx .95$ for an interaction of effect size $f = .15$; Cohen, 1988). The variability parameters are listed in Table 2. There were no significant effects on the starting point range s_z (divided by the maximum range a; Fs (1, 73) < 3.58, ps > .06), whereas drift rate variability η was higher in the No-PM control group, $F(1, 73) = 5.37$, $p = .02$, $\eta_p^2 = .07$. Finally, the range of nondecision time s_t was larger with an embedded PM task and when importance of this PM task was emphasized, as indicated by main effects of group $F(1, 73) = 5.37$, $p = .02$, $\eta_p^2 = .07$, and a group × importance interaction, $F(1, 73) = 5.55$, $p = .02$, $\eta_p^2 = .07$. The range of nondecision time possibly captured portions of the increased variability that was observed in RTs when the PM task was embedded or emphasized.

The results of Experiment 2 are straightforward. Emphasizing the importance of a PM task improved PM and slowed the latencies in the ongoing LDT, which was reflected in increased boundary separation and nondecision time. As in Experiment 1, the interference effect of PM mapped onto these same parameters. The changes in boundary separation were consistent with slight changes in LDT accuracy that were found when a PM task was added or when importance of this task was emphasized. Importantly, boundary shifts alone did not fully account for the RT effects. We consistently found additional changes in nondecision time.

Model fit. Goodness of fit was again evaluated on three levels. None of 75 KS tests for the models for each participant indicated significant deviations of predicted from empirical cumulative RT distributions ($\alpha = .05$), with a mean *p* value of .90. Again, predicted mean RT and accuracy for each individual correlated strongly with observed mean RT ($r = .99$) and accuracy ($r = .97$). Across all participants and blocks ($n = 150$), the mean absolute deviation in RT was 4 ms and 0.6 % in accuracy. Finally, model fit at the aggregate level is displayed in the middle panel of Figure 5. The empirical cumulative RT distributions, including correct and error responses from all participants in a group, agreed satisfactorily with the predicted cumulative RT distributions.

Experiment 3: Focal/Nonfocal Targets

The MPV (McDaniel & Einstein, 2000, 2007) proposes that the resource-demands of a PM task – and therefore the magnitude of task interference – will depend on conjoint characteristics of the PM- and the ongoing-task. *Nonfocal* tasks are proposed to require executive resources (reflected in larger interference), because attention must be shifted to target-features that are irrelevant for ongoing-task decisions. By contrast, *focal* tasks may involve relatively spontaneous intention retrieval, resulting in little or possibly no task interference (Einstein et al., 2005; but see Smith et al., 2007). A PM target is focal to the ongoing task if its relevant features are processed in the routine of this task. For instance, encoding a particular target-word (e.g., *tiger*) likely activates semantic features. If a subsequent ongoing task, such as the LDT, involves semantic judgments, then this target is focal to ongoing-task processing (McDaniel & Einstein, 2000, 2007; see Maylor, 1996, and Morris, Bransford, & Franks, 1977 for related concepts). Processing of a PM target is nonfocal if the ongoing task does not require to direct attention to the relevant target features. For instance, encoding a single letter as a target (as in Experiments 1 and 2) may activate graphemic/phonemic features that are less likely processed in service of a subsequent ongoing task that requires semantic judgments (and not a single-letter search; cf. Scullin et al., 2010). In this experiment, we compared focal/nonfocal effects with the diffusion model because it is unclear so far how underlying ongoing-task decision processes contribute to changes in the magnitude of the PM interference effect as a function of target focality. When using an ongoing LDT, previous research indicated that search- or monitoring-difficulty for initial-letter targets and for word targets is the same (unlike for syllable-targets; cf. Scullin et al., 2010). Therefore, we compared focal word targets and nonfocal initial-letter targets in this experiment in order to study task interference without concern that focal/nonfocal effects are confounded with mere target monitoring-difficulty.

The present goal was not to test whether task interference can or cannot be completely eliminated with focal targets under particular conditions (Einstein & McDaniel, 2010; Einstein et al., 2005; Smith, 2010; Smith et al., 2007). Regardless of this question, there is converging evidence that focal targets can reduce the amount of task interference imposed by a PM task (assumed to reflect resource demands), and can boost PM performance (e.g., Kliegel, Jäger, & Phillips, 2008). Thus far, however, studies of focal versus nonfocal tasks have relied on observed mean RT and accuracy; the specific ongoing-task decision processes under different degrees of focal processing have as yet not been examined in detail (but see McDaniel & Einstein, in press, for emerging neuroscientific approaches). We employed the diffusion model to examine – ceteris paribus – whether the processes underlying ongoing-task decisions differ quantitatively or qualitatively with different PM targets (focal vs. nonfocal).

Method

Participants and design. Sixty participants (44 female) volunteered for credit points in psychology courses or received monetary compensation. Their mean age was 23 years, ranging from 18 to 35 years. All participants completed two blocks of the LDT and were randomly assigned to one of two groups. In the PM group, 30 participants had an embedded PM task in both blocks of the LDT: one block with focal PM targets (*words*), and another block with nonfocal PM targets (*initial letters*). In the No-PM group, 30 participants completed the same LDT task, but did not receive instructions for a PM task.

Materials. The strings in the two filler sets were the same as in Experiment 2. For the PM task, two sets of 15 target words each were selected and matched for word length. Each target set contained five words with the initial letters *G*, *H*, and *M*, respectively. One of the target sets contained only very-low-frequency words and the other set contained only low-frequency words. During the experiment, three words with each of the initial letters were randomly selected by the program as targets for the focal block. The remaining twelve words were selected as targets for the nonfocal block. Thus, participants saw the stimuli from a target set either as focal or nonfocal targets, randomized across participants and conditions in order to keep the target events comparable for both conditions.

Procedure. The procedures matched those of the previous experiment, with the exceptions noted here. Before the focal block, participants in the PM group encoded three target words that started with the letters *G*, *H*, and *M*, respectively. They were told to remember to press the *F1* key whenever they saw one of these words during the subsequent block. Each of the target words occurred four times in the focal block. Before the nonfocal block, participants encoded the initial-letter targets *G*, *H*, and *M* and were asked to remember to press the *F1* key if a word started with one of these letters. In contrast to the previous experiments, the initial-letter targets occurred only on word strings for comparability with the focal condition. Participants were explicitly informed that the initial letters could only occur on word strings. As in the previous experiments, each initial letter occurred on four different strings to prevent participants from associating an initial letter with a particular word, which could functionally convert a nonfocal target into a focal target (cf. Einstein et al., 2005). In the No-PM group, participants memorized the alphanumeric code "F1" as well as three target words before one block and three target letters before the other block. This kept the retrospective-memory load comparable across conditions (focal vs. nonfocal) and groups (PM vs. No-PM). The order of the two blocks was counterbalanced. That is, approximately half of the participants in each group saw the focal targets (words) before the first block and the nonfocal targets (initial letters) before the second block, and vice versa. At the end of the experiment, participants received two recognition-memory tests, one for the words and one for the letters, in which the three target items were randomly mixed with three foils, respectively.

Results and Discussion

Prospective memory. In line with previous studies (Einstein et al., 2005), PM performance was higher with focal targets (M = .81, SE = .04) than with nonfocal targets (M = .32, SE = .05), and this effect was large, $t(29) = 7.70$, $p < .01$, $d = 1.41$.

Retrospective memory. In the postexperimental tests, recognition memory for word targets (M = .97, SE = .02) and for letter targets (M = .95, SE = .02) did not differ ($t < 1$) and was near ceiling. When queried, 28 of the 30 participants in the PM group remembered the correct PM action, which likely underestimates true action recall. The two participants who failed to recall the action clearly showed PM performance above zero; it can therefore be assumed that participants understood the PM task, and retrospective memory demands were similarly low for focal and for nonfocal targets.

Ongoing-task performance. We calculated split-plot ANOVAs with the factors *target type* (focal vs. nonfocal) and *group* (PM vs. No-PM) on response latencies and accuracy in the ongoing LDT (Table 1). Response latencies were higher with a nonfocal PM task, $F(1, 58) = 12.75$, $p < .01$, $\eta_p^2 = .18$, and in the group with an embedded PM task, $F(1, 58) = 9.82$, $p < .01$, $\eta_p^2 = .15$. Both factors were qualified by an interaction, $F(1, 58) = 16.69$, $p < .01$, $\eta_p^2 = .22$. Single comparisons within the PM group revealed significantly higher RTs with nonfocal targets than with focal targets, $t(29) = 4.42$, $p < .01$, $d = 0.81$. As expected, no RT differences occurred within the No-PM group, in which participants only memorized these targets without an associated intention, $t < 1$. As suggested by previous research (Einstein et al., 2005; Scullin et al., 2010), nonfocal initial-letter targets may require higher degrees of monitoring for successful PM than focal word targets that are routinely processed in the service of an LDT. However, in comparison with the No-PM control group, performing the PM task increased the latencies in both the focal condition and the nonfocal condition. The RT results are in line with MPV in that the magnitude of interference is reduced in the focal condition relative to the nonfocal condition. The fact that task interference was not completely eliminated in the focal condition may be due to the use of multiple targets and neutral instructions that did not emphasize the ongoing task (Einstein & McDaniel, 2010).

The proportion of correct responses was slightly higher when the PM group performed the LDT with embedded nonfocal PM targets, but these effects were not statistically reliable, largest $F(1, 58) = 1.23$, smallest $p = .27$.

Diffusion model results. We fit diffusion models to the ongoing-task data separately for each participant and each block, yielding 60×2 submodels. Each submodel was based on 456 nontarget trials minus the excluded responses (0.83% of the nontarget trials). The parameter estimates for boundary separation a, nondecision time T_{er}, and drift rate v, are in the bottom panels of Figure 3.

Boundary separation a was larger in the nonfocal PM condition and with an embedded PM task, as indicated by main effects of target type, $F(1, 58) = 8.03$, $p < .01$, $\eta_p^2 = .12$, and group, $F(1, 58) = 9.40$, $p < .01$, $\eta_p^2 = .14$, which were qualified by an interaction of both factors $F(1, 58) = 8.29$, $p < .01$, $\eta_p^2 = .13$. Single comparisons showed that participants adopted more conservative speed-accuracy settings within the PM group in the block with nonfocal PM targets than with focal targets, $t(29) = 3.45$, $p < .01$, $d = 0.63$, whereas no such boundary shift emerged within the No-PM group, $t < 1$. Participants may have anticipated more difficult target detection for the nonfocal PM task (initial letters) and thus performed the LDT more cautiously, as reflected in the increase of boundary separation. In line with our previous experiments, more cautious speed-accuracy settings contributed to the PM interference effect in the LDT. Compared with the No-PM control group, performing the PM task led to an increase of boundary separation in both the nonfocal block, $t(46.37) = 3.27$, $p < .01$, $d = 0.84$, as well as the focal block, $t(51.68) = 2.70$, $p < .01$, $d = 0.70$.

Nondecision time T_{er} was longer with nonfocal PM targets than with focal PM targets, $F(1, 58) = 4.84$, $p = .03$, $\eta_p^2 = .08$, whereas the main effect of adding a PM task failed to reach significance, $F(1, 58) = 3.72$, $p = .06$. There was an interaction of both factors, $F(1, 58) = 8.34$, $p = .01$, $\eta_p^2 = .13$. Paired t tests indicated that nondecision time in the PM group was higher with nonfocal targets than with focal targets, $t(29) = 3.03$, $p < .01$, $d = 0.55$, and no such effect emerged in the No-PM group, $t < 1$. In contrast to the previous experiments, nondecision time of the PM group was longer in only one of the conditions with an embedded PM task: Compared with the No-PM control group, performing the PM task increased nondecision time with nonfocal targets, $t(58) = 2.47$, $p = .02$, $d = 0.64$, but not with focal targets $t(58) = 1.08$, $p = .29$. Thus, parameter T_{er} contributed to the interference effect of PM only in the nonfocal condition.

Again, no effects on drift rate emerged. There was a tendency for drift rates to be lower in the PM group than in the No-PM group, but this effect was not statistically reliable, $F(1, 58) = 2.75$, $p = .10$. Finally, the variability parameters did not differ as a function of group or target type (Table 2). There were no significant effects on the ranges of nondecision time s_t, starting point s_z (all Fs < 1), or drift rate variability η, Fs $(1, 58) < 2.64$, ps > .10.

In sum, Experiment 3 replicated expected findings with focal/nonfocal targets at the behavioral data level (Einstein et al., 2005; Kliegel et al., 2008). The diffusion-model analysis showed that the effects of different target types on RTs are explained by underlying changes in boundaries and nondecision time, but not in drift rates. In comparison with the No-PM control group, performing a PM task with nonfocal targets had the same effects on boundary separation and nondecision time as in the previous experiments. In the focal condition, a qualitatively different pattern emerged: boundary separation alone accounted for increased latencies when a PM task was added. The MPV proposes that focal tasks discourage the use of monitoring (Einstein & McDaniel, 2010). The results for nondecision time may indicate that this parameter is influenced by explicit strategic monitoring, which would not be expected in the focal task.

Model fit. None of 60 KS tests for the models for each participant indicated significant deviations from empirical RT distributions ($\alpha = .05$), with a mean p value of .87. Predicted mean RT and accuracy for each individual correlated almost perfectly with observed mean RT ($r = .99$) and accuracy ($r = .98$). Across all participants and blocks ($n = 120$), the mean absolute deviation in RT was 5 ms and 0.3 % in accuracy. Displays of model fit at the aggregate level (bottom panels of Figure 5) indicate acceptable agreement between the predicted and empirical RT distributions.

General Discussion

We applied Ratcliff's (1978) diffusion model in this study to disentangle the processes contributing to the cost- or interference-effect of event-based PM (Hicks et al., 2005; Smith, 2003). To date, ongoing-task performance has been extensively analyzed to assess the resource demands of PM (Einstein et al., 2005; Marsh et al., 2003; Smith et al., 2007). However, the conclusions typically relied on mean RT in isolation. A notable limitation of such analyses is that RT and accuracy can show divergent patterns. For instance, many PM studies employing an ongoing LDT have found increases in RT when participants performed a PM task, but no differences in accuracy (Marsh et al., 2003; Smith, 2003; Smith et al., 2007). The differences in RTs suggest that PM can have strong impact on information processing in the ongoing task, while the accuracy data suggest that such impact is small or nonexistent. These seemingly inconsistent results can be addressed through the application of models that relate the two dependent measures – RT and accuracy – to the underlying processes (e.g., Ratcliff et al., 2006). Within the framework of an explicit process model, such as the diffusion model, there are separable possible reasons for observed differences in RTs and accuracy: nondecisional processes, speed-accuracy settings (boundary separation), and the rate of information accumulation (drift rate). A diffusion-model analysis can therefore be theoretically more informative and constraining in the debate what interference effects can or cannot reveal about PM (Einstein & McDaniel, 2010; Smith, 2010).

In line with previous studies (e.g., Smith et al., 2007), ongoing-task latencies were longer in the PM conditions than in the No-PM conditions. The diffusion model indicates that two separable components, increased boundary separation and longer nondecision time (from 45 to

90 ms) contribute to this interference effect. Differences in drift rates were not observed. Across experiments, the patterns of nondecision time and boundary separation suggest that different types of processes, involved in the *prospective component* of a PM task, may influence these parameters. The diffusion model thus offers a less monolithic view on interference effects in the PM paradigm.

That is, both boundary separation and nondecision time increased with an embedded PM task in Experiment 1, but regression analyses indicated that primarily the changes of nondecision time (ΔT_{er}), and not of boundaries or of drift rate, accounted for variance in PM. Moreover, given that nondecision time was clearly influenced by task importance in Experiment 2, this component may well be affected by strategic monitoring or checking for targets (cf. Smith & Bayen, 2004). Additional support for this view comes from Experiment 3. The MPV (McDaniel & Einstein, 2000) suggests that focal tasks discourage strategic monitoring, and nondecision time was in fact reduced in the focal relative to the nonfocal condition. The possibility that these differences in nondecision time reflect differences in retrospective memory-load or in monitoring-difficulty, is unlikely, as the two target types (focal vs. nonfocal) were equated on both dimensions (cf. Scullin et al., 2010). Finally, the focal PM condition and the No-PM condition did not differ in nondecision time, consistent with MPV's proposal that focal tasks do not require an explicit monitoring strategy.

We suggest that monitoring for PM targets may affect nondecision time in two ways. As a first possibility, early stimulus encoding, which is necessary *before* a decision process can begin, may be slowed if processing resources are taken away from the ongoing task. Specifically, Smith, Ratcliff, and Wolfgang (2004; cf. Ratcliff & Smith, 2010) showed that if attention is shifted to miscued locations (i.e., away from the location of subsequent occurrence of the stimulus), this prolongs nondecision time for these unattended stimuli. Comparable attentional effects may well emerge if (a) resources are absorbed by a monitoring process, and (b) if attention must be shifted to target-features that are unrelated to the ongoing task (McDaniel & Einstein, 2007), thus accounting for the focal/nonfocal effects on nondecision time. In this way, strategic monitoring could interfere with early perceptual encoding in service of the ongoing-task decisions. A second possibility is a strategy of sequential checks for the presence of targets. When participants check whether strings are PM targets before or after their lexical decisions, they may be briefly withholding their ongoing-task decisions or responses, respectively. In fact, Scullin et al. (2010) found that a majority of individuals in a monitoring experiment adopted a relatively constant strategy: Approximately half of the participants made their target decisions first, whereas the other half preferred to make their lexical decisions first. Such a strategy would likely map on the nondecision component of RT, which captures the processes before and after the actual decision phase. In related vein, Klauer et al. (2007) suggested that task-set switching and motor response-coordination can affect nondecision time. However, such processes, which may be involved in active monitoring as well (Bisiacchi, Schiff, Ciccola, & Kliegel, 2009), are less pronounced in the present model analyses, because we excluded all PM target- and posttarget-trials.

Recently, several studies examined the attentional demands of the prospective component with focal targets (e.g., Kliegel et al., 2008; Scullin et al., 2010; Smith et al., 2007). Do the current findings have implications for the resource-demands associated with focal tasks? The results of the third experiment indicate that strategic monitoring is reduced with focal relative to nonfocal targets. At first glance, the lack of a difference[7] in nondecision time between the focal PM and No-PM condition might appear problematic for the PAM theory (Smith, 2008, 2010) if one incorrectly equated strategic monitoring and preparatory attentional processes. However, Smith et al. (2007) suggest that the nature of preparatory attentional processes can include strategic monitoring (as described above), but is likely to include more subtle shifts in the way a task is approached overall. That is, a feature of preparatory attentional processes is that they establish a state in which the individual is prepared, in a general sense, to make a response that differs from the ongoing task response. In this view, preparatory attentional processes can include, but are not equated with a checking or monitoring for target events. Increased boundary separation in the PM compared to No-PM conditions may thus reflect a state of readiness to perform the PM task (Smith, 2008; see also Guynn's, 2003, 2008, two-process model of monitoring). Similarly, shifts of the boundaries are consistent with the proposal that participants are guided by strategic allocation policies when they approach a given task-set (Hicks et al., 2005; Marsh et al., 2005). As outlined in the introduction, our hypothesis was that additional PM instructions may increase the perceived complexity how to perform the ongoing task, thereby inducing boundary shifts (Ratcliff & Van Dongen, 2009). Similarly, participants may also anticipate more difficult target detection with nonfocal relative to focal targets and increase their boundaries (cf. Einstein & McDaniel, 2008, 2010).

The results across experiments indicate a robust and consistent effect on the boundaries when participants receive PM instructions. An interesting implication of these findings is that ongoing-task accuracy can increase with an additional PM task, because the ongoing activities are approached with more cautiousness. Such effects may have been masked in previous research, because most of the variability in LDT-performance appeared in the RTs (e.g., Marsh et al., 2003; Smith et al., 2007). In the diffusion model, however, even smaller effects in error rates can imply larger criterion-shifts if accuracy is generally high (i.e., over .90), and such effects may only be detectable with a process model, but not with traditional analyses (Ratcliff, Spieler, & McKoon, 2000).

On the whole, the results provide new information regarding the processes that contribute to the interference effect seen in many PM studies. That is, critical characteristics of PM – task importance and nonfocal targets – increase ongoing-task RTs via speed-accuracy settings and nondecision time; and exactly these processes change when a PM task must be remembered. Our results resemble findings from aging studies with the diffusion model, in which the effects of aging mapped on boundary separation and nondecision time (but not on drift rates) in the LDT and many other paradigms (Ratcliff, Thapar, et al., 2004; Ratcliff et al., 2000, 2006). The diffusion model analysis helped us localize these components of task interference more precisely and provides information that was not gleaned from the standard analyses of mean RT or accuracy.

References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Bisiacchi, P. S., Schiff, S., Ciccola, A., & Kliegel, M. (2009). The role of dual-task and task-switch in prospective memory: Behavioural data and neural correlates. *Neuropsychologia, 47*, 1362–1373.

Breneiser, J. E., & McDaniel, M. A. (2006). Discrepancy processes in prospective memory retrieval. *Psychonomic Bulletin & Review, 13*, 837–841.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*, 153–178. doi: 10.1016/j.cogpsych.2007.12.002

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Cohen, A.-L., Jaudas, A., & Gollwitzer, P. M. (2008). Number of cues influences the cost of remembering to remember. *Memory & Cognition,* 36, 149–156. doi: 10.3758/MC.36.1.149

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409-422). Amsterdam: Elsevier.

Craik, F. I. M. & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and Cognitive Processes* (pp. 191-211). New York: Plenum.

Criss, A. H. (2010). Differentiation and Response Bias in Episodic Memory: Evidence From Reaction Time Distributions. *Journal of Experimental Psychology: Learning, Memory,* *and Cognition, 36*, 484–499. doi: 10.1037/a0018435

Crovitz, H. F., & Daniel, W. F. (1984). Measurements of everyday memory – Toward the prevention of forgetting. *Bulletin of the Psychonomic Society, 22*, 413–414.

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 717–726.

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 115–141). Mahwah, NJ: Erlbaum.

Einstein, G. O., & McDaniel, M. A. (2008). Prospective memory and metamemory: The skilled use of basic attentional and memory processes. In A. S. Benjamin & B. Ross (Eds.), *The Psychology of Learning and Motivation*, Vol. 48 (pp. 145–173). San Diego, CA: Elsevier.

Einstein, G. O., & McDaniel, M. A. (2010). Prospective memory and what costs do not reveal about retrieval processes: A commentary on Smith, Hunt, McVay, and McConnell (2007). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1082–1088. doi: 10.1037/a0019184

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., et al. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General, 134*, 327–342. doi: 10.1037/0096-3445.134.3.327

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. doi: 10.3758/brm.41.4.1149

Graf, P., & Uttl, B. (2001). Prospective memory: A new focus for research. *Consciousness and Cognition, 10*, 437–450. doi: 10.1006/ccog.2001.0504

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology, 38*, 245–256.

Guynn, M. J. (2008). Theory of strategic monitoring: Instantiating a retrieval mode and periodic target checking. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 53–76). Mahwah, NJ: Erlbaum.

Harris, J. E. (1984). Remembering to do things: A forgotten topic. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory, actions, and absentmindedness* (pp. 71-92). London: Academic Press.

Hicks, J. L., Marsh, R. L., & Cook, G. I. (2005). Task interference in time-based, event-based, and dual intention prospective memory conditions. *Journal of Memory and Language, 53*, 430–444. doi: 10.1016/j.jml.2005.04.001

Horn, S. S., Bayen, U. J., & Smith, R. E. (in press). What can the diffusion model tell us about prospective memory? *Canadian Journal of Experimental Psychology*.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93*, 353–368. doi: 10.1037/0022-3514.93.3.353

Kliegel, M., Jäger, T., & Phillips, L. H. (2008). Adult age differences in event-based prospective memory: a meta-analysis on the role of focal versus nonfocal cues. *Psychology and Aging, 23*, 203–208. doi: 10.1037/0882-7974.23.1.203

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2001). Varying the importance of a prospective memory task: Differential effects across time- and event-based prospective memory. *Memory, 9*, 1–11.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2004). Importance effects on performance in event-based prospective memory tasks. *Memory, 12*, 553–561. doi: 10.1080/09658210344000099

Kolmogorov, A. (1941). Confidence limits for an unknown distribution function. *Annals of Mathematical Statistics, 12*, 461–463.

Laming, D. R. J. (1968). *Information theory of choice reaction time*. New York: Wiley.

Lehrl, S. (1995). *Mehrfachwahl-Wortschatz-Test (MWT-B).* [Multiple Choice Vocabulary Test]. Erlangen: Straube.

Luce, R. D. (1986). *Response times*. New York: Wiley.

Marsh, R. L., & Hicks, J. L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 336–349.

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 68–75. doi: 10.1037/0278-7393.31.1.68

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2006). Task interference from prospective memories covaries with contextual associations of fulfilling them. *Memory & Cognition, 34*, 1037–1045.

Marsh, R. L., Hicks, J. L., Cook, G. I., Hansen, J. S., & Pallos, A. L. (2003). Interference to ongoing activities covaries with the characteristics of an event-based intention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 861–870. doi: 10.1037/0278-7393.29.5.861

Marsh, R. L., Hicks, J. L., & Watson, V. (2002). The dynamics of intention retrieval and coordination of action in event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 652–659. doi: 10.1037//0278-7393.28.4.652

Maylor, E. A. (1996). Age-related impairment in an event-based prospective-memory task. *Psychology and Aging, 11*, 74–78.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14*, S127–S144.

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Thousand Oaks, CA: Sage.

McDaniel, M. A., & Einstein, G. O. (in press). The neuropsychology of prospective memory in normal aging: A componential approach. *Neuropsychologia.*

McDaniel, M. A., Guynn, M. J., Einstein, G. O., & Breneiser, J. (2004). Cue-focused and reflexive-associative processes in prospective memory retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition, 30*, 605–614. doi: 10.1037/0278-7393.30.3.605

McNerney, M., & West, R. (2007). An imperfect relationship between prospective memory and the prospective interference effect. *Memory & Cognition, 35*, 275–282.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7,* 134–140. doi:

10.1016/s1364(03)00028-7

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519–533.

Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 269-310). Cambridge, MA: MIT Press.

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

Perea, M., Rosa, E., & Gomez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics, 67*, 301–314.

Pew, R. W. (1969). The speed-accuracy operating characteristic. *Acta Psychologica, 30*, 16–26.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111*, 159–182. doi: 10.1037/0033-295x.111.1.159

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*, 873–922.

Ratcliff, R., Philiastides, M. G., & Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 6539–6544. doi: 10.1073/pnas.0812589106

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*, 347–356.

Ratcliff, R. & Smith, P. L. (2010). Perceptual discrimination in static and dynamic noise: The temporal relation between perceptual encoding and decision making. *Journal of Experimental Psychology: General, 139*, 70–94.

Ratcliff, R., Spieler, D., & McKoon, G. (2000). Explicitly modeling the effects of aging on response time. *Psychonomic Bulletin & Review, 7*, 1–25.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19*, 278–289. doi: 10.1037/0882-7974.19.2.278

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin & Review, 13*, 626–635.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9*, 438–481.

Ratcliff, R., & Van Dongen, H. P. A. (2009). Sleep deprivation affects multiple distinct cognitive processes. *Psychonomic Bulletin & Review, 16,* 742–751. doi: 10.3758/pbr.16.4.742

Schönpflug, W. (1969). N-Gramm Häufigkeiten in der deutschen Sprache. I. Monogramme und Digramme. [N-gram frequency in German. I. Monograms and digrams]. *Zeitschrift für experimentelle und angewandte Psychologie, 16*, 157–183.

Schouten, J. F., & Bekker, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica, 27*, 143–153.

Scullin, M. K., McDaniel, M. A., Shelton, J. T., & Lee, J. H. (2010). Focal/nonfocal cue effects in prospective memory: Monitoring difficulty or different retrieval processes? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 736–749. doi: 10.1037/a0018971

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal-lobe damage in man. *Brain, 114*, 727–741.

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 347–361. doi: 10.1037/0278-7393.29.3.347

Smith, R. E. (2008). Connecting the past and the future: Attention, memory, and delayed intentions. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 27–50). Mahwah, NJ: Erlbaum.

Smith, R. E. (2010). What costs do reveal and moving beyond the cost debate: Reply to Einstein and McDaniel (2010). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1089–1095.

Smith, R. E., & Bayen, U. J. (2004). A multinomial model of event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 756–777. doi: 10.1037/0278-7393.30.4.756

Smith, R. E., Hunt, R. R., McVay, J. C., & McConnell, M. D. (2007). The cost of event-based prospective memory: Salient target events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 734–746. doi: 10.1037/0278-7393.33.4.734

Smith, P. L., Ratcliff, R. & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision Research*, 44, 1297–1320.

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 101–117. doi: 10.1037/0278-7393.1.101

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging, 25,* 377–390. doi: 10.1037/a0018022.

Stoltzfus, E. R., Hasher, L., & Zacks, R. T. (1996). Working memory and aging: Current status of the inhibitory view. In J. T. E. Richardson, R. W. Engle, L. Hasher, R. H. Logie, E. R. Stoltzfus, & R. T. Zacks (Eds.), *Working memory and human cognition* (pp. 66–88). Oxford: University Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*, 550–592. doi: 10.1037//0033-295x.108.3.550

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32*, 1206–1220.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39*, 767–775.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52*, 1–9. doi: 10.1016/j.jmp.2007.09.005

Wagenmakers, E. J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive*

Psychology, 21, 641–671. doi: 10.1080/09541440802205067

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58*, 140–159. doi: 10.1016/j.jml.2007.04.006

Wagenmakers, E. J., van der Maas, H. L. J., & Grasman, R. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14*, 3–22.

West, R., Krompinger, J., & Bowry, R. (2005). Disruptions of preparatory attention contribute to failures of prospective memory. *Psychonomic Bulletin & Review, 12*, 502–507.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology, 54*, 39–52.

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 547–565.

Footnotes

[1] Characteristically, the target events occur only on a small fraction of all trials. Thus, the secondary task must be performed very rarely, in contrast to dual-task or task-switching paradigms (Monsell, 2003; Pashler, 1994). In this line, it has been suggested that the degree to which an intention is represented in working-memory distinguishes “vigilance” tasks from “PM proper” tasks (cf. Graf & Uttl, 2001).

[2] Accumulation within a given trial with drift ξ varies randomly, following a Gaussian distribution $N(\xi,s)$. The standard deviation s governs the amplitude of the random noise (the diffusion coefficient) and is fixed at 1 for all present model analyses as a scaling parameter. Different values of s would rescale the absolute values of some of the parameters, but would not change their relations.

[3] Although accuracy is generally high in the LDT, a number of factors could have increased accuracy in previous applications in the PM paradigm even further. Typically, PM studies used sets with high-frequency or medium-frequency words only, and self-paced responding (e.g., Marsh, et al. 2002; Smith et al., 2007). In combination with smaller numbers of trials, this could have contributed to near-ceiling effects, which renders a diffusion-model analysis problematic.

[4] Latencies below the first quartile minus three times the interquartile-range or above the third quartile plus three times the interquartile-range were defined as extreme RT outliers (Tukey, 1977), separately for each phase and each participant. The additional absolute cutoff was imposed (cf. Ratcliff, Gomez, et al., 2004), because a few aberrantly fast or slow responses from individuals whose overall speed was relatively fast or slow would have remained in the data set otherwise.

[5] Supplemental model analyses with drift rates free to vary between the three item types replicated typical effects of lexical strength (i.e., word frequency and wordlikeness) on drift rates (Ratcliff, Gomez, et al., 2004), but did not change the basic conclusions concerning the PM task. Specifically, separate drift rates for the three item types in the No-PM control groups were similar in magnitude to those from Ratcliff, Gomez, et al.'s experiments 1 and 5.

[6] Power analyses were conducted with the program G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009).

[7] Note that the focus of Experiment 3 was to assess focal/nonfocal effects, and not whether interference-effects on parameters can completely be eliminated with focal targets. The interpretation of such null-effects is problematic because it can require unrealistically large sample sizes. That is, to detect smaller effects ($d = 0.2$; Cohen, 1988) on nondecision time in Experiment 3 with sufficient power of .80 would have required $N = 620$ participants (one-sided test). Our achieved power was .80 to detect medium-large effects ($d = 0.65$), and .92 to detect large effects ($d = 0.8$) on nondecision time, respectively. Thus, smaller effects on nondecision time cannot fully be excluded, and this parameter was numerically increased in the focal PM condition relative to the No-PM condition.

Table 1

Mean Latencies and Accuracy in the Ongoing Lexical Decision Task as a Function of Group (PM vs. No-PM Control) and Condition in Experiments 1 to 3.

Experiment and Condition	Response Time			Accuracy		
	PM	No-PM	$t(df)$ [a]	PM	No-PM	$t(df)$ [a]
	M (*SE*)	*M* (*SE*)		*M* (*SE*)	*M* (*SE*)	
Experiment 1						
Phase 1	745 (24)	742 (32)	0.08	.85 (.01)	.82 (.02)	1.51
Phase 2	880 (36)	722 (29)	3.44**	.86 (.01)	.82 (.02)	1.92
Experiment 2						
OI	903 (35)	783 (27)	2.71**	.88 (.01)	.82 (.01)	3.31**
PMI	986 (43)	796 (29)	3.61**	.89 (.01)	.82 (.01)	4.00**
Experiment 3						
Focal	772 (26)	696 (22)	2.23*	.83 (.02)	.82 (.01)	0.57
Nonfocal	855 (38)	690 (24)	3.66**	.84 (.02)	.82 (.02)	1.04

Notes. Values in parentheses are standard errors. Response times are in milliseconds. Accuracy = proportion of correct responses; OI = Ongoing Task Important; PMI = Prospective Memory Task Important.

[a] Independent-samples test between PM group and No-PM group; $df = 45$ (Exp. 1); $df = 73$ (Exp. 2); $df = 58$ (Exp. 3); * $p < .05$; ** $p < .01$

Table 2

Variability Parameters of the Diffusion Model as a Function of Group (PM vs. No-PM Control) and Condition in Experiments 1to 3.

Experiment and Condition	s_z		s_t		η	
	PM	No-PM	PM	No-PM	PM	No-PM
	M (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Experiment 1						
Phase 1	.29 (.02)	.27 (.02)	0.21 (0.01)	0.21 (0.02)	0.34 (0.05)	0.31 (0.05)
Phase 2	.31 (.02)	.31 (.01)	0.28 (0.04)	0.19 (0.02)	0.33 (0.06)	0.31 (0.03)
Experiment 2						
OI	.32 (.02)	.27 (.01)	0.28 (0.03)	0.23 (0.01)	0.26 (0.03)	0.35 (0.04)
PMI	.29 (.02)	.28 (.02)	0.32 (0.03)	0.22 (0.01)	0.26 (0.03)	0.34 (0.03)
Experiment 3						
Focal	.31 (.02)	.30 (.01)	0.21 (0.01)	0.20 (0.01)	0.39 (0.04)	0.35 (0.03)
Nonfocal	.29 (.02)	.30 (.01)	0.22 (0.02)	0.20 (0.01)	0.33 (0.04)	0.30 (0.03)

Notes. The scaling parameter *s* was fixed at 1. Parameter s_z was divided by the maximum range of *z* (i.e., by parameter *a*).

Figure 1. Illustration of the diffusion model. The vertical axis is the decision axis, and the horizontal axis is the time axis. Diffusion processes start at point *z* and move over time until the upper boundary (positioned at *a*) or the lower boundary (positioned at zero) is reached. Decision time distributions for decisions associated with the upper boundary (A) and the lower boundary (B) are shown. The different sizes of the shaded areas indicate that processes reach the upper boundary in most cases, reflecting the positive drift ($v > 0$). Observed response time (RT) is the sum of the decision time and a nondecision component (T_{er}) that represents the duration of processes such as perceptual encoding and response execution.

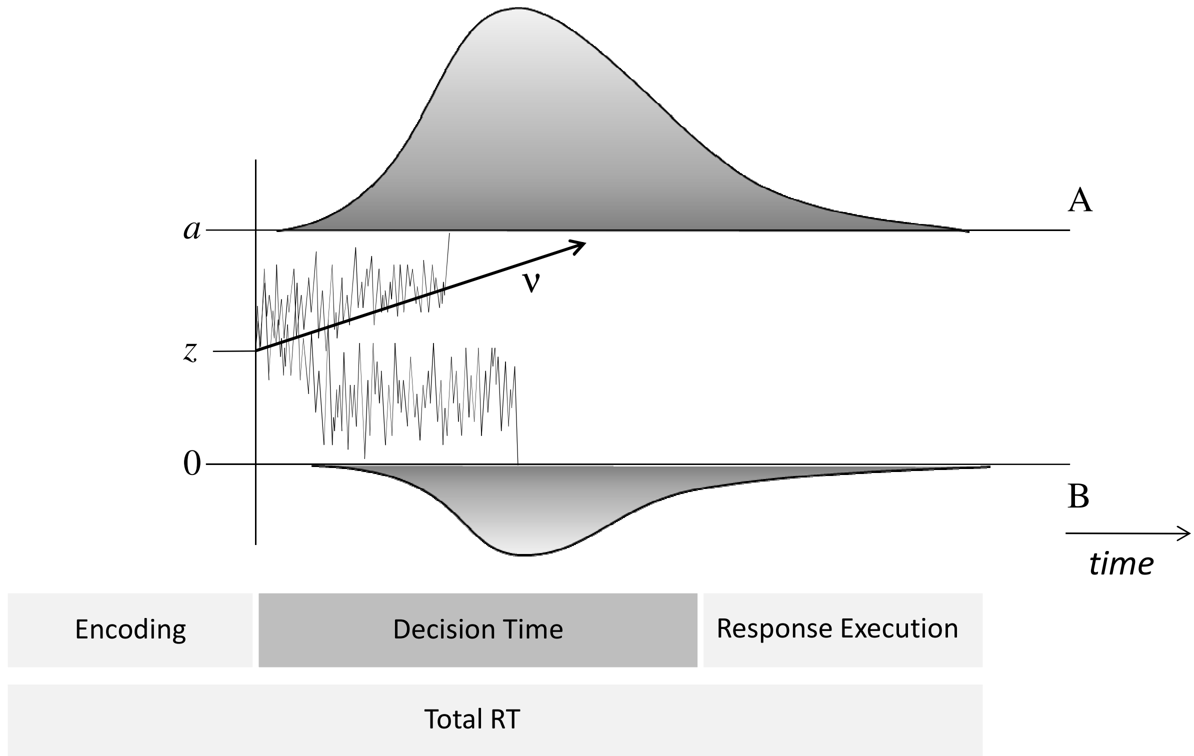


Figure 2. Individual change in mean RT and mean accuracy for participants in the PM condition and No-PM Condition (Experiment 1). Individual difference scores were calculated by subtracting Phase-1 means from Phase-2 means.

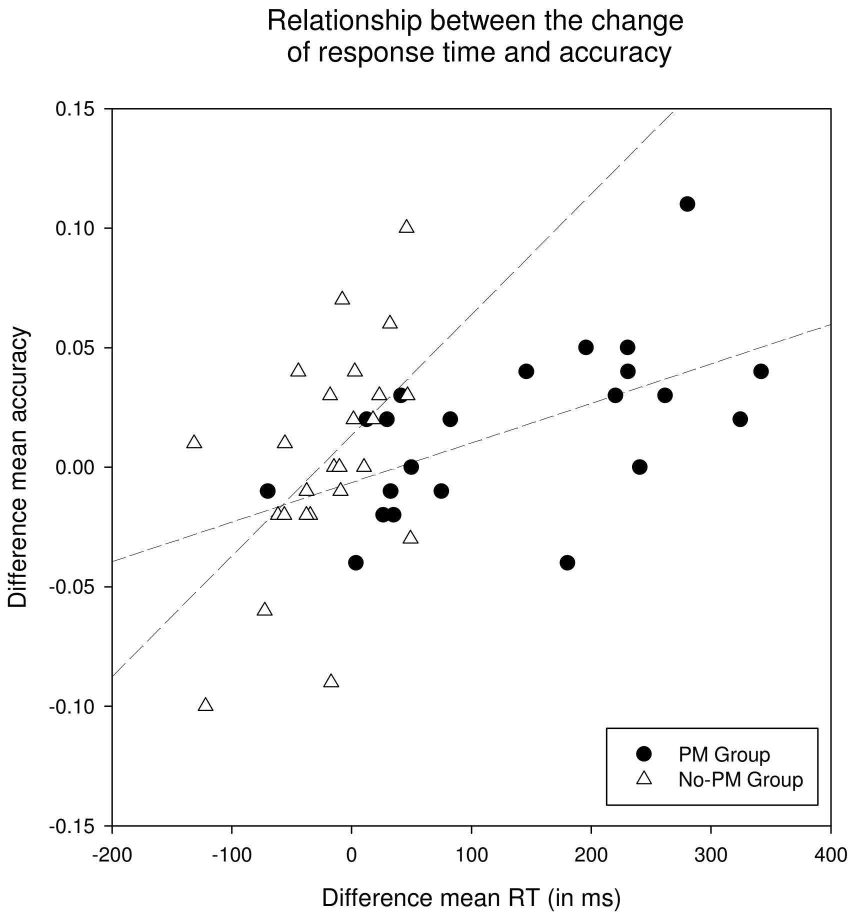


Figure 3. Mean estimates of the diffusion model parameters in Experiments 1 to 3 as a function of group (PM vs. No-PM) and condition. Error bars represent standard errors. The scaling parameter *s* is fixed at 1 for all model analyses. PM = Prospective Memory; OI = Ongoing Task Important; PMI = Prospective Memory Task Important; RT = Response Time.

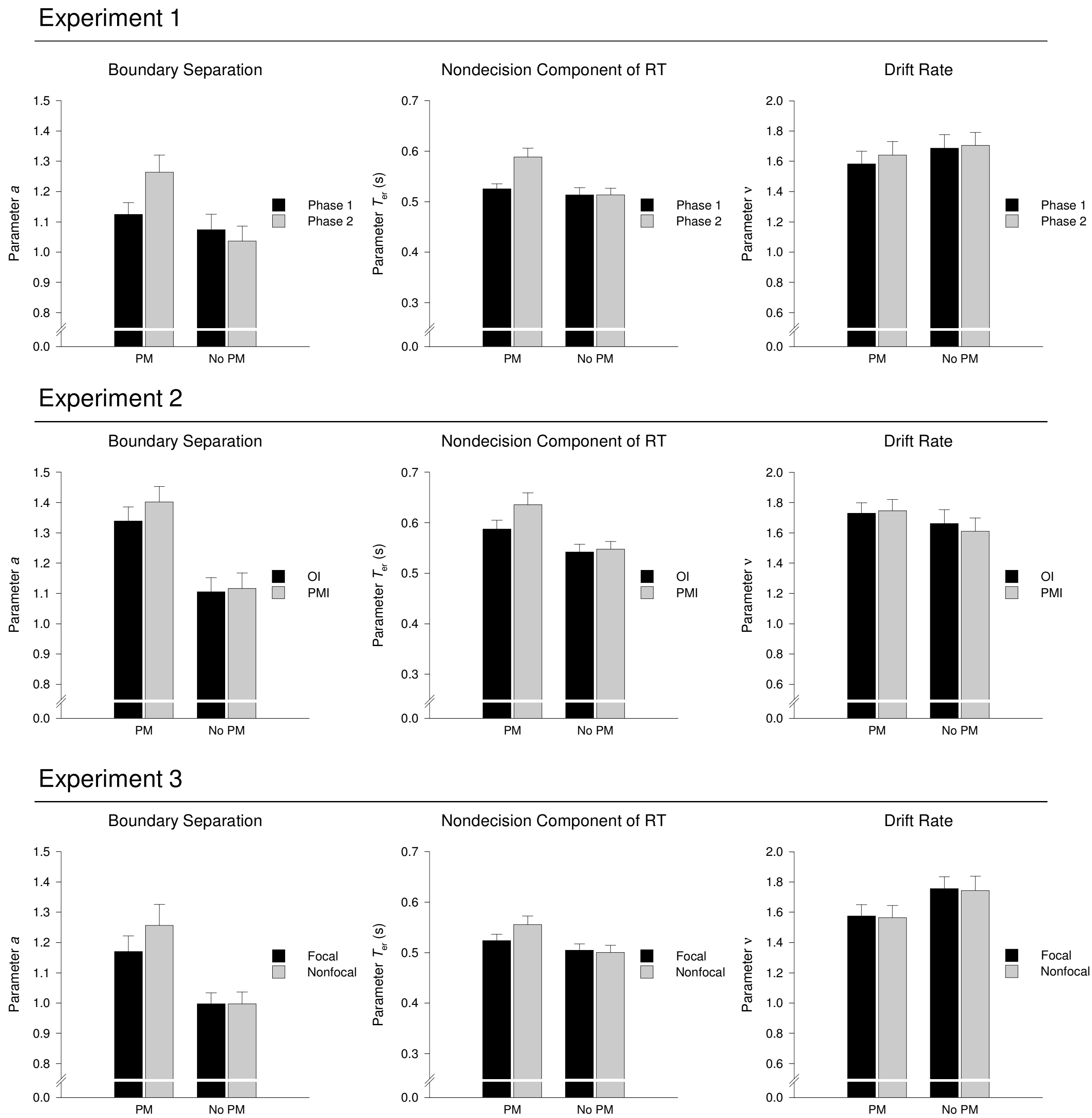


Figure 4. Regression of PM performance on the change in boundary separation Δa, nondecision time ΔT_{er}, and drift rate Δv between Phase 1 (Ongoing task only) and Phase 2 (Ongoing task + embedded PM task) of Experiment 1 (PM group). Beta values represent standardized regression coefficients. PM = prospective memory.

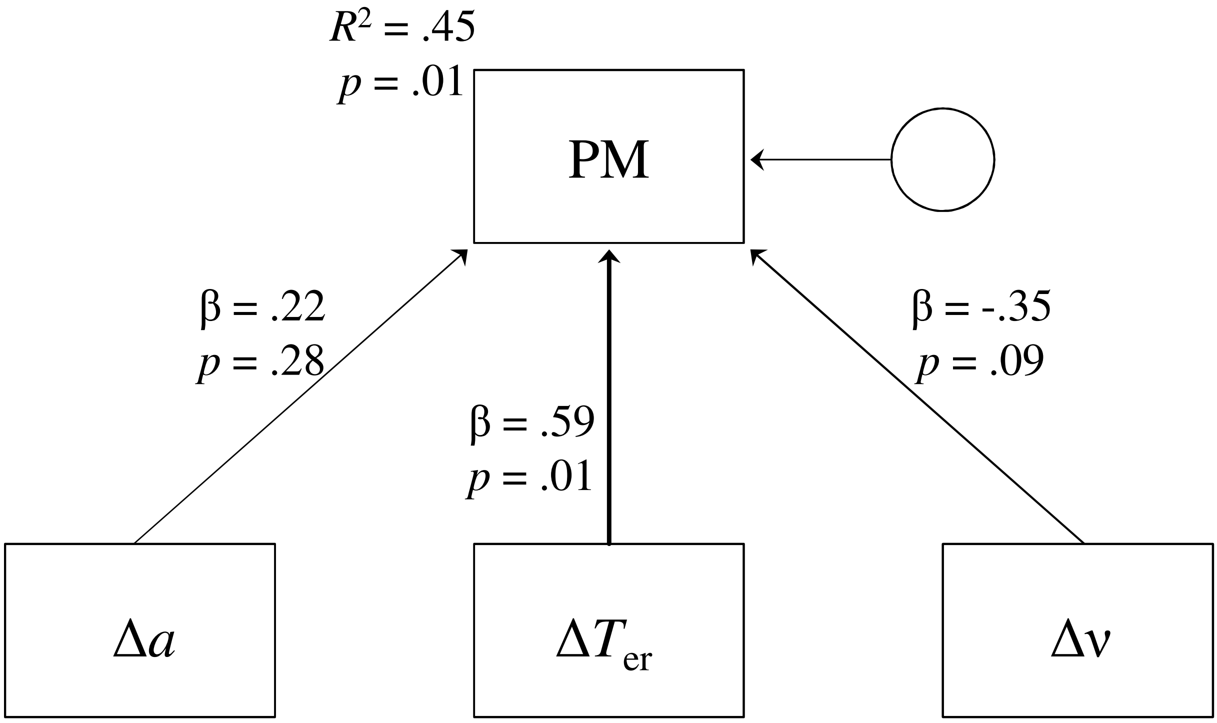
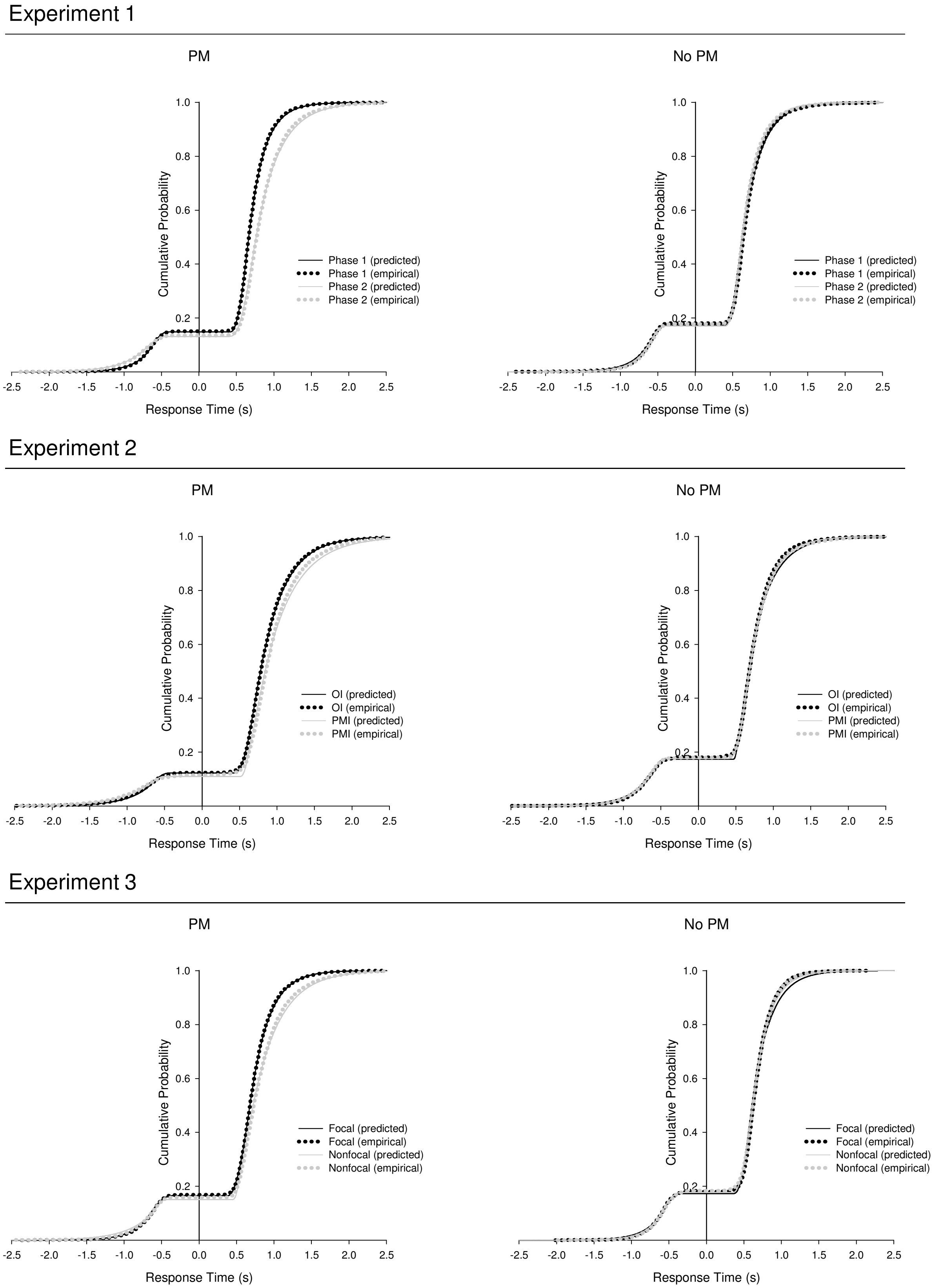


Figure 5. Model fits for Experiments 1 to 3. The graphs show the empirical (observed) and the predicted cumulative distribution functions of response time, aggregated over participants within a condition and group. Negative values on the horizontal axis are latencies of error responses (multiplied by –1), and positive values are latencies of correct responses. PM = Prospective Memory; OI = Ongoing Task Important; PMI = Prospective Memory Task Important.


Experiment 1
PM
No PM
Cumulative Probability
Response Time (s)
Phase 1 (predicted)
Phase 1 (empirical)
Phase 2 (predicted)
Phase 2 (empirical)
Experiment 2
PM
No PM
Cumulative Probability
Response Time (s)
OI (predicted)
OI (empirical)
PMI (predicted)
PMI (empirical)
Experiment 3
PM
No PM
Cumulative Probability
Response Time (s)
Focal (predicted)
Focal (empirical)
Nonfocal (predicted)
Nonfocal (empirical)
1.0
0.8
0.6
0.4
0.2
-2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5

Adult Age Differences in Interference From a Prospective-Memory Task: A Diffusion-Model Analysis

Sebastian S. Horn and Ute J. Bayen

Heinrich-Heine-Universität Düsseldorf

Rebekah E. Smith

The University of Texas at San Antonio

Author Note

Sebastian S. Horn and Ute J. Bayen, Institut für Experimentelle Psychologie, Heinrich-Heine-Universität Düsseldorf. Rebekah E. Smith, Department of Psychology, The University of Texas at San Antonio.

We thank Christina Engels, Maike Lex, and Sarah Müller for assistance with data collection. Correspondence concerning this article should be addressed to Sebastian Horn or Ute J. Bayen, Institut für Experimentelle Psychologie, Abteilung Mathematische und Kognitive Psychologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, Geb. 23.02, D-40225 Düsseldorf, Germany. E-mail: s.horn@hhu.de

Abstract

People often slow down their ongoing activities when they must remember an intended action, known as the *cost- or interference effect* of prospective memory (PM). Only few studies have examined adult age differences in PM interference, and the specific reasons underlying such differences are not well understood. The authors used a model-based approach to reveal processes underlying PM interference and age differences in these processes. Older and younger adults first performed a block of an ongoing lexical decision task alone. An embedded event-based PM task was added in a second block. Simultaneously accounting for the changes in RT distributions and error rates induced by the PM task, Ratcliff's (1978) diffusion model was used to decompose the nonlinear combination of speed and accuracy into psychologically meaningful components. The PM task increased response cautiousness and peripheral nondecision time in both age groups. Moreover, it reduced older adults' processing efficiency. In sum, the findings suggest that there are multiple, but similar reasons underlying PM-task interference in both age groups. Supplemental materials can be downloaded from *http://*

Keywords: Prospective Memory, Aging, Diffusion Model, Monitoring and Attentional Resources, Older Adults

Adult Age Differences in Interference From a Prospective-Memory Task: A Diffusion-Model Analysis

Prospective memory (PM) involves remembering intended actions after a delay, such as remembering to buy groceries after work. PM is essential for an independent lifestyle, particularly when getting older (e.g., Schnitzspahn, Ihle, Henry, Rendell, & Kliegel, 2011). PM tasks involve the retrieval of information from the past (the *retrospective* component) and also require us to remember *that* there was something we intended to do (*prospective* component). Research on PM has largely focused on how intentions come into mind at the relevant moment and whether this process requires attentional resources. To address these questions in the laboratory, the PM task and target events (e.g., "remember to press the spacebar when you see the word *tiger*") are embedded in an ongoing task (e.g., a lexical decision task; LDT). One fairly pervasive finding is the *cost-* or *interference effect of PM* (Smith, 2003). That is, response times (RTs) on nontarget trials of the ongoing task are often increased in the presence of a PM task relative to performing the ongoing task alone. It has been proposed that interference indicates resources being devoted to monitoring (e.g., Guynn, 2003) or for preparatory attentional processes (Smith, 2003). Although there is debate whether these processes are necessary for all PM tasks, theoretical views concur that the prospective component involves resource-demanding processes leading to interference effects in *nonfocal* tasks. Nonfocal tasks require a shift from the processing routine during the ongoing task towards the relevant PM-target features (see McDaniel, Einstein, & Rendell, 2008, for details) and show reliable age-related PM differences favoring young adults (e.g., Henry, MacLeod, Phillips, & Crawford, 2004). In the current study, we investigated the cognitive processes underlying task interference and how these processes may vary in young and older adults.

Although interference is an important indicator of allocation of attention towards the PM task, the specific processes giving rise to the slowing are not well understood. Thus, one important research goal is to identify processes that contribute to PM-task interference. Moreover, a limitation to previous analyses of age-related differences in PM interference stems from scaling dependency and unequal baseline performance between young and older adults (Perfect & Maylor, 2000). That is, conclusions about task-induced changes in RTs between age groups (treatment × age interactions) are not meaningful if the assumption of a common interval scale across age groups is violated, if the relationship between latent processes and RT is unclear, or if the nonlinearity of speed-accuracy functions is ignored (see Verhaeghen, 2000; Wagenmakers, Krypotos, Criss, & Iverson, 2012).

We addressed these issues by applying the diffusion model that explicitly models the nonlinear combination of speed and accuracy and does not treat RT as a process-pure variable. Prior applications of the model have shown that participants' speed-accuracy criteria often differ as a function of age, with older adults sacrificing speed for accuracy (e.g., Ratcliff, Thapar, Gomez, & McKoon, 2004). We therefore expected longer RTs and higher LDT accuracy for older relative to young adults. We also included a PM task embedded in the LDT and, based upon prior research (e.g., Gao et al., 2012; McDaniel et al., 2008; Smith & Bayen, 2006), we expected PM interference for both age groups. Our goal was to model whether the same or qualitatively different processes are responsible for PM interference in young and older adults, whether there are age-related differences in the size of these effects, and to confirm findings from previous applications of the diffusion model in PM studies with younger adults.

The Diffusion Model

The diffusion model (Ratcliff, 1978) has been successfully applied to many speeded two-choice tasks and has contributed to our understanding of the processes underlying age-related slowing in LDT (Ratcliff et al., 2004), signal detection, and recognition memory (Ratcliff, Thapar, & McKoon, 2006), among others. By decomposing accuracy and the whole RT distributions for correct and error responses into psychologically substantive components, the model provides a comprehensive and detailed approach to examining *why* RTs are fast or slow and *why* few or many errors are made. For instance, the hypothesis of a decrease in processing speed with aging (Salthouse, 1996) was refined by showing that there are separable ways that such slowing can be produced. As one major finding, older adults set more conservative speed-accuracy criteria and have longer nondecision times (including motor execution) across many tasks (e.g., Ratcliff et al., 2004, 2006). Across many applications, the diffusion model has provided a close fit to observed data, and the discriminant validity of its parameters has been demonstrated (Voss, Rothermund, & Voss, 2004).

Model Parameters

One core model assumption is that decisions are based on continuous information accumulation over time, defined by a Wiener diffusion process moving from a starting point until one of two decision boundaries is crossed (Figure 1). Accumulation within a trial is noisy, implying that diffusion processes for equivalent items will not deterministically reach the same boundary at the same time and will thus produce errors and RT distributions. The model version we used includes the following parameters.

Mean drift rate $\boldsymbol{v}$. This parameter quantifies the average speed of information uptake. Drift rate represents processing efficiency, with higher absolute values leading to both faster and more accurate decisions, thus reflecting participant ability or task difficulty.

Boundary separation $\boldsymbol{a}$. The distance a between the two decision boundaries quantifies the amount of evidence required until a decision is made, representing the speed-accuracy setting. By increasing *a,* the decision maker can trade speed for accuracy (a “cautious” setting), or vice versa.

Mean of nondecision time T_{er}. Observed RT can be additively split into decision- and nondecision time (i.e., RT = DT+T_{er}). The latter includes the processes *before and after* the actual decision phase (such as encoding and motor execution). Changes in T_{er} imply shifts of the entire RT distribution without changing accuracy.

Variability parameters η, s_z, s_t. The model assumes variation of some parameters across trials. Drift rate varies normally around mean v with standard deviation η. The starting point and nondecision time are uniformly distributed around mean z and T_{er} with ranges s_z and s_t, respectively. With inter-trial variability of drift rate η, the model can account for slower RTs of incorrect responses than of correct responses, and variability in the starting point (s_z) can account for a reverse pattern (e.g., Wagenmakers, Ratcliff, Gomez, and McKoon, 2008).

Modeling Ongoing-Task Performance in the PM Paradigm

Although various ongoing tasks used in PM research could provide appropriate data for application of the model, we used a LDT as it has been studied in detail with the diffusion model (see Ratcliff et al., 2004, Wagenmakers et al., 2008). Only a few PM studies have applied the model, and these were exclusively with young adults (Boywitt & Rummel, 2012; Horn, Bayen, & Smith, 2011; Horn, Smith, Bayen, & Voss, 2008). In the present study, we focused on age comparisons in processes underlying PM-task interference. From results with younger adults, we predicted that PM tasks induce cautiousness (boundary-separation parameter a), reflecting a shift in the approach to the ongoing task (Horn et al., 2008, 2011). Moreover, we assumed that the time before or after an ongoing-task decision (the nondecision-time parameter T_{er}) is prolonged if participants strategically check for the presence of PM targets (Horn & Bayen, 2012).

Method

Participants

Forty-six young adults (11 male; *M* =13.51 years of education) and 43 community-dwelling older adults (18 male; *M* =11.95 years of education) participated. The participants, all native German speakers, were recruited through newspaper advertisements or a department pool and received money or psychology course credit. Individuals with major health problems that may affect cognitive functioning were excluded. Older adults scored higher than young adults on verbal-semantic knowledge, but lower in the digit-symbol test and short-term memory span (details in Table 1).

Design

Participants were randomly assigned to either the PM group (22 young, 21 older), which first performed a baseline block of the LDT alone, followed by a second block with an embedded event-based PM task, or to the No-PM Control group (24 young, 22 older), which performed the LDT alone in both blocks. The design thus included the factors *Age* (young vs. older), *Experimental Group* (PM vs. Control), and *Block* (first vs. second).

Materials

We selected 252 low-frequency (*M* = 4.46 occurrences per million, *SD* = 0.50) and 252 very-low-frequency German words (*M* = 0.35, *SD* = 0.48) from the CELEX lexical database. Word length ranged from 8 to 12 letters (*M* = 9.88, *SD* = 1.27). Pronounceable pseudowords were generated from each word by randomly replacing one interior vowel/umlaut with a different vowel/umlaut. From this pool, we created two nontarget sets for the ongoing LDT and two target sets for the PM task, matched for word frequency and length. Each nontarget set contained 120 low-frequency words, 120 very-low-frequency words, and 240 pseudowords (created from the words of the other nontarget set). The two target sets each contained 6 words and 6 pseudowords with initial letters *G*, *H*, or *M*, occurring equally often within a target set. None of the nontarget items started with these letters. During the experiment, the program selected one filler and one target set, from which items were randomly drawn without replacement, under the restriction of equal numbers of words and pseudowords and of low-frequency and very-low-frequency words within each block. The possible combinations of target and distractor sets were counterbalanced across participants and conditions.

Procedure

Participants completed 24 practice trials, followed by two blocks of 252 LDT trials (50% words) each. Each trial started with a fixation cross that remained in the center of the screen for 400,450,500,550,or 600ms with fixation display times randomly selected with equal probability to prevent anticipatory responding. Subsequently, a string appeared in black upper-case letters (sans-serif font, 24pt.) in the center of a white screen. Participants were asked to categorize each string as a word or nonword as fast and accurately as possible by pressing the *J*- or *F*-key. Key-response mappings were counterbalanced across participants. A blank screen followed each response for 200ms.

After the first block, participants were informed that they should memorize some letters. Participants in the PM group were then asked to remember to press the *spacebar* (instead of the *F* or *J* key) if a string started with one of the letters *G*, *H*, or *M* in the subsequent block. In both groups, participants were informed that their memory for these letters would be tested later, and the letters were then presented in random order for 5s each. Participants reproduced these letters by typing them into the keyboard. If mistakes occurred, the letters were again displayed until perfect recall. The experiment then continued with a filler digit-span task (2min), followed by the second block. There was no further mention of the PM task.

The 12 PM target events appeared within trial ranges (starting from Trials 17-20), with 17 to 23 nontarget trials between any two targets. Finally, participants completed a questionnaire, cognitive tests, and an old-new recognition task in which the three target letters and three lures appeared one at a time in random order. Sessions lasted about 1h.

Results

Prospective Memory and Target Recognition

Young adults were better than older adults in PM performance (concurring with many laboratory studies, Henry et al., 2004) and post-task recognition of target letters (Table 1).

Ongoing-Task Performance

For analyses of ongoing-task performance, as is standard in PM research, we excluded PM responses (hits, false alarms) and the two trials following these responses (e.g., Smith & Bayen, 2006). Applying the same cutoffs for the LDT as Ratcliff et al. (2004), we removed RTs smaller than 350ms or greater than 4000ms for older adults (1.1% of the data) and RTs smaller than 300ms or greater than 3000ms for young adults (0.9% of the data). Table 2 shows descriptive statistics.

Baseline performance. Older adults' lexical decisions in Block 1 were slower [$F(1,85)=35.73$, $\eta_p^2=.30$, $p< .01$] but more accurate than those of young adults [$F(1,85)= 38.15$, $\eta_p^2=.31$, $p<.01$]. As expected for the baseline block in which all participants worked on the ongoing LDT only, there were no differences (nor interactions) between experimental groups (PM, No-PM), *F*s< 1.

The effects of the PM task. Both younger and older adults' RTs increased substantially with an embedded PM task in the second block (see Table 2), and this interference effect was larger for older adults [$F(1,41)=5.05$, $\eta_p^2=.11$, $p= .03$, for the *Age* $\times$*Block* interaction in the PM groups].[1] Moreover, the slowing was positively correlated with PM performance for young [$r=.66$, $p<.01$] and older adults [$r=.86$, $p<.01$].

Variability in accuracy is usually smaller in the LDT than variability in latencies (e.g., Ratcliff et al., 2004). Older adults' accuracy was very high and not affected by PM instructions. Yet, we observed a statistically reliable increase in young adults' accuracy with a PM task, replicating previous findings (e.g., Horn & Bayen, 2012). These results emphasize the importance of simultaneously modeling speed-accuracy changes.

Diffusion-Model Analysis

Parameter estimation and model fit. We fit diffusion models separately to the data from each participant and block. For parameter estimation, we used the fast-dm program which minimizes the maximal vertical distance between the observed and predicted cumulative RT distributions (Voss & Voss, 2008). This estimation method is relatively robust against outliers and provides a useful option if trial numbers are suboptimal for quantile-based categorization of RT distributions (we also tried other estimation approaches, which yielded similar outcomes; Vandekerckhove, Tuerlinckx, & Lee, 2011).[2] To examine the impact of our PM manipulation in this first study with older adults, all model parameters were allowed to vary between blocks, reflecting the fact that no data are as of yet available that would allow strong a-priori constraints for the PM paradigm. Because lexical variables were not the focus of this study, we coded responses as correct (upper boundary) or incorrect, yielding an average measure of drift rate across words and nonwords.[3]

The models provided good individual fits to the data from older and young adults, as indicated by indices *Z* (the vertical deviance between observed and predicted CDFs; M_{older} =.02; M_{young} =.03) and *p* (the probability value of the Kolmogorov-Smirnov statistic under the null hypothesis of perfect fit; all *p*-values >.89). The supplemental materials include further details about model fits and estimates of variability parameters.

Modeling results: Baseline. We examined baseline performance (Block 1) between age and experimental groups. Figure 2 shows the distributions of best-fitting parameters. Drift rate did not differ between age groups [$F(1,85)=1.88$, $p=.18$]. However, older adults were substantially more cautious (increased boundary separation) [$F(1,85)=55.06$, $\eta_p^2=.39$, $p<.01$], and their nondecision times were longer than those of young adults [$F(1,85)=12.25$, $\eta_p^2=.13$, $p<.01$]. These results are consistent with previous research (Ratcliff et al., 2004), showing that the main source underlying adult-age differences in LDTs are not the drift rates and reveal that older adults were slower but more accurate due to more conservative speed-accuracy settings. As expected for the baseline block, there were no main effects (nor interactions) involving the experimental groups (all *F*s<1.56).

PM and drift rate. We analyzed the change in parameters of main interest between the baseline and PM blocks within participants. Young adults' drift rate (*v*) did not change with an embedded PM task [$t(21)=1.34$, $p=.20$]. Older adults' drift rate decreased [$t(20)=2.98$, $\eta_p^2=.31$, $p<.01$]. This suggests that PM interfered with the efficiency of processing the letter strings in this age group, as older adults may be forced to trade off their limited attentional resources between the ongoing and the PM task. In the control groups, drift rates did not change for young nor older adults, *t*s<1.[4]

PM and boundary separation *a*. Boundary separation increased after assignment of the PM task, confirming previous findings (e.g., Horn et al., 2008). Both young [$t(21)=4.14$, $\eta_p^2=.45$,

p<.01] and older adults [t(20)=2.96, η_p^2=.30, p<.01] responded more cautiously and required more information for ongoing decisions. Such changes were not observed in the control groups with young [t<1] or older adults [t(21)=1.76, p =.09]. There is evidence that participants perceive ongoing decisions as more complex with embedded PM tasks, inducing criterion shifts to respond more cautiously to detect PM targets (Boywitt & Rummel, 2012).

PM and nondecision time. Nondecision time (T_{er}) increased substantially with an additional PM task for young [t(21)=6.46, η_p^2=.67, p<.01] and older adults [t(20)=4.68, η_p^2=.52, p<.01]. This component is likely prolonged if participants check whether strings are PM targets *before/after* ongoing lexical decisions (see Horn et al., 2011). Consistent with Gao et al. (2012), this finding may also point to prolonged encoding- and/or motor speed .

Possibly due to practice, nondecision time decreased in the young-adult control group [t(23)=3.24, η_p^2=.31, p<.01]; this effect was not reliable for older adults [t(21)=1.84, p =.08].

Discussion

In line with previous work (Ratcliff et al., 2004, 2006), older adults' baseline lexical decisions were slower but more accurate than those of young adults due to more conservative speed-accuracy settings and longer nondecision time. By applying the diffusion model that decomposes the nonlinear combination of speed and accuracy into psychologically meaningful components (Ratcliff, 1978), we could take these age-related baseline differences into account when examining PM interference, which we found for both age groups. The embedded PM task induced cautiousness and prolonged nondecision time in both age groups to a similar extent. Thus, one portion of the slowing can be attributed to shifts in speed-accuracy criteria. The counterintuitive implication of this finding is that intervening ongoing activities may be approached more slowly but with *higher* accuracy if participants must remember an intention.

Such effects are detectable via the diffusion model, but are often masked in comparisons of average performance if accuracy is relatively high, as in the LDT (Horn et al., 2011).

The changes in nondecision time concur with Gao et al. (2012) insofar as lower motor speed may be an important component underlying PM-task interference. Also, participants may engage in intermittent target-checking before/after their ongoing lexical decisions in demanding PM tasks (cf. Guynn, 2003). In addition to increased cautiousness and prolonged nondecision time, older adults showed reduced drift rates, which points to reduced processing efficiency if a PM task is added. Overall, however, the modeling largely revealed that the same processes contribute to PM-task interference in young and older adults. The *age* $\times$ *task type* interaction, found in RTs, was *not* mirrored in the latent processes (cf. Wagenmakers et al., 2012). The general pattern is thus consistent with aging theories assuming qualitatively common mechanisms underlying interference effects (see Verhaeghen, 2000, for discussion of parallelisms in task interference).

In this study, increases in RTs with an embedded PM task were larger for older than younger adults (see also Gao et al., 2012; McDaniel et al., 2008). Other studies did not find age-related differences (Einstein, McDaniel, & Scullin, 2011), sometimes with a tendency towards larger interference for young adults (Smith & Bayen, 2006). Besides methodological caveats mentioned above,[1] a parsimonious explanation for this apparent inconsistency is that resource-allocation policies likely differ across individuals and tasks. At least with more demanding PM tasks, reduced attentional capacity may force older adults to either sacrifice ongoing-task performance for maintaining adequate levels of PM, or vice versa. Consequently, age-related differences in PM performance favoring young adults were *more* pronounced whenever interference did *not* differ between age groups (Einstein et al., 2011; Smith & Bayen, 2006).

The modeling approach provides insights not attainable with mere analyses of mean RTs.

Interesting avenues for further research may be to apply the diffusion model to investigate less demanding PM tasks or the effects of differential task emphasis.

References

Boywitt, C.D., & Rummel, J. (2012). A diffusion model analysis of task interference effects in prospective memory. *Memory & Cognition, 40,* 70–82.

Einstein, G.O., McDaniel, M.A., & Scullin, M.K. (2011). Prospective memory and aging: Understanding the variability. In M. Naveh-Benjamin & N. Ohta (Eds.), *Memory and Aging: Current Issues and Future Directions.* Philadelphia, PA: Psychology Press.

Gao, J.L., Cheung, R.T.F., Chan, Y.S., Chu, L.W., & Lee, T.M.C. (2012). Increased prospective memory interference in normal and pathological aging: different roles of motor and verbal processing speed. *Aging, Neuropsychology, and Cognition.* doi:10.1080/13825585.2012. 672948

Guynn, M.J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology, 38,* 245–256.

Henry, J.D., MacLeod, M.S., Phillips, L.H., & Crawford, J.R. (2004). A metaanalytic review of prospective memory and aging. *Psychology and Aging,19*, 27–39.

Horn, S.S., & Bayen, U.J. (2012). *Interference From Remembering to Remember: A Diffusion Model Account.* Manuscript submitted for publication.

Horn, S.S., Bayen, U.J., & Smith, R.E. (2011). What can the diffusion model tell us about prospective memory? *Canadian Journal of Experimental Psychology,65*, 69–75.

Horn, S.S., Smith, R.E., Bayen, U.J., & Voss, A. (2008). Analyzing the cost of prospective memory with the diffusion model. *International Journal of Psychology,43,*787.

McDaniel, M.A., Einstein, G.O., & Rendell, P.G. (2008). The Puzzle of inconsistent age-related declines in prospective memory: A multiprocess explanation. In M. Kliegel, M.A.

McDaniel, and G.O. Einstein (Eds.), *Prospective Memory: Cognitive, Neuroscience, Developmental, and Applied Perspectives* (pp.141–160). Mahwah, NJ: Erlbaum.

Perfect, T.J., & Maylor, E.A. (2000). Rejecting the dull hypothesis: The relation between method and theory in cognitive aging research. In T.J. Perfect & E.A. Maylor (Eds.), *Models of cognitive aging* (pp.1–18). Oxford: University Press.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19,* 278–289.

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin & Review,13,* 626–635.

Salthouse, T.A. (1996). The processing speed theory of adult age differences in cognition. *Psychological Review,103*, 403–428.

Schnitzspahn, K.M., Ihle, A., Henry, J.D., Rendell, P.G., & Kliegel, M. (2011). The age-prospective memory-paradox: An exploration of possible mechanisms. *International Psychogeriatrics, 23,* 583–592.

Smith, R.E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 347–361.

Smith, R.E., & Bayen, U.J. (2006). The source of adult age differences in event-based prospective memory: A multinomial modeling approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 623–635.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M.D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods, 16,* 44–62.

Verhaeghen, P. (2000). The parallels in beauty's brow: Time-accuracy functions and their implications for cognitive aging theories. In T.J. Perfect and E.A. Maylor (Eds.), *Models of cognitive aging* (pp.50–86). Oxford: University Press.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32,* 1206–1220.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52,* 1–9.

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A.H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition, 40,* 145–160.

Wagenmakers, E.J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58,* 140–159.

Footnotes

[1] The supplementary online materials include a comprehensive ANOVA table. Only one three-way interaction involving *Age* emerged, namely for RTs, which should be interpreted with caution. As noted in the Introduction, it is preferable to address the potential problem of scale-dependent ("removable") interactions by focusing on latent parameters (Wagenmakers et al., 2012).

[2] Although Bayesian hierarchical parameter estimation would be a fruitful approach for future aging studies, we did not apply this approach here because numerical problems caused slow convergence for some participants.

[3] With this coding, any possible biases towards a decision alternative cancel out and do not affect the mean starting point position (i.e., z can be fixed at $a/2$). The online supplement 3 includes details.

[4] As expected, two-way *Experimental Group* × *Block* interactions emerged across all model parameters and RTs.

Table 1. *Participant Characteristics, Cognitive Tests, and Memory Performance*

Variable	Young Adults		Older Adults		
	M	*SD*	*M*	*SD*	η_p^2
Age	23.0	3.5	70.0	5.0	
Vocabulary [a]	32.2	2.2	33.1	1.5	.05 *
Processing speed [a]	88.3	13.4	61.8	11.9	.53 ***
Digit span [a]	19.6	3.4	16.3	3.2	.20 ***
Prospective memory [b]	.63	.27	.44	.31	.11 *
Post-task recognition [c]	.99	.05	.83	.33	.11 **

Note. [a] raw test scores. Vocabulary = multiple-choice vocabulary test (MWT-A; Lehrl, Merz, Burkhard, & Fischer, 1991). Processing speed (digit-symbol substitution) and digit span (combined forward/backward span) were measured with the German version of the Wechsler Adult Intelligence Scale (WIE; von Aster, Neubauer, & Horn, 2006). [b] Prospective memory (PM) was measured as the proportion of target events to which participants made a correct PM response (i.e., a spacebar press on a PM target or within the subsequent two LDT trials). Post-task recognition of letters was measured as Hit Rate minus False-Alarm Rate. η_p^2 = size of age effect; *p <.05; **p <.01; ***p <.001

Table 2. *Response Times and Accuracy in the Ongoing Lexical Decision Task*

Variable	Group	Block 1		Block 2		η_p^2
		M	(*SE*)	*M*	(*SE*)	
RT	Young					
	PM	980	(63)	1199	(70)	.66 ***
	No-PM Control	931	(60)	902	(67)	.18 *
	Older					
	PM	1312	(65)	1693	(71)	.63 ***
	No-PM Control	1350	(63)	1296	(70)	.20 *
Accuracy	Young					
	PM	.91	(.01)	.93	(.01)	.28 **
	No-PM Control	.90	(.01)	.90	(.01)	.02
	Older					
	PM	.95	(.01)	.95	(.01)	.04
	No-PM Control	.96	(.01)	.95	(.01)	.08

Note. The PM groups received instructions for the PM task after Block 1; RT= response time (ms); Accuracy= proportion of correct lexical decisions; η_p^2 = effect size for within-subject change Block2 – Block1; **p* <.05; ***p* <.01; ****p* <.001

Figure 1. Illustration of the diffusion model (Ratcliff, 1978). The vertical axis is the decision-related strength-of-evidence axis, and the horizontal axis is the time axis. Diffusion processes start at point *z* and move over time until the upper boundary (positioned at *a*) or the lower boundary (positioned at zero) is reached. Decision time distributions associated with the upper and the lower boundaries are shown. The two oscillating sample tracks illustrate that different boundaries can be reached with the same (positive) drift rate due to random influence, as given by the diffusion constant *s* (which is set to 1 in all present analyses). Total RT is the sum of the decision time and a nondecision component T_{er} that represents the duration of processes such as encoding and motor response execution.

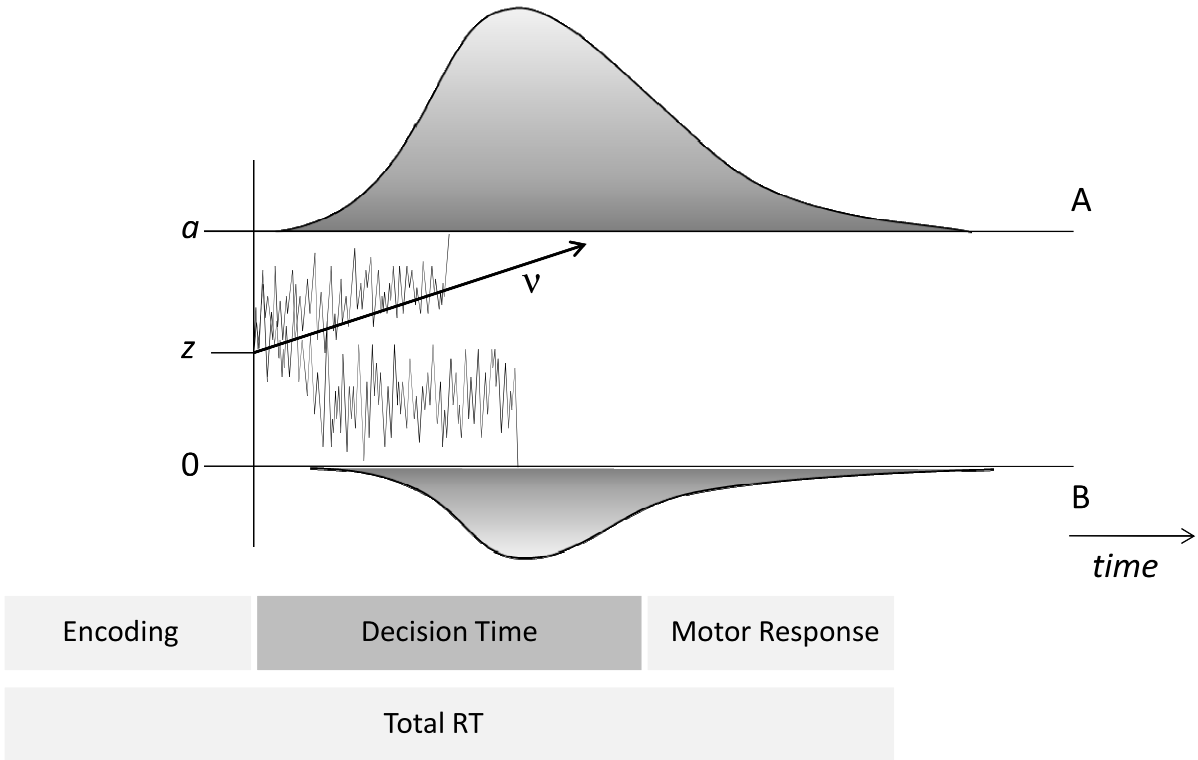
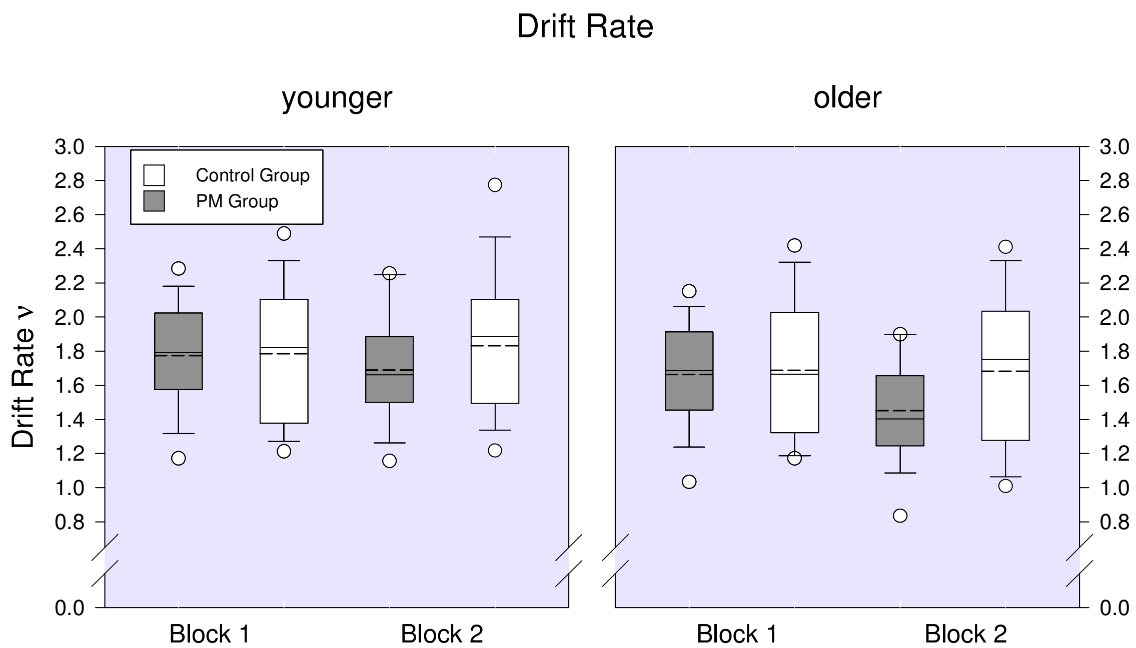
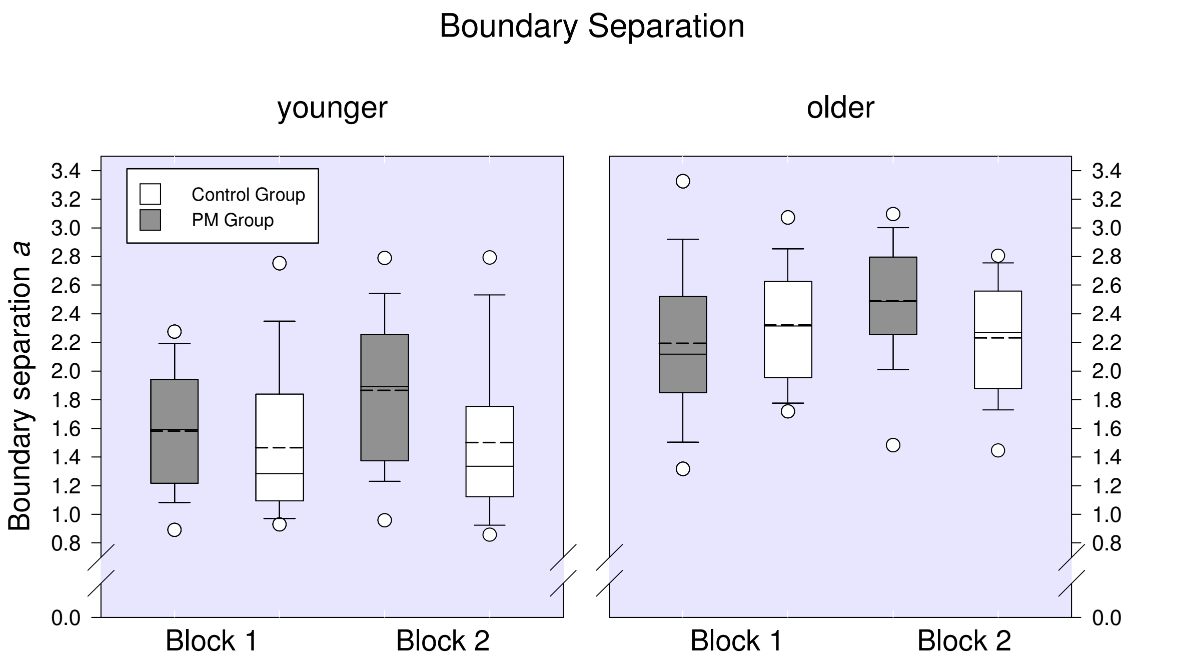
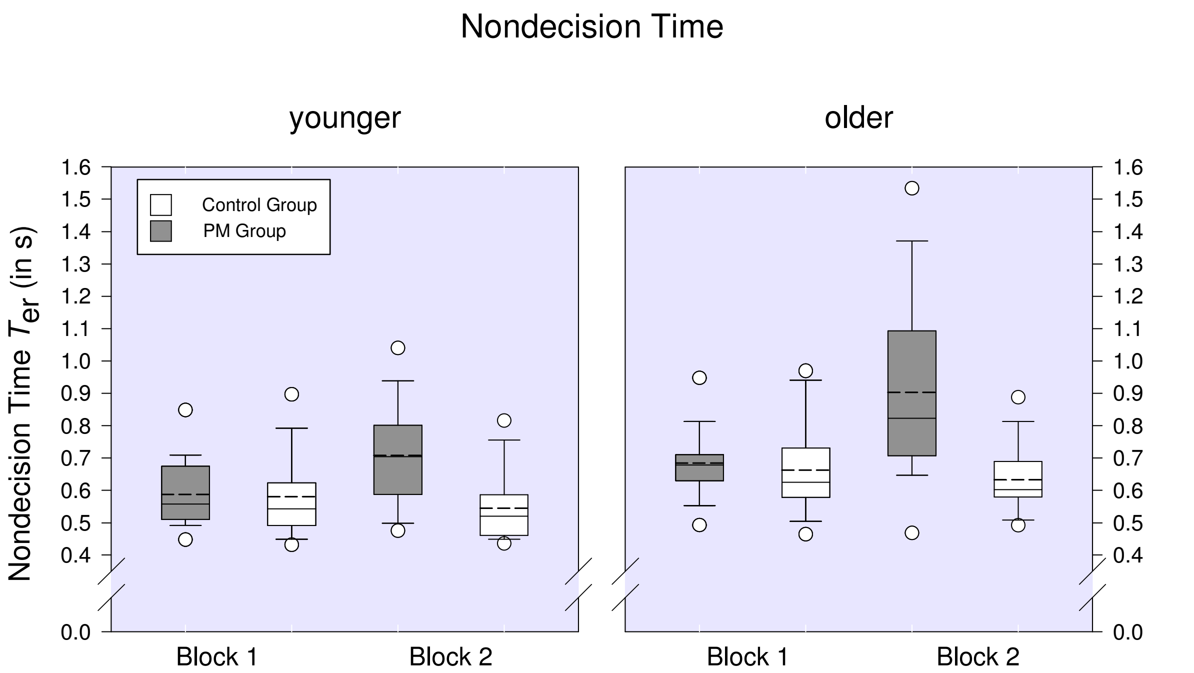


Figure 2. Best-fitting diffusion-model parameters. Drift rate v, boundary separation a, and nondecision time T_{er} as a function of *Age* (young vs. older), *Experimental Group* (PM, No-PM Control) and *Block* (first vs. second). The box-and-whisker plots include the medians (solid lines) and means (dashed lines). Whiskers (error bars) below/above the boxes mark the 10th and 90th percentiles, and dots mark the 5th and 95th percentiles. Scaling parameter (diffusion constant) s=1.




Boundary Separation
younger
older
Boundary separation a
Control Group
PM Group
Block 1
Block 2
Block 1
Block 2


Nondecision Time
younger
older
Nondecision Time Ter (in s)
Control Group
PM Group
Block 1
Block 2
Block 1
Block 2

Supplement 1. Best fitting variability parameters of the diffusion model.

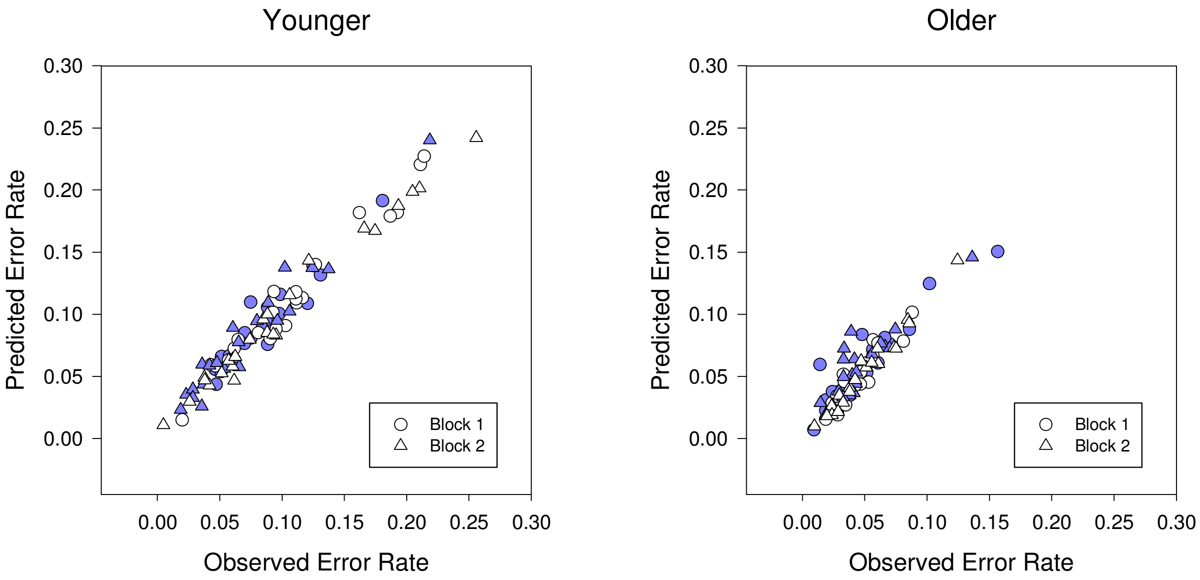
Measure	Group	Block 1		Block 2	
		M	(*SE*)	*M*	(*SE*)
η	Younger				
	PM	.42	(.05)	.46	(.04)
	No-PM	.36	(.04)	.34	(.04)
	Older				
	PM	.46	(.05)	.38	(.04)
	No-PM	.41	(.05)	.39	(.04)
s_t (in s)	Younger				
	PM	.04	(.02)	.04	(.01)
	No-PM	.05	(.01)	.02	(.01)
	Older				
	PM	.03	(.02)	.03	(.01)
	No-PM	.04	(.02)	.02	(.01)
s_z	Younger				
	PM	.58	(.06)	.65	(.06)
	No-PM	.51	(.06)	.47	(.06)
	Older				
	PM	.71	(.06)	.92	(.06)
	No-PM	.72	(.06)	.70	(.06)

Notes. Means of across-trials variability in drift rate (η), starting point (s_z), and nondecision time (s_t). Diffusion constant s = 1. Standard errors are in parentheses; PM = Prospective Memory.

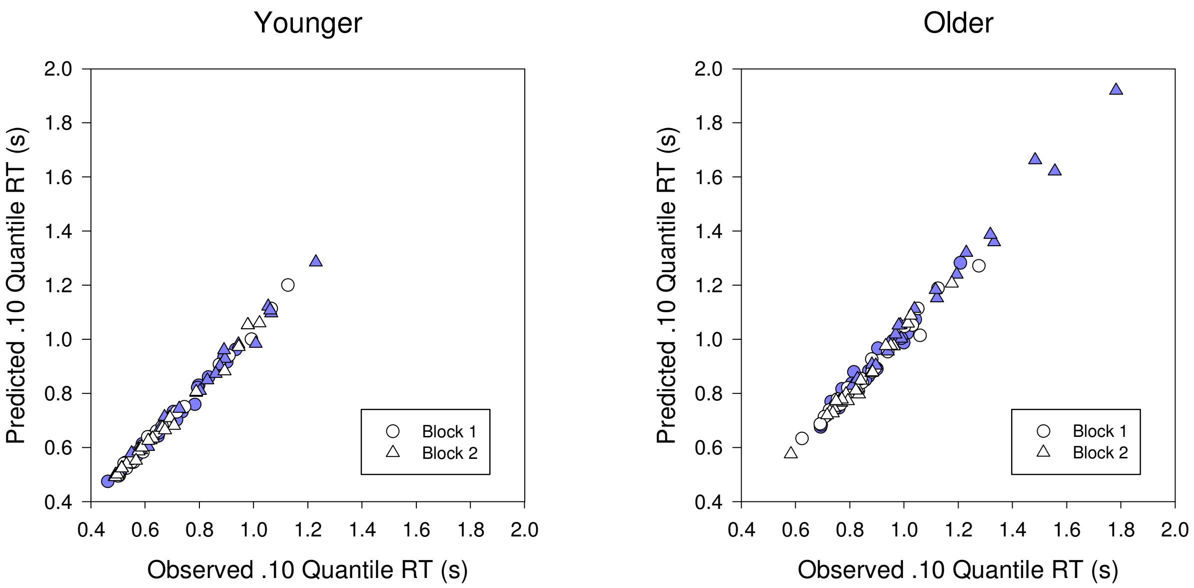
Supplement 2. Individual fits of the diffusion model. The present diffusion-model analysis rests on the Kolmogorov–Smirnov statistic (KS; Kolmogorov, 1941) as objective function for parameter estimation, which represents the maximum vertical distance between the predicted and the observed cumulative distribution functions (CDFs) of RT. The parameter values were determined in such a way that this objective statistic was minimized. In our analysis, none of the corresponding *p*-values indicated significant deviations from the empirical data in Block 1 (average *p*s: M_{older} =.998, SD_{older} =.007; M_{younger} =.988, SD_{younger} =.022) or in Block 2 (M_{older} =.992, SD_{older} =.021; M_{younger} =.987 SD_{younger} =.022).

In addition to this quantitative assessment, we plotted the models' predicted against observed (empirical) individual values to determine fit across the sample qualitatively. In the graphs below, diagonal lines with a slope of +1 would indicate perfect fit. The upper panels (*2a*) show observed and predicted mean error rates for each participant in the PM group (blue) and in the No-PM Control group (white) as function of *Age Group* and *Block*. Panels *2 b-f* show observed and predicted quantiles of correct responses (the .1, .3, .5, .7, and .9 quantile of the RT distribution). As shown, correspondence is relatively high and the residuals do not indicate systematic biases in model predictions. Overall, the diffusion model adequately reproduces the effects on individual RTs and accuracy in the PM paradigm for both younger and older adults.

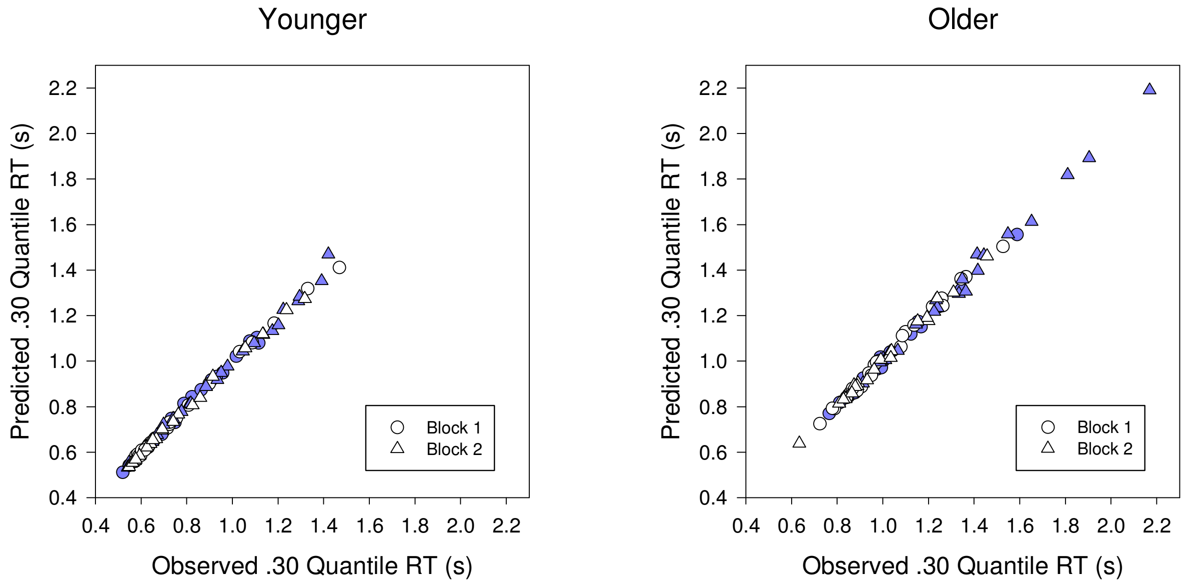
(*2a*) Error Rate.



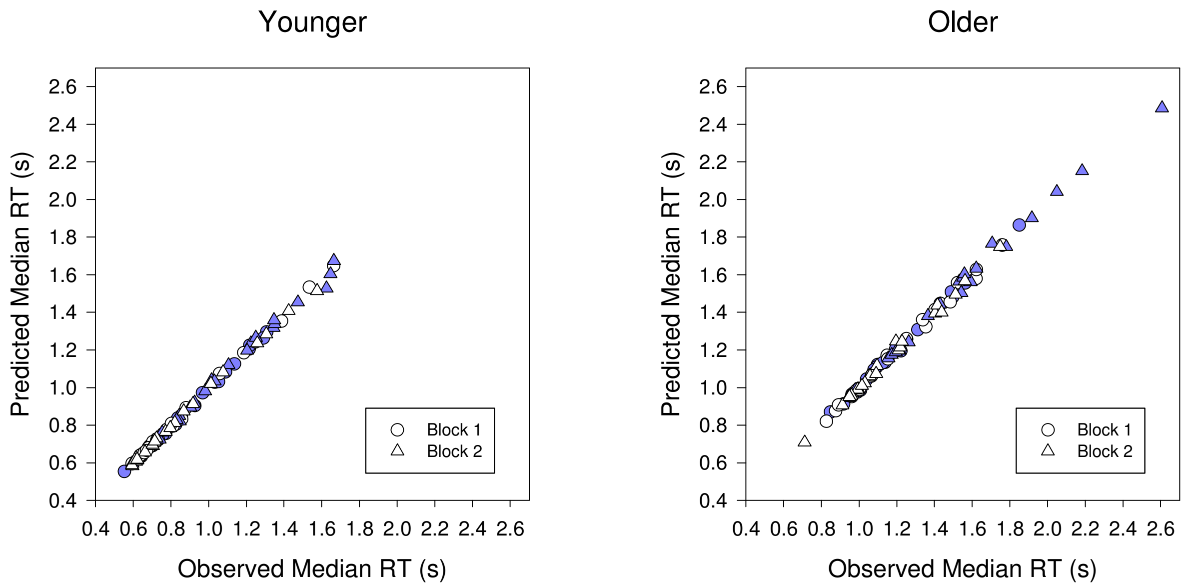
(*2b*) .1 RT Quantile.



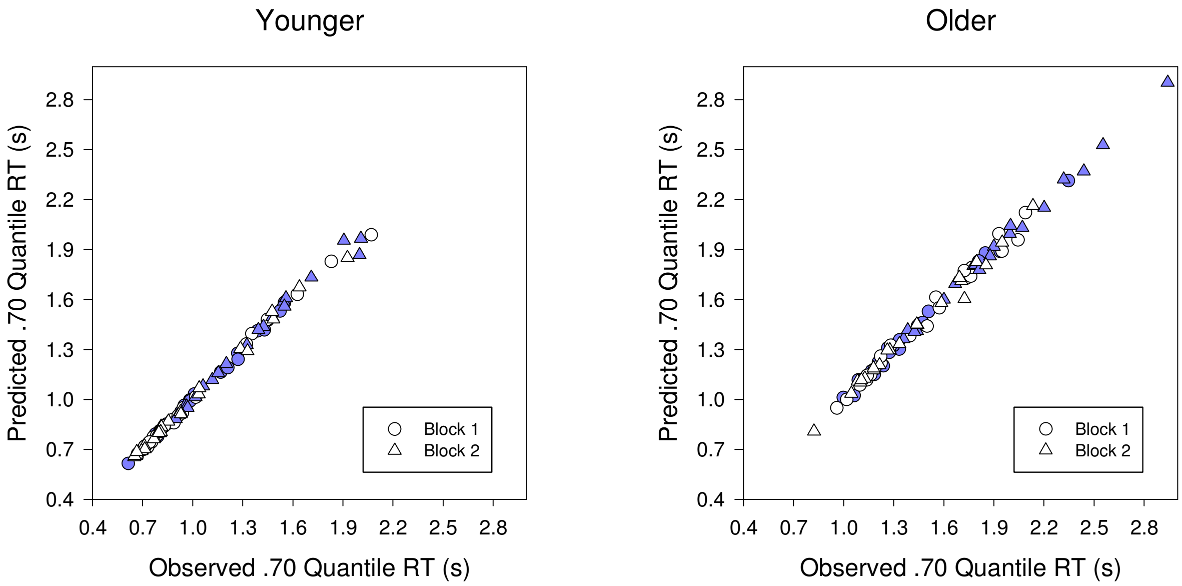
(*2c*) .3 RT Quantile



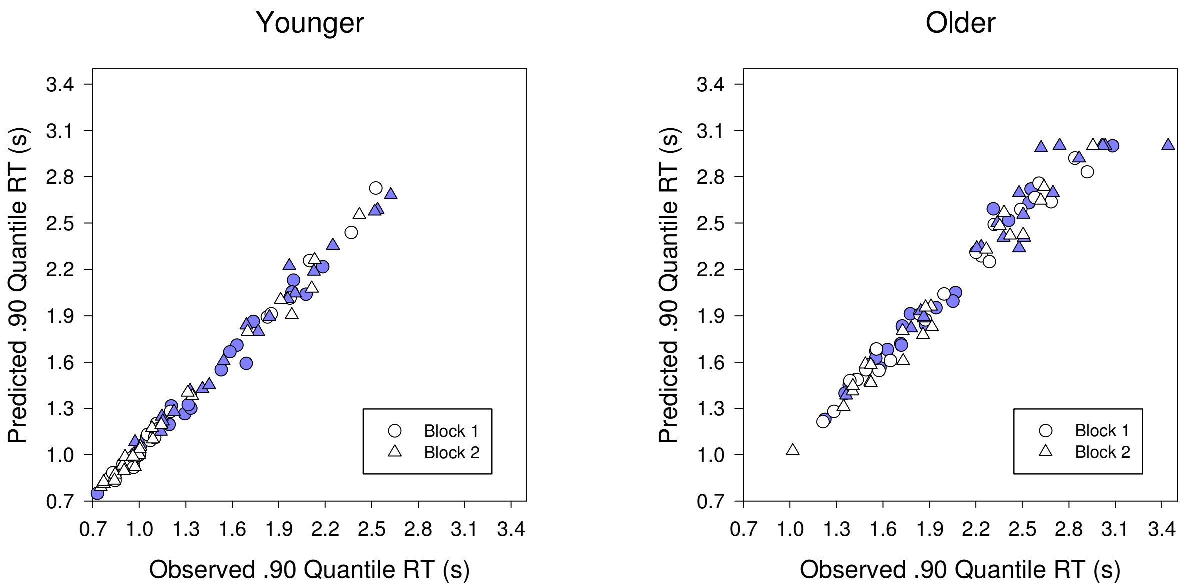
(*2d*) .5 RT Quantile (Median)



(*2e*) .7 RT Quantile



(*2f*) .9 RT Quantile



Supplement 3. Starting point $z = a/2$. The relation of the starting point z to the boundaries is an indicator of decision bias. If a decision maker does not favor one response over the other, the starting point position is equidistant from both decision boundaries (z = a/2), which holds for many experiments (e.g., Wagenmakers, van der Maas, & Grasman, 2007). In the present analysis, we coded responses as *correct* (upper boundary) vs. *incorrect* (lower boundary), yielding an average measure of drift rate across words and nonwords. Note that with this coding, any bias towards a decision alternative would lead participants to favor correct and error responses equally often if stimulus types are presented in equal proportions, as in the present study (50% words, 50% nonwords). If present, any bias would thus not affect the mean starting point position (i.e., z can be fixed at $a/2$), but may contribute to starting-point variability s_z (see Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007, for the same approach).

Decision-bias checks. Although not a necessary precondition for the present modeling, we additionally compared the relative speed of correct vs. error responses as function of stimulus type. This allows to examine whether biases towards a decision alternative were present (Wagenmakers et al., 2007). An a-priori bias for a particular decision (e.g.,“word”) would imply faster correct than error responses for the corresponding stimulus type (words), and the reverse pattern of RT for the other type (nonwords). ANOVAs on RTs indicated that *String Type* (word vs. nonword) did not interact with *Correctness* (correct response vs. error) for younger adults in any Group, or for older adults in the PM Group (all Fs< 1), but for older adults in the No-PM Control Group [$F(1,21)=7.29$, $p< .05$]. However, further visual examination showed that the latter interaction was not of a cross-over type, which would be expected with a systematic bias (see Wagenmakers et al., 2007), and was not reliable after Bonferroni-correction for multiple testing. In sum, the present data do not indicate systematic biases towards a particular decision.

Supplement 4. ANOVA tables for between-subjects factors *Age* (younger, older adults) and *Experimental Group* (PM, NoPM), and within-subjects factor *Block* (first, second), on ongoing-task response time (RT), accuracy, drift rate *v*, boundary separation *a*, and nondecision time T_{er}.

Variable	Factor	η_p^2	F (1,85)	p
RT	*Group*	.08	7.73	< .01
	Age	.33	41.89	< .001
	Age × Group	.00	.00	.96
	Block	.36	47.07	< .001
	Block × Group	.49	82.27	< .001
	Block × Age	.04	3.39	.07
	Block × Age × Group	.07	6.17	.02
Accuracy	*Group*	.02	1.42	.24
	Age	.24	26.55	< .001
	Age × Group	.03	2.64	.11
	Block	.04	3.32	.07
	Block × Group	.04	3.73	.06
	Block × Age	.05	4.08	< .05
	Block × Age × Group	.00	.01	.94
Drift rate *v*	*Group*	.02	2.03	.16
	Age	.05	4.19	.04
	Age × Group	.00	0.12	.73
	Block	.06	5.00	.03
	Block × Group	.09	8.77	< .01
	Block × Age	.03	2.46	.12
	Block × Age × Group	.01	0.44	.51
Boundary separation *a*	*Group*	.03	2.67	.11
	Age	.40	56.91	< .001
	Age × Group	.01	0.90	.34
	Block	.15	14.92	< .001
	Block × Group	.20	21.76	< .001
	Block × Age	.01	0.73	.40
	Block × Age × Group	.01	1.04	.31
Nondecision time T_{er}	*Group*	.16	15.87	< .001
	Age	.16	15.79	< .001
	Age × Group	.01	1.09	.30
	Block	.25	27.70	< .001
	Block × Group	.42	60.92	< .001
	Block × Age	.04	3.93	.051
	Block × Age × Group	.04	3.19	.08

References

Klauer, K.C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93*, 353–368.

Kolmogorov, A. (1941). Confidence limits for an unknown distribution function. *Annals of Mathematical Statistics, 12*, 461–463.

Wagenmakers, E.J., van der Maas, H.L.J., & Grasman, R. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14,* 3–22.

Neuropsychologia 50 (2012) 67–76

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Does frequency matter? ERP and behavioral correlates of monitoring for rare and frequent prospective memory targets

Daniela Czernochowski*, Sebastian Horn, Ute J. Bayen

Institute for Experimental Psychology, Heinrich-Heine-University, Düsseldorf, Germany

ARTICLE INFO

Article history:
Received 10 February 2011
Received in revised form 26 October 2011
Accepted 28 October 2011
Available online 3 November 2011

Keywords:
Prospective memory
ERPs
Executive functions
Prefrontal cortex

ABSTRACT

Behavioral and event-related potential (ERP) correlates of monitoring in an event-based prospective memory (PM) task were compared during blocks with rare versus frequent PM target presentations relative to an ongoing-task only condition. For both rare and frequent PM conditions, behavioral interference costs in terms of longer reaction times (RTs) were observed. Likewise, during both PM blocks a sustained ERP positivity with a frontal focus was identified on ongoing-task trials. While PM target identification and RT interference costs were larger during the PM-frequent relative to the PM-rare condition, the same sustained frontal positivity was observed during both PM blocks. These findings suggest that successful monitoring is associated with the adoption of a more general prospective retrieval mode, irrespective of target frequency. Moreover, preparatory attentional modulations directed at relevant target features played an important role for subsequent PM performance, as evident in larger P2 amplitudes during PM blocks.

1. Introduction

Prospective memory (PM) refers to the delayed realization of intentions, for instance remembering to pass a message to your colleagues the next time you see them. Frequently, such intentions will not be realized during the course of a busy day, sometimes with serious consequences. PM has found increasing interest among memory researchers during the last two decades, both in terms of the cognitive processes involved and its practical implications for everyday functioning (e.g., for maintaining health and independence in old age; see Kliegel, McDaniel, & Einstein, 2008, for an overview). Theorists agree that generally, monitoring processes play an important role in successfully maintaining delayed intentions and in initiating the intended action in the appropriate situation. However, it remains unclear precisely how these monitoring processes support PM performance, and whether resource allocation towards PM monitoring – associated with costs on ongoing activities – is always necessary.

According to the preparatory attentional and memory processes (PAM) theory of PM (e.g., Smith, 2003; Smith & Bayen, 2004, 2005), attentional resources dedicated to monitoring are necessary for successful PM performance regardless of task characteristics. By contrast, the multi-process framework (Einstein & McDaniel, 2005; McDaniel & Einstein, 2000) postulates that delayed intentions can be retrieved spontaneously without active monitoring for PM target occurrence under certain task conditions. This may be the case, for instance, when PM targets are highly associated with the appropriate target action. Both theories concur that monitoring processes are necessary to successfully carry out delayed PM intentions under many task conditions, for instance when several PM targets are used (Einstein & McDaniel, 2005; McDaniel & Einstein, 2000; Smith, Hunt, McVay, & McConnell, 2007). The retrieval mode and target checking theory (Guynn, 2003, 2008) attempts to specify monitoring processes more directly. According to this framework, a sustained prospective retrieval mode enables individuals to treat stimuli as cues for retrieving PM intentions, and hence is necessary for PM. This strategic process is sometimes complemented by periodic item checking for relevant features defining PM targets.

In the laboratory, monitoring processes are usually inferred on the basis of behavioral costs to an ongoing task. This so-called prospective interference effect (Marsh, Hicks, Cook, Hansen, & Pallos, 2003; Smith, 2003) is typically observed in terms of longer ongoing-task reaction times (RTs) in the presence (versus absence) of a PM task. Participants first encode PM target stimuli and a specified response to be initiated whenever a PM target is presented. PM target events are then embedded within the ongoing task. Typically, several target events occur (either recurring or different stimuli) to allow for a more reliable measurement of PM performance. Consistent with PM requirements outside the laboratory, in most behavioral studies PM targets occur rarely (i.e., in 2–10% of all ongoing-task trials). Importantly, the character of the PM task may vary considerably when PM targets occur more often

* Corresponding author at: Institute for Experimental Psychology, Heinrich-Heine University Düsseldorf, R 23.03.U1.82, 40204 Düsseldorf, Germany.
Tel.: +49 211 8113863; fax: +49 211 8115037.
E-mail address: d.czernochowski@hhu.de (D. Czernochowski).

0028-3932/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2011.10.023

(see also Monsell, 2003 for related findings in explicit task-switching paradigms). Graf and Uttl (2001) suggested that so-called "vigilance tasks" and "PM proper tasks" form two ends of a continuum. According to this framework, at one end of this continuum (termed "vigilance"), processing resources are allocated to PM-target monitoring completely, and hence, behavioral costs to ongoing activities may be substantial. On the other side of this continuum (termed "PM proper"), PM intentions stay less active in working memory, as many ongoing trials intervene. Consequently, tasks with frequent and rare target occurrences would differ in the degree of monitoring necessary to support successful PM performance, as intentions are more likely kept active in working memory with frequent compared to rare PM targets. Recent evidence suggests that the amount of attention devoted to the PM task is subject to adjustments based on task experience. That is, behavioral costs are reduced when PM targets are not presented despite PM instructions (Loft, Kearney, & Remington, 2008). Following this logic, even more attentional resources might be invested to monitor for subsequent PM targets when they occur frequently, and hence behavioral costs might increase substantially. Alternatively, it might be less demanding to maintain a PM intention when it must be carried out more often, because successfully detected PM targets may serve as reminders of the PM intention. For instance, passing a message to one colleague may remind you of your intention to pass this message to someone else as well. Following this idea, less PM monitoring would be necessary for successful PM performance if targets occur frequently. So far, the precise nature of PM monitoring processes and their relation to variations in target frequency remains open.

To further characterize the cognitive processes underlying performance in PM tasks, neural activity has been measured along with behavioral performance. To this end, predominantly two methods have been applied in previous studies, namely functional magnetic resonance imaging (fMRI) and event-related potentials (ERPs). A recent fMRI study (Reynolds, West, & Braver, 2009) compared transient and sustained neural activity during a PM task. Transient activity during ongoing trials was taken to reflect periodic item checking (see also Guynn, 2003), but was not observed during ongoing trials. By contrast, sustained activity during ongoing trials was taken to reflect controlled processes associated with strategic monitoring, and was observed in a network of brain areas including the anterior prefrontal cortex (PFC; specifically lateral Brodmann area 10). Notably, this region has been implied previously in so-called *stimulus-independent* cognitive processing directed away from external stimuli and towards internal mental representations (Burgess, Dumontheil, & Gilbert, 2007; Burgess, Gilbert, & Dumontheil, 2007). Moreover, sustained activity in the anterior PFC was correlated with faster RTs for PM targets, thus providing further evidence for a functional role of sustained activity in the anterior PFC for PM performance (Reynolds et al., 2009).

So far, several investigations have used ERPs to assess the neural activity during PM tasks (e.g., West & Bowry, 2005; West & Krompinger, 2005; West, McNerney, & Krauss, 2007; West, McNerney, & Travers, 2007; West, Scolaro, & Bailey, 2011; Zöllig et al., 2007; see also a recent review by West, 2011). For instance, West, Bowry, and Krompinger (2006) first mentioned a sustained positivity at frontal electrodes on ongoing trials in PM-blocks compared to blocks without PM instructions. Consistent with the well-established role of the frontal cortex for executive control functions, all investigations reported ERP modulations at frontal electrodes.[1] Three recent studies focused on the prospective interference effect in particular (i.e., the comparison of ongoing-task trials with and without PM instructions; Chen, Huang, Jackson, & Yang, 2009; Chen, Huang, Yang, Ren, & Yue, 2007; Knight, Ethridge, Marsh, & Clementz, 2010; see also West et al., 2006). Ongoing trials in PM-blocks were associated with larger ERP amplitudes compared to ongoing trials without PM instructions in several short time windows between 200 ms and the end of the recording epoch (400 or 600 ms). Notably, later time windows in which sustained processes related to maintaining a delayed intention over time should be particularly pronounced, were not evaluated in these investigations.

ERP modulations with latencies around 200 ms at frontal electrodes have been reported previously in the context of other experimental manipulations. For instance, larger P2 amplitudes have been reported for trials in which deviant item features are task relevant (e.g., Luck & Hillyard, 1994; Potts, 2004; Ruz & Nobre, 2008) or under conditions of increased arousal (e.g., following caffeine intake; Ruijter, Lorist, Snel, & De Ruiter, 2000). However, these modulations were restricted to the P2 time range, and were not sustained over several hundred milliseconds. Given this transient time course, it seems unlikely that they reflect a sustained retrieval mode and/or continuous preparatory attentional processes. As the P2 modulations described above occur particularly under conditions of interference from competing stimuli or task demands, they have been taken as evidence for a focus towards specific relevant item features as a result of top-down control of attention (see also Luck, Woodman, & Vogel, 2000). In a recent PM study, Knight et al. (2010) examined ongoing trials with and without a PM intention. In this investigation, PM targets were defined by the conjunction of two features (words printed in red) embedded in a lexical decision task. Modulations of ERP components around 200 ms post-stimulus onset were taken as evidence for preparatory attention aimed at identifying relevant stimulus features defining PM targets (see also Guynn, 2003, 2008). The specific timing and location of these effects may have been affected by this particular choice of relevant features and by their conjunction. Notably, the ERP epochs only spanned 600 ms, consistent with previous ERP studies (Chen et al., 2007, 2009). Hence, ERP analyses between ongoing-task trials with and without PM intentions have been restricted to stimulus evaluation during early time windows; sustained ERP differences more likely associated with maintaining a controlled retrieval mode across trials have not been evaluated so far. It thus remains unclear how these early ERP modulations relate to the *sustained* fMRI activity in anterior PFC reported by Reynolds et al. (2009), which has been directly associated with actual PM performance.

One central aspect that may influence how participants monitor for PM targets and that differs between the studies described so far is the frequency of PM target events. In investigations examining the neural correlates of controlled processes during PM tasks, typically relatively frequent PM targets have been used. For instance, in two ERP studies examining correlates of prospective monitoring (Chen et al., 2007, 2009), PM targets occurred in 20% of all trials, and in the fMRI study detailed above (Reynolds et al., 2009) PM targets occurred in 11% of all trials. By contrast, behavioral studies typically rely on fewer target events (Einstein & McDaniel, 2005). For instance, two recent behavioral studies demonstrate that PM interference effects are smaller when a particularly large number of intervening ongoing trials is presented between PM targets (fixed number of 89 versus 32 between each PM target presentation,

[1] Note that some ERP investigations reported a sustained frontal positivity, coupled with an additional occipital-parietal slow-wave negativity (see West et al., 2006, and West, McNerney, & Travers, 2007, for details on more posterior ERP correlates of monitoring). Posterior activity has been predominantly observed with an average reference. By contrast, a mastoid reference was used in the present analysis, as well as in the investigations by Chen et al. (2007, 2009). With a mastoid reference, the frontal positivity was not associated with a negative deflection at opposite sites of the scalp (see Luck, 2005).

corresponding to 3 versus 1% PM target frequency; Loft & Yeo, 2007, Experiment 3), or when no PM targets are presented following PM instructions (Loft et al., 2008). If PM target frequency influences the degree or the type of monitoring processes recruited to support PM performance, the neural correlates of PM monitoring recently proposed might be characteristic for conditions with frequent PM targets only. This would have important implications, because the conclusions drawn from neuro-scientific approaches to PM may strongly differ from behavioral paradigms (or everyday situations) with typically fewer target events. Given the potential impact of PM target frequency on the nature of monitoring, it is surprising that so far monitoring for frequent versus rare PM target occurrences has not been directly compared. We addressed this open question by explicitly examining whether ERP and behavioral correlates of PM monitoring are affected by frequent versus rare PM target occurrences. PM target stimuli were selected according to two criteria. First, they were not particularly salient. Second, attending to the feature that identified stimuli as PM targets was not critical for ongoing-task performance (i.e., non-focal PM target stimuli). Under these conditions, both the multi-process framework and PAM theory concur that monitoring processes are necessary for successful PM performance (McDaniel & Einstein, 2000; Smith et al., 2007). Hence, we expected behavioral costs in the ongoing task for both PM conditions relative to the control condition without PM instructions. As each successfully identified PM target could serve as a reminder of the PM task, we expected better PM performance for the frequent compared to rare PM condition.

ERPs were employed to further characterize monitoring processes. To allow the examination of slow-wave brain activity, the present ERP epochs spanned 1000 milliseconds following stimulus presentation. We expected neural correlates of PM monitoring to be evident for both frequent and rare PM conditions relative to the control condition. In line with the proposed role of the frontal cortex for maintaining and carrying out delayed intentions, we expected sustained ERP modulations predominantly over frontal electrode sites (see also footnote 1). Based on the recent fMRI findings described above and the assumption of a functional relationship between monitoring and successful PM performance, we expected an association between PM performance and the ERP correlates of monitoring. If adopting a prospective retrieval mode is a prerequisite for identifying PM targets (as suggested by Guynn, 2003, 2008), the ERP correlate for PM monitoring should be observable whenever participants successfully identify PM targets, *irrespective* of PM target frequency. By contrast, differences in the ERP correlates of PM monitoring for frequent and rare PM conditions would suggest differences in the degree of monitoring processes supporting PM performance, as suggested by Graf and Uttl (2001). Consistent with the process of target checking (Guynn, 2008) and recent ERP findings (e.g., Knight et al., 2010), we expected a P2 modulation reflecting selective attention to relevant features defining PM targets.

2. Methods

2.1. Participants

Nineteen students completed the experiment. We had to exclude three participants from the analyses. Of these, two never pressed the PM key in one or both of the PM blocks, and one misunderstood the PM instructions and responded to a different stimulus. The final sample thus consisted of 16 participants (3 males; mean age 23.4 years, range 19–29 years). All participants were native speakers of German, right-handed and reported to have normal or corrected-to-normal vision. Participants reported themselves to be in good physical and mental health. All participants gave informed consent and received course credit or a monetary incentive for participation.

2.2. Materials

The stimuli were strings of 8–12 letters, which were either German words or pronounceable non-words. The words were low-frequency German words from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995), occurring either 4–5 or 0–1 times per million. We created the non-words by randomly replacing all vowels/umlauts in the words with randomly chosen different vowels/umlauts.

For the experimental blocks, we created two target sets and two filler sets that were matched for word frequency and string length, respectively. Each target set contained a total of 116 target strings (starting with the letters *G*, *H*, or *M*) and each filler set contained 1384 non-target strings (strings with all remaining starting letters). During the experiment, the program randomly chose one filler set and one target set, and the stimuli were randomly drawn from these sets without replacement. The four possible target-list × filler-list combinations were counterbalanced across participants. To avoid repetition priming, a given stimulus never occurred both as a word and as the corresponding non-word within a single session. For the practice phase, we chose 24 words and created 24 corresponding non-words in a similar way.

2.3. Procedure

The three conditions of the experiment were presented in blocks of 500 trials each and were separated by a pause of 3 min in which participants filled out a questionnaire. Each block was further divided into four equal parts of 125 trials separated by short breaks of 1 min to avoid fatigue.

Prior to the actual experiment, 48 ongoing lexical-decision-task trials were included as practice. During practice, a visual feedback "too slow" was given if participants did not respond within 1500 ms, and the feedback "wrong" whenever an error occurred in the lexical decision task. Feedback was given during practice to ensure that participants performed as fast and as accurately as possible. At the end of practice, participants received feedback regarding the proportion of correct responses and their average RT in the entire practice block.

Frequency of PM target occurrence was manipulated within participants. All participants first completed a control block of ongoing-task trials alone, then two blocks with the embedded PM task. The concern was to avoid carry-over effects from the PM conditions to the control condition (i.e., continued monitoring for PM targets). After completion of the control block, participants received one PM block with rare target occurrence, and one PM block with frequent target occurrence. The order of these two blocks was counterbalanced across participants. During the ongoing lexical-decision task, in each trial, a fixation cross appeared with a mean onset latency of 1250 ms, randomly jittered between 1150 and 1350 ms, to prevent anticipatory responding. Each item was presented at the center of a white screen in black upper case letters (font size 24). Participants were asked to categorize each item as a word or a non-word, with equal emphasis on speed and accuracy. Participants gave their response with the left or right index finger via button press on a response box. The assignment of response keys (left versus right) to the *word* and *non-word* response options was counterbalanced across participants. Each item was displayed until the participant gave a response. Following the response, a blank screen was presented for 500 ms. Participants received no performance feedback during the actual experiment.

After completion of the ongoing-task control block, participants received instructions for the first PM block. One half of the participants were asked to press the far right key, the other half to press the far left key on the response box whenever a letter string started with the letters *G*, *H*, or *M*. The target letters were presented before each PM block, one at a time for 5 s each. To ensure that all participants had encoded the three letters, they were asked to count backwards from 100 in steps of 3s for 30 s, and to then reproduce the target letters. If participants were not successful, the presentation of the target letters was repeated.

During the PM blocks, PM targets were embedded into the ongoing task. The exact distance between two targets was randomized within a certain trial-range to discourage a strategy of counting trials, particularly for the frequent PM target condition. In the rare-target condition, PM targets appeared on 3% of the trials. There were thus 16 targets, 4 in each block of 125 trials, with 22 to 27 ongoing trials between any two PM targets. In the frequent-target condition, targets appeared on 20% of the trials, with a minimum of two and a maximum of six ongoing trials between any two PM target trials. In this condition, there were thus 100 targets, 25 in each block of 125 trials. The total experiment lasted about 90 min. Participants were debriefed after completion.

2.4. EEG recording and ERP data preprocessing

EEG activity was recorded from 40 scalp sites placed according to the extended 10–20 system with sintered Ag/AgCl electrodes using a ground on the right forehead. Vertical and horizontal electrooculogram (EOG) was recorded from electrodes placed above and below the left eye and at the outer canthus of each eye. Electrode impedance was kept below 5000 Ω. The activity of all scalp electrodes was initially referenced to the right mastoid and re-referenced offline to averaged mastoids. EEG and EOG were recorded continuously with a NyAmp Express amplifier (DC; 100 Hz high-frequency cutoff; 1000 Hz digitization rate). EEG data were further analyzed using Vision Analyzer 2 software (Brain Products, Gilching, Germany). The continuous EEG was downsampled to 500 Hz and a .1–30 Hz band-pass filter was applied offline. If single channels showed artifacts, a spherical spline algorithm (Perrin, Pernier, Bertrand, & Echallier, 1989) was used for interpolation, with a maximum of two channels interpolated for a given participant. Trials containing voltage steps > 100 μV (e.g., muscular activity) were removed prior to the removal of eye movements. Eye movements were corrected via independent component analysis (Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997), using the ocular correction ICA tool in Vision Analyzer 2, based on an artifact-free segment with a length of 180 s. Finally, trials with remaining artifacts (i.e., voltage steps > 50 μV, voltage differences of >100 μV within a 200 ms time window) were removed.

EEG epochs extended from 200 ms prior to stimulus onset until 1000 ms, for a total duration of 1200 ms. Averages were constructed for each participant for a total of five conditions[2] (mean trial numbers and range are given in parentheses): ongoing trials in the control block (ongoing only: 415.13, 240–488), ongoing trials during the PM-rare block (PM-rare: 336.13, 187–452), ongoing trials during the PM-frequent block (PM-frequent 199.75, 118–271). Ongoing-task trials occurring within 30 s before a PM-rare target hit (before hit: 64.53, 22–108) and before a PM-rare target miss (before miss: 69.93, 25–107) were analyzed in 15 participants who contributed a minimum of 20 artifact-free trials for these analyses. Due to PM target frequency, this analysis was not feasible for the frequent PM condition.

We chose an α-level of .05 for all analyses. Greenhouse-Geisser corrections for the violation of sphericity were used if necessary and are reflected in the *p*-values along with the respective epsilon values (ε) and uncorrected degrees of freedom. Partial eta squared (η_p^2) is reported as an estimate of effect size for main and interaction effects. Following the statistical analyses, a lowpass filter of 15 Hz was applied to the EEG waveforms and is reflected only in the figures.

3. Results

3.1. Behavioral performance

Behavioral analyses focused on two aspects. First, we compared hit rates and RTs on PM targets between the PM-rare and PM-frequent conditions. Second, we assessed ongoing-task RTs and accuracy (i.e., the proportion of correct lexical decisions) in separate repeated-measures ANOVAs including the factor Condition (ongoing only, PM-rare, PM-frequent). The two trials following a PM target were excluded from analyses of ongoing-task performance to avoid inclusion of residual response processes prompted by PM targets. Consistent with the cutoff criteria used for ERP analyses, outliers (RTs faster than 200 ms or slower than 3000 ms) were removed prior to statistical analyses. Moreover, error trials were discarded for the RT analyses (see also Marsh et al., 2003).

3.1.1. PM performance

Consistent with expectations, participants were very accurate in detecting PM targets in the frequent-PM condition (PM target hit rate: *M* = 89.3%, SD = 9.4; range 69–100 out of 100 PM targets). In the PM-rare condition, PM target hit rate was considerably lower (*M* = 55.5%; SD = 18.9; range 5–15 out of 16 PM targets), $F(1,15) = 52.17$, $p < .001$, $\eta_p^2 = .78$. RTs on PM target hits were faster during the PM-frequent (*M* = 822 ms, SD = 121 ms) compared to the PM-rare condition (*M* = 1027 ms, SD = 313 ms), $F(1,15) = 9.78$, $p < .01$, $\eta_p^2 = .40$.

3.1.2. Interference costs on ongoing-task performance

Accuracy in the ongoing lexical decision task was consistently high across the three conditions (ongoing-only condition: *M* = 95%, SD = 3.8%; PM-rare condition: *M* = 96%, SD = 2.6%; PM-frequent condition: *M* = 95%, SD = 2.4%), $F(2,30) = 1.18$, $p = .32$, $\eta_p^2 = .07$). Consistent with prior investigations (e.g., Marsh et al., 2003), no interference was evident in terms of reduced accuracy in the ongoing task.

By contrast, costs were observed in terms of longer RTs, as suggested by a main effect of Condition, $F(2, 30) = 49.39$, $p < .0001$, $\eta_p^2 = .77$. Paired contrasts indicated that RTs were longer for trials in the PM-rare condition (*M* = 795 ms, SD = 129 ms) relative to the ongoing-only condition (*M* = 665 ms, SD = 93 ms), $p < .01$, as well as for the PM-frequent condition (*M* = 847 ms, SD = 130 ms) relative to the PM-rare condition ($p < .01$). An interaction of behavioral RT costs and the order in which participants completed the PM rare and frequent conditions was observed, $F(1,14) = 6.62$, $p < .05$, $\eta_p^2 = .32$, indicating that responses during the PM-rare block were slightly slower for participants who completed the PM-frequent block first. Importantly, reliable RT interference costs in the same direction (i.e., larger for the PM-frequent condition) and of similar effect sizes were still observed in both counterbalanced conditions, $\eta_p^2 = .79$ (PM rare first) and .85 (PM frequent first), *F*s (2,14) > 25.63, *p*s < .0001.

3.2. ERP results

3.2.1. ERP correlates of prospective monitoring

As illustrated in Fig. 1, both PM conditions were associated with very similar waveforms, namely a sustained positivity relative to trials in the control block. To compare ERP correlates of prospective monitoring for the PM-rare and PM-frequent condition, mean amplitudes between 600 and 900 ms following stimulus cue onset were evaluated in a mixed-model ANOVA with the factors Condition (ongoing-only, PM-rare, PM-frequent) × Anterior-Posterior Electrode site (AP) × Left-Central-Right Electrode site (LCR). For this analysis, we selected 9 electrode sites covering lateral and midline parts of the scalp (F7, Fz, F8, T7, Cz, T8, P7, Pz, P8). To further characterize the topographical distribution of ERP effects across these electrode sites, reliable interactions with the factor Condition were followed up with subsidiary ANOVAs where appropriate. Following our central hypothesis, two planned contrasts were specified for the factor Condition, namely PM-rare versus ongoing-only, and PM-frequent versus ongoing-only.

The overall ANOVA with the factors Condition × AP × LCR revealed a main effect of condition, $F(2,30) = 5.25$, $p < .05$, $\eta_p^2 = .26$, which was modulated by two-way interactions with the factors AP and LCR and a three-way interaction (see Table 1). Subsidiary ANOVAs indicated reliable main effects of condition at frontal and central electrode sites, and interactions of Condition × LCR at frontal, central and parietal electrode sites (all *p*s < .05). Planned contrasts for each electrode revealed reliable ERP differences relative to ongoing-task only at F7, Fz, T7, Cz and Pz for the PM-rare condition, and at F7, Fz, T7 and Cz for the PM-frequent condition (all *p*s < .05). Consistent with the proposed functional relevance for PM monitoring, for both contrasts, effect sizes were largest at frontal electrodes F7 and Fz (see Table 1), as also evident in Fig. 2. The order in which participants completed the PM rare and frequent conditions did not influence the magnitude of the ERP correlate for PM monitoring [interaction with block order: $F(2,28) < 1$, $p = .49$].

We then examined a potential relationship between these proposed neural correlates of PM monitoring and PM target performance measures. Assuming that larger ERP amplitude differences should be associated with larger performance benefits, we correlated the magnitude of the ERP differences between PM and control conditions (i.e., the mean of the five electrodes with reliable effects) with mean RTs of PM target hits (Reynolds et al., 2009). Indeed, larger ERP correlates of PM monitoring were associated with faster responses to PM targets for the PM-frequent ($r = -.53$) and for the PM-rare condition ($r = -.44$), *p*s < .05.[3] By contrast, larger

[2] In the present investigation, PM targets could be both words and non-words, hence the classification as a word or non-word was irrelevant for the PM task. Importantly, no reliable main effects or interactions with Stimulus Type (word versus non-word) were found in the time windows of the present analyses: all *F*s < 2.46, all *p*s > .14 (for 600–900 ms), all *F*s < 1.49, all *p*s > .21 (for the P2 analysis, 160–210 ms). Therefore, word and non-word stimuli were collapsed for all subsequent analyses.

[3] Controlling for individual differences in general processing speed by considering baseline RT in the ongoing task as a covariate did not change this pattern of results (partial *r*s = −.43 and −.53, for the PM-rare and PM-frequent conditions, respectively; all *p*s < .05).

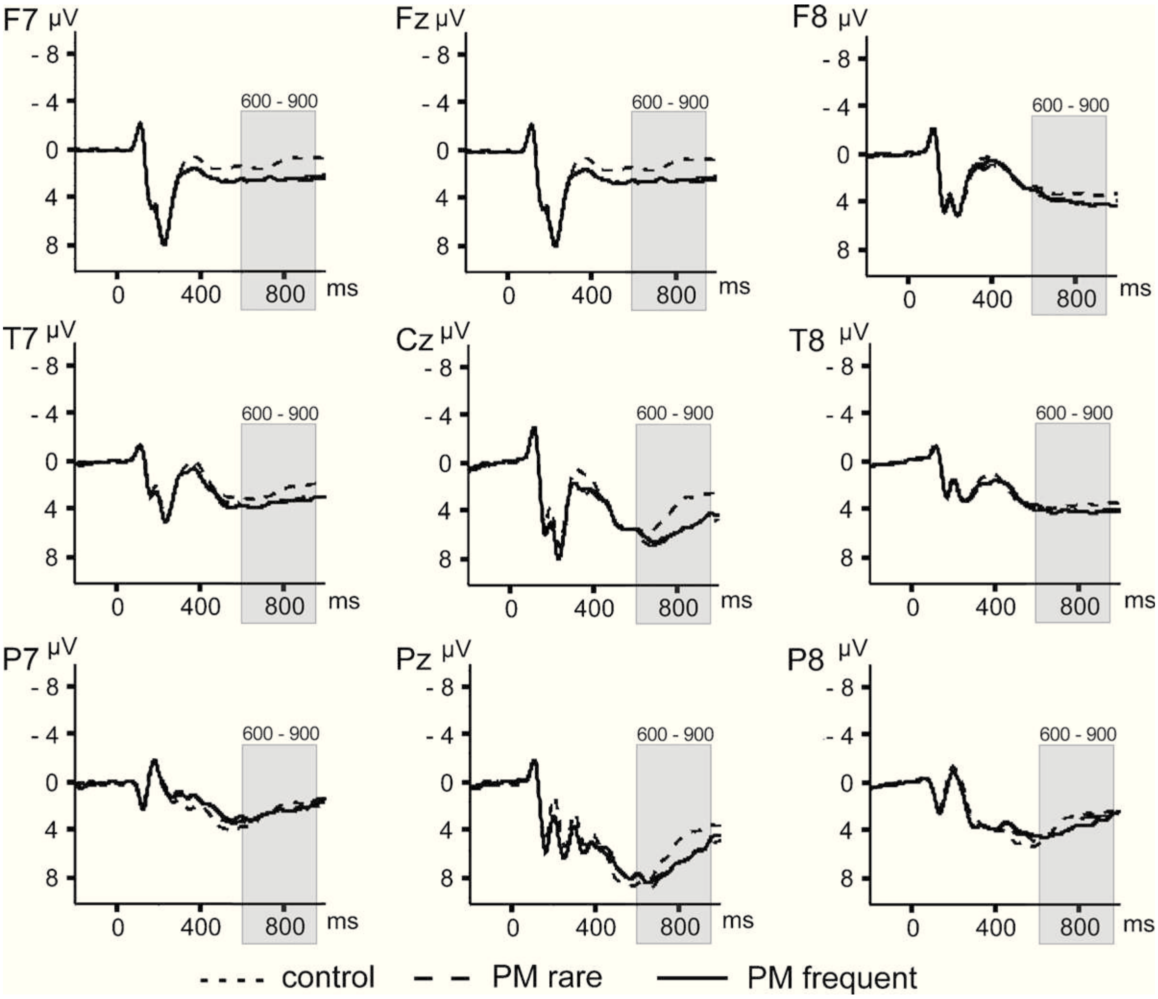


Fig. 1. ERP waveforms for ongoing-only/control (dotted line), PM-rare (dashed line) and PM-frequent conditions (solid line) at nine electrode sites used for the analyses of the correlate for PM monitoring ($n = 16$ participants). Reliable PM monitoring effects were found at electrodes F7, Fz, T7 and Cz for both PM conditions and at Pz for the PM-rare condition.

Table 1
ANOVA results for mean ERP amplitudes between 600 and 900 ms for ongoing-task trials in the control block, in the PM-frequent and in the PM-rare conditions.

Time window/Specific electrode site	Factor or contrast	F	df1, df2	p	ε	η^2_p
600–900 ms	Condition	5.25	2, 30	<.05		.26
	Condition × AP	4.83	4, 60	<.05	.57	.24
	Condition × LCR	6.65	4, 60	<.0001		.31
	Condition × LCR × AP	3.09	8, 120	<.05	.58	.17
Frontal (F7, FZ, F8)	Condition	11.01	2, 30	<.0001		.42
	Condition × LCR	7.87	4, 60	<.0001		.34
Central (T7, CZ, T8)	Condition	4.65	2, 30	<.05		.24
	Condition × LCR	5.12	4, 60	<.01	.66	.25
Parietal (P7, PZ, P8)	Condition	1.58	2, 30	.22		.10
	Condition × LCR	4.60	4, 60	<.01		.24
F7	PM rare versus CON	16.73	1, 15	<.01		.53
	PM frequent versus CON	20.55	1, 15	<.0001		.58
FZ	PM rare versus CON	22.63	1, 15	<.01		.60
	PM frequent versus CON	15.70	1, 15	<.01		.51
T7	PM rare versus CON	5.98	1, 15	<.05		.29
	PM frequent versus CON	7.16	1, 15	<.05		.32
CZ	PM rare versus CON	13.94	1, 15	<.01		.48
	PM frequent versus CON	5.41	1, 15	<.05		.27
PZ	PM rare versus CON	6.61	1, 15	<.05		.31
	PM frequent versus CON	3.69	1, 15	.07		.20

Note. CON – control block, AP – anterior-posterior electrode sites, LCR – left, central or right electrode sites. Contrast for single electrodes are reported for those electrodes with reliable main effects of condition. $N = 16$ participants.

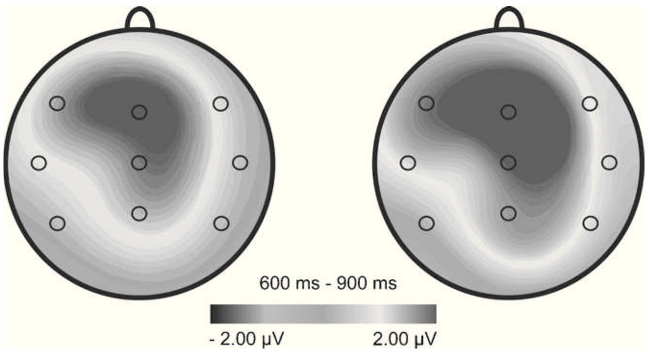


Fig. 2. Topographical map of the ERP correlate for prospective monitoring, for PM-rare (left) and PM-frequent condition (right) relative to ongoing-task trials in the control block between 600 and 900 ms. Note the extremely similar topography of both effects.

ERP amplitudes were not associated with higher PM target hit rate (*p*s > .14).

3.2.2. *ERP correlates of early attentional aspects of stimulus evaluation*

To determine whether PM monitoring also affected early attentional aspects of stimulus evaluation, we evaluated the activity during the P2 time window in a corresponding ANOVA. As illustrated in Fig. 3, two positive peaks emerged in the ERP waveforms around 200 ms following stimulus onset. By contrast, only one peak at 185 ms was evident in the difference waveforms (i.e. rare and frequent PM targets minus ongoing trials only, respectively), suggesting both peaks in the ERP waveforms originated in a common positivity superimposed on a smaller negative peak. Hence, the time window for the evaluation of the P2 effect was centered around the peak of the difference wave, and spanned 160–210 ms. To account for the mid-central topography of the P2, analyses focused on the following electrode sites: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

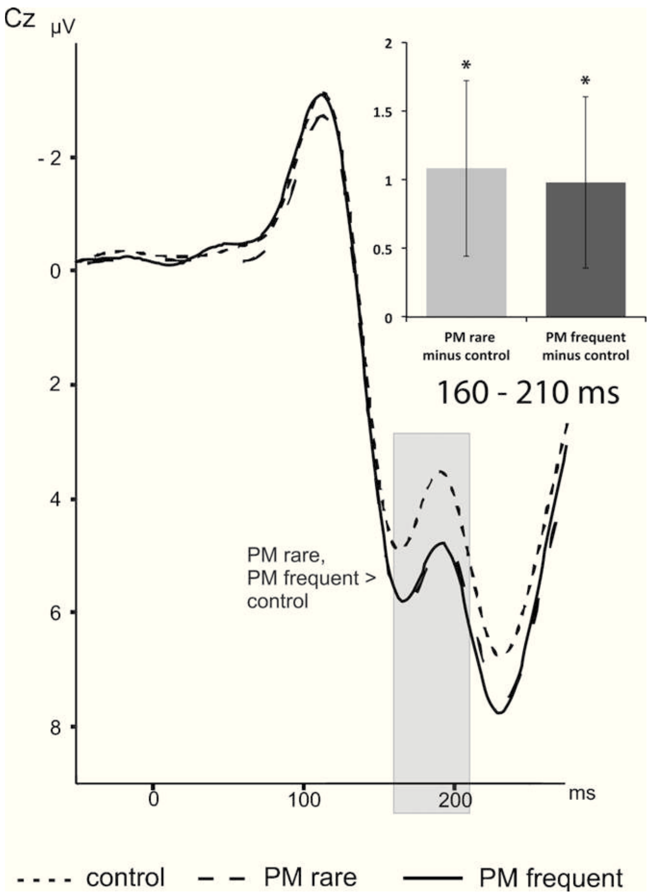


Fig. 3. Detailed view on the early part of the waveforms depicted in Fig. 1 at electrode site Cz. The P2 effect was evaluated at nine selected electrode sites between 160 and 210 ms, and reflected larger amplitudes during PM blocks relative to the control condition. The bar graphs (top right) illustrate reliable ERP amplitude differences across these nine electrode sites for both PM blocks relative to the control condition between 160 and 210 ms (error bars reflect 95% confidence intervals).

Consistent with the mid-central topography of the P2, a main effect of condition without interactions by electrode sites was observed, $F(2,30) = 9.68$, $p < .01$, $\eta_p^2 = .39$. As illustrated in more detail at electrode Cz in Fig. 3, planned contrasts revealed larger amplitudes for PM-rare relative to the control condition, 4.5 versus 3.4 μV, $F(1,15) = 13.03$, $p < .01$, $\eta_p^2 = .47$, and PM-frequent compared to the control condition, 4.4 versus 3.4 μV, $F(1,15) = 11.11$, $p < .01$, $\eta_p^2 = .43$. The order in which participants completed the PM rare and frequent conditions did not influence the magnitude of the P2 modulation [interaction with block order: $F(2,28) = 1.12$, $p = .34$].

3.2.3. *ERP correlates preceding prospective hits and misses*

To compare ERP correlates of PM monitoring preceding a PM target hit or miss, mean amplitudes between 160–210 ms and 600–900 ms for trials occurring within 30 s before a PM-rare target hit or miss were compared to ongoing trials in the control condition. As illustrated in Fig. 4 and detailed in Table 2, both before PM target hits and misses, a sustained positivity relative to the control condition was observed between 600 and 900 ms, consistent with the general pattern described above for all ongoing trials. As illustrated in Fig. 5, the P2 effect was reduced in magnitude to a trend, $F(2,26) = 3.01$, $p = .087$, $\eta_p^2 = .19$. Planned contrasts revealed a reliable P2 effect for trials preceding PM target hits, $F(1,14) = 14.36$, $p < .01$, $\eta_p^2 = .53$, 4.6 versus 3.3 μV, but not for trials preceding PM target misses (4.2 versus 3.3 μV, $p = .23$). To assess whether the P2 effect was completely absent or rather delayed in latency for trials preceding PM target misses, the following time window between 210 and 260 ms was also evaluated for this condition. This post-hoc analysis revealed a reliable P2 effect in this later time window, $F(1,14) = 30.11$, $p < .0001$, $\eta_p^2 = .68$, 7.0 versus 5.5 μV for trials preceding PM target misses versus ongoing-task only, indicating a delayed P2 effect for trials preceding PM target misses.

4. Discussion

Behavioral and ERP correlates of PM monitoring during blocks with rare and frequent PM target presentations were directly compared to an ongoing-task only condition in the same participants. For both rare and frequent PM conditions, we found behavioral interference costs in terms of longer RTs. A sustained positivity with frontal focus was identified on ongoing-task trials during both PM blocks. While PM target identification and RT costs were higher during the frequent relative to the rare PM target condition, the same sustained frontal positivity was observed during both PM blocks. Three open questions will be discussed in the upcoming sections: in Section 4.1 we will assess evidence for a functional association of the observed sustained frontal ERP effect with PM performance. In Section 4.2, the functional relevance of the observed P2 modulation for PM monitoring will be discussed. In Section 4.3, we will turn to the main question whether PM performance for rare versus frequent targets is supported by distinct monitoring processes.

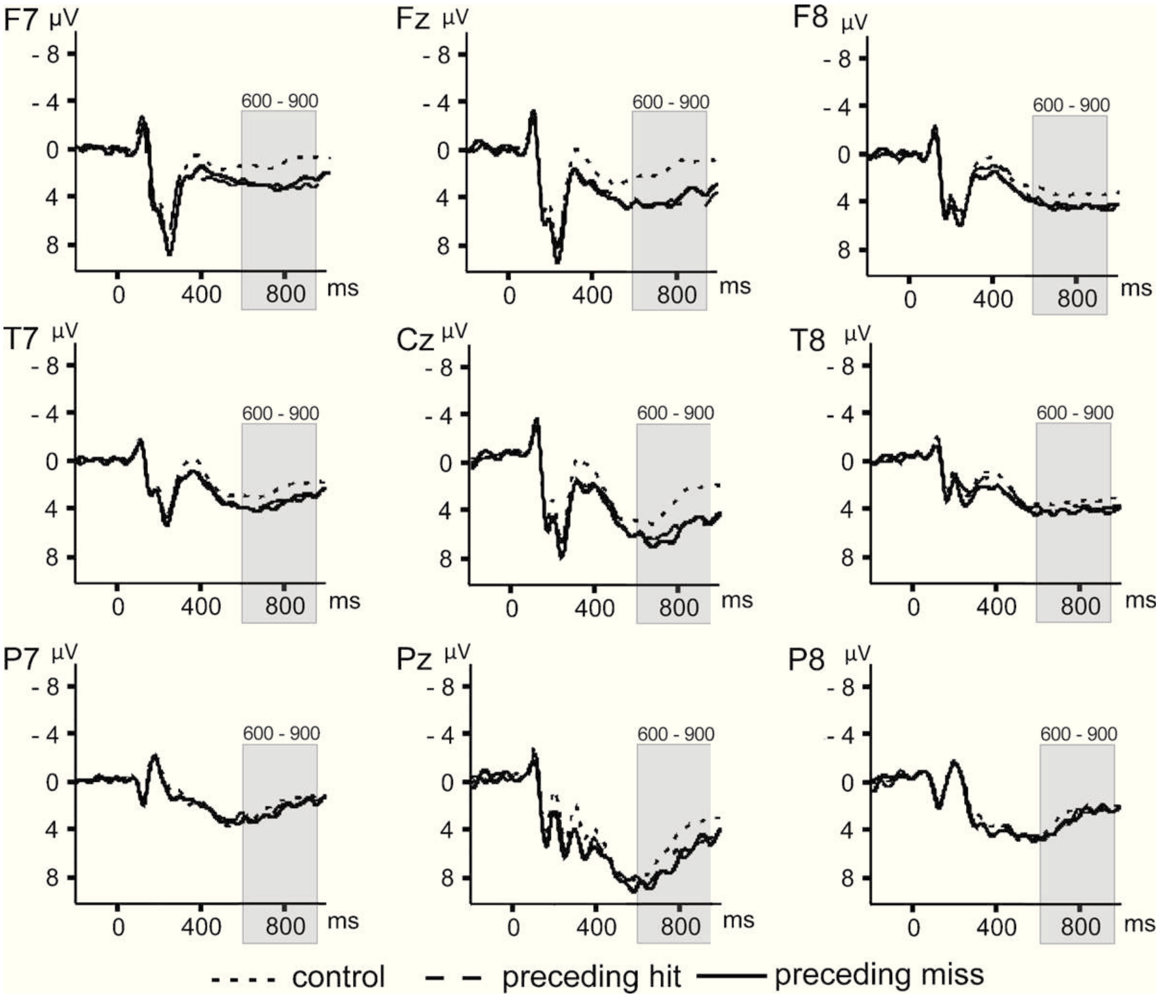


Fig. 4. ERP waveforms for trials preceding PM-rare target hits (dashed line) and misses (solid line) relative to ongoing trials in the control block (dotted line) at nine electrode sites used for the analyses of the correlate for PM monitoring preceding rare PM target hits and misses. Data are shown for 15 participants with sufficient artifact-free trials. Between 600 and 900 ms, we found reliable PM monitoring effects for trials preceding both PM hits and misses at F7, Fz, F8, T7, Cz, and Pz.

Table 2
ANOVA results for mean ERP amplitudes between 600 and 900 ms for ongoing-task trials in the control block, and ongoing-task trials preceding PM-frequent target hits and misses.

Time window/Specific electrode site	Factor or contrast	F	df1, df2	p	ε	$\eta_p^2 =$
600–900 ms	Condition	12.40	2, 28	<.0001		.47
	Condition × LCR	6.55	4, 56	<.01		.32
Left (F7, T7, P7)	Condition	7.35	2, 28	<.01		.34
	Condition × AP	5.61	4, 56	<.01	.62	.29
Midline (FZ, CZ, PZ)	Condition	13.54	2, 28	<.0001		.49
Right (F8, T8, P8)	Condition	6.07	2, 28	<.01		30
F7	Before hit versus CON	21.83	1, 14	<.0001		.61
	Before miss versus CON	19.43	1, 14	<.01		.58
FZ	Before hit versus CON	23.65	1, 14	<.0001		.63
	Before miss versus CON	23.57	1, 14	<.0001		.67
F8	Before hit versus CON	11.91	1, 14	<.01		.46
	Before miss versus CON	9.71	1, 14	<.01		.41
T7	Before hit versus CON	5.56	1, 14	<.05		.27
	Before miss versus CON	15.95	1, 14	<.01		.53
CZ	Before hit versus CON	13.60	1, 14	<.01		.49
	Before miss versus CON	15.84	1, 14	< .01		.53
PZ	Before hit versus CON	6.62	1, 14	<.05		.32
	Before miss versus CON	4.97	1, 14	<.05		.26

Note. $N = 15$ participants with sufficient artifact-free trials.

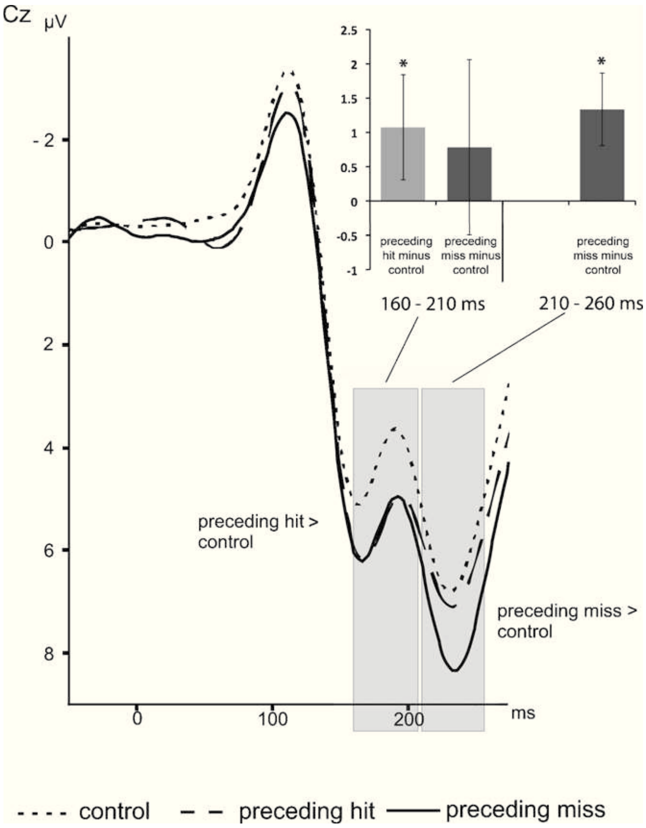


Fig. 5. Detailed view on the early part of the waveforms depicted in Fig. 4 at electrode site Cz. The P2 effect was evaluated at nine selected electrode sites between 160 and 210 ms, and larger amplitudes were observed during trials preceding PM target hits relative to the control condition. By contrast, for trials preceding PM target misses reliable P2 effects were evident between 210 and 260 ms compared to the control block. The bar graphs (top right) illustrate reliable ERP amplitude differences for trials preceding PM target hits relative to the control condition between 160 and 210 ms, and for trials preceding PM target misses between 210 and 260 ms (error bars reflect 95% confidence intervals).

4.1. *Is the sustained frontal ERP effect functionally related to PM performance?*

As predicted, a sustained positive ERP modulation was observed, with largest effects over fronto-central and left frontal electrode sites. Consistent with the notion that adopting a prospective retrieval mode is a prerequisite for successful PM performance (Guynn, 2003, 2008), the same ERP correlate was observed *irrespective* of PM target frequency. Although no firm conclusions regarding the underlying neural generators can be drawn based on the topography on the scalp, its frontal topography is in line with a growing number of neuroimaging studies employing fMRI (e.g., Burgess, Dumontheil, et al., 2007; Reynolds et al., 2009) and Positron Emission Topography (e.g., Burgess, Scott, & Frith, 2003) reporting PM-related activity in the anterior PFC. The anterior PFC is a region that has been implicated in various higher-order cognitive functions, for instance monitoring for contextual details during memory retrieval (e.g., Ranganath, Johnson, & D'Esposito, 2000) and the maintenance of an (episodic) retrieval mode (e.g., Duzel et al., 1999; see also Guynn, 2003). In general, and beyond specific task requirements, the anterior PFC has been associated with the coordination of several related cognitive operations in the service of a common behavioral goal (see Ramnani & Owen, 2004, for a review). Recent evidence also suggests a functional specificity within the anterior PFC: Whereas the medial part of the anterior PFC has been associated with stimulus-oriented processing (for a review, see Burgess, Dumontheil, et al., 2007; Burgess, Gilbert, et al., 2007), the lateral rostral part of the PFC in particular has been linked to stimulus-independent processing, in which the focus of attention is directed away from stimulus features and towards mental representations (Burgess, Gilbert, et al., 2007). Likewise, in order to successfully identify PM targets, an internal representation of a delayed intention needs to be maintained over time, consistent with the observed activity in anterior PFC during PM tasks (e.g., Burgess, Dumontheil, et al., 2007; Reynolds et al., 2009).

Consistent with the results by Reynolds et al. (2009), in the present investigation, a larger sustained frontal effect was associated with *faster* PM target detection. Although fMRI and ERPs are based on two very different types of physiological data, both measures were associated with the same behavioral index for PM performance. This converging evidence emphasizes the important function of sustained frontal neural activity for PM monitoring, in particular for maintaining delayed intentions.

4.2. *Are P2 modulations during PM blocks also functionally related to PM monitoring?*

Consistent with prior ERP studies (Chen et al., 2007, 2009; Knight et al., 2010), a P2 modulation was identified that distinguished the two PM blocks from the control condition. During both PM blocks, larger P2 amplitudes were observed between 160 and 210 ms. Consistent with the idea that these ERP effects reflect increased attention to specific stimulus aspects that are relevant for PM target identification, the present results and those obtained in previous ERP studies differ in the precise topography and timing of the observed P2 modulations. For instance, in the study by Knight et al. (2010), PM targets were defined as words appearing in a particular color (red), and corresponding ERP modulations of mean P2 amplitudes were most prominent over occipital regions, thought to be relevant for color discrimination. In our study, in which color was not a relevant feature for target identification, the P2 modulation had a central topography. Notably, the P2 modulation in the present study also did not differ between the two PM blocks, in which the same target features (initial letters) were relevant. By contrast, the observed P2 modulation was delayed for trials preceding rare PM target misses. This unexpected ERP finding may be related to less focused attention preceding PM target misses. It is conceivable that attention might be captured by stimulus features that are not necessarily relevant for PM target identification initially, and is re-directed at task-relevant features later on. However, this ERP pattern needs to be replicated in future studies before a conclusive interpretation can be drawn.

4.3. *Evidence for distinct monitoring as a function of PM target frequency?*

In this first investigation explicitly comparing monitoring for frequent and rare PM targets, behavioral data indicate performance differences between the PM blocks: we observed longer RTs and improved PM target identification for the frequent PM condition. These findings are consistent with behavioral data reported by Brandimonte, Ferrante, Feresin, and Delbello (2001) who manipulated the instructions and the training phase for a PM task. In a standard PM condition, the ongoing task was explained first, followed by PM task instructions. In a so-called vigilance condition, PM instructions emphasized the dual-task nature of the PM paradigm and participants were additionally reminded of the PM task during a training phase whenever they missed a relatively frequent PM target (which occurred in 17% of the training trials). Participants in the vigilance condition missed fewer PM targets and had slower RTs compared to those in the standard PM condition. These combined instruction- and training-manipulations likely influenced participants' perception of the relative importance of the PM task. Thus, participants who had reasons to believe that the PM task was very important for overall performance showed better PM performance and slower RTs in the ongoing task, consistent with increased effort devoted to the PM task (cf. Smith & Bayen, 2004). Similarly, in the present experiment, both higher PM target identification and longer RTs were observed for the frequent-PM condition. Hence, participants may have achieved higher PM target identification rates at the expense of ongoing-task performance. Increased RT costs for frequent compared to rare PM targets suggest that more attentional resources were devoted to monitoring for PM targets when frequently reminded of PM target occurrence. By contrast, ERPs provide no indication for distinct monitoring processes for frequent versus rare PM target events. How can this apparent discrepancy be explained?

While behavioral measures capture the result of a decision process, ERPs provide the opportunity to examine the cognitive processes leading up to this decision. From a participant's perspective, PM target frequency likely influenced further aspects of how to approach the given task set. Despite equivalent instructions, a task set with more frequently occurring PM targets may be perceived differently, leading to different strategies or allocation policies (Marsh et al., 2005). Recent evidence suggests that participants modify their initial attention allocation away from the PM task if PM targets are no longer presented, or fail to be presented when they are expected following PM instructions (Loft et al., 2008). Notably, participants whose RT cost increased for the PM-frequent relative to the PM-rare conditions showed a concomitant increase in ongoing-task accuracy ($r = .48$, $p < .05$). This pattern of performance suggests that participants emphasized ongoing-task speed in the PM-rare condition and ongoing-task accuracy in the PM-frequent condition, possibly reflecting different criteria how to perform the ongoing task (i.e., a speed-accuracy tradeoff; cf. Horn, Bayen, & Smith, 2011). Recent fMRI studies suggest that the striatum and the basal ganglia are important for such speed-accuracy tradeoff settings (Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; Forstmann et al., 2010). Both structures are located relatively deep within the brain. Hence, neural activity originating from these structures would not be easily detectable on the surface of the skull using ERPs, possibly accounting for similar monitoring effects across the two PM conditions. This differential criterion setting might also account for the fact that behavioral PM interference costs were larger for those participants who completed the PM-frequent condition before the PM rare condition. Importantly, for both counterbalancing conditions PM interference costs were reliable and effect sizes were similar.

Finally, when comparing PM-rare and PM-frequent conditions, further task characteristics are likely to vary inherently along with PM target frequency. Most notably, each correctly identified frequent PM target presumably serves as an additional reminder for the PM task (see also Loft et al., 2008, for a related argumentation). In the present investigation, this factor also affected RT-variability, which is thought to reflect fluctuations in the efficiency of PM monitoring (West, Krompinger, & Bowry, 2005). That is, responses to the three ongoing trials preceding PM target hits in the PM-rare condition were about 100 ms slower than those preceding target misses (Scullin, McDaniel, & Einstein, 2010; West et al., 2005). Notably, no corresponding RT differences were observed during the frequent-PM condition, suggesting that more constant PM monitoring might take place during the PM-frequent compared to the PM-rare condition. Thus, the mean prospective interference effect across trials might reflect a combination of the amount of monitoring plus further factors, such as criterion setting and response selection, and may not be a unique indicator of monitoring (see also Scullin et al., 2010, for a related argument; West et al., 2005).

Irrespective of target frequency, the same sustained ERP modulation was observed during PM blocks, in line with the idea of a general prospective retrieval mode (Guynn, 2003, 2008). By contrast, periodic target checking may or may not complement prospective retrieval mode for successful PM performance. Moreover, like criterion setting and response selection, target checking would not necessarily be consistently time-locked to stimulus presentation, and hence not be readily observable in ERPs or fMRI (cf. Reynolds et al., 2009, who also did not find evidence for transient activity during ongoing-task trials). Thus, it is conceivable that the perceived importance of the PM task mediated the frequency or intensity of target checking, and as a result may have influenced RTs, but not ERPs. Future studies need to address how target checking and differential criterion setting may be affected by higher perceived task importance.

5. Conclusion

The present data add to extant evidence that PM monitoring is functionally related to PM performance. A sustained frontal ERP effect appears to be a correlate of PM monitoring, and more specifically, maintaining an internal representation of the delayed intention until it is re-activated upon cue presentation. Identical ERP correlates of monitoring for frequent compared to rare PM targets support the notion that successful PM monitoring is associated with the adoption of a prospective retrieval mode irrespective of PM target frequency. Moreover, these results suggest that brain responses associated with PM monitoring in previous neuro-scientific research generalize to paradigms using considerably fewer targets. Due to an increase of perceived importance of the PM-frequent condition, participants may be more inclined to monitor for frequent versus rare PM targets at the expense of slower RTs in the ongoing task. While such task priorities can be expected to heavily influence RTs, they would not necessarily be detectable in ERPs. A second process, the modulation of attention directed at relevant target features, plays an important role for detecting PM targets. Our electrophysiological data suggest distinct roles of this P2 modulation and the sustained frontal effects for a more specific attentional focus on relevant prospective target features and prospective monitoring, respectively.

Acknowledgements

The authors would like to thank Nicole Röder for assistance in stimulus preparation, participant recruitment, and EEG data preprocessing, Nicole Röder and André Haese for data collection, and Siegmund Switala for technical assistance.

References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends in Neurosciences*, *33*, 10–16.

Brandimonte, M. A., Ferrante, D., Feresin, C., & Delbello, R. (2001). Dissociating prospective memory from vigilance processes. *Psicologica*, *22*, 97–113.

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Science*, *11*, 290–298.

Burgess, P. W., Gilbert, S. J., & Dumontheil, I. (2007). Function and localization within rostral prefrontal cortex (area 10). *Philosophical Transactions of the Royal Society London Section B Biological Sciences*, *362*, 887–899.

Burgess, P. W., Scott, S. K., & Frith, C. D. (2003). The role of the rostral frontal cortex (area 10) in prospective memory: A lateral versus medial dissociation. *Neuropsychologia*, *41*, 906–918.

Chen, Y., Huang, X., Jackson, T., & Yang, H. (2009). Effect of characteristics of target cues on task interference from prospective memory. *Neuroreport*, *20*, 81–86.

Chen, Y., Huang, X., Yang, H., Ren, G., & Yue, C. (2007). Task interference from event-based prospective memory: An event-related potentials study. *Neuroreport*, *18*, 1951–1955.

Duzel, E., Cabeza, R., Picton, T. W., Yonelinas, A. P., Scheich, H., Heinze, H. J., et al. (1999). Task-related and item-related brain processes of memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 1794–1799.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science*, *14*, 286–290.

Forstmann, B. U., Anwander, A., Schafer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., et al. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 15916–15920.

Graf, P., & Uttl, B. (2001). Prospective memory: A new focus for research. *Consciousness and Cognition: An International Journal*, *10*, 437–450.

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology*, *38*, 245–256.

Guynn, M. J. (2008). Theory of monitoring in prospective memory: Instantiating a retrieval mode and periodic target checking. In Kliegel, McDaniel, & Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 53–76). New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates.

Horn, S. S., Bayen, U. J., & Smith, R. E. (2011). What can the diffusion model tell us about prospective memory? *Canadian Journal of Experimental Psychology*, *65*, 69–75.

Kliegel, M., McDaniel, M. A., & Einstein, G. O. (Eds.). (2008). *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives*. New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates.

Knight, J. B., Ethridge, L. E., Marsh, R. L., & Clementz, B. A. (2010). Neural correlates of attentional and mnemonic processing in event-based prospective memory. *Frontiers in Human Neuroscience*, *4*, 1–13.

Loft, S., Kearney, R., & Remington, R. (2008). Is task interference in event-based prospective memory dependent on cue presentation? *Memory & Cognition*, *36*, 139–148.

Loft, S., & Yeo, G. (2007). An investigation into the resource requirements of event-based prospective memory. *Memory & Cognition*, *35*, 263–274.

Luck, S. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

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

Luck, S., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Science*, *4*, 432–440.

Makeig, S., Jung, T. P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 10979–10984.

Marsh, R. L., Hicks, J. L., Cook, G. I., Hansen, J. S., & Pallos, A. L. (2003). Interference to ongoing activities covaries with the characteristics of an event-based intention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 861–870.

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*, 68–75.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, *14*, 127–144.

Monsell, S. (2003). Task switching. *Trends in Cognitive Science*, *7*, 134–140.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*, 184–187.

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

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience*, *5*, 184–194.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *Journal of Neuroscience*, *20*, 108–112.

Reynolds, J. R., West, R., & Braver, T. (2009). Distinct neural circuits support transient and sustained processes in prospective memory and working memory. *Cerebral Cortex*, *19*, 1208–1221.

Ruijter, J., Lorist, M. M., Snel, J., & De Ruiter, M. B. (2000). The influence of caffeine on sustained attention: An ERP study. *Pharmacology, Biochemistry and Behavior*, *66*, 29–37.

Ruz, M., & Nobre, A. C. (2008). Attention modulates initial stages of visual word processing. *Journal of Cognitive Neuroscience*, *20*, 1727–1736.

Scullin, M. K., McDaniel, M. A., & Einstein, G. O. (2010). Control of cost in prospective memory: Evidence for spontaneous retrieval processes. *Journal of Experimental Psychology: Learning Memory and Cognition*, *36*, 190–203.

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 347–361.

Smith, R. E., & Bayen, U. J. (2004). A multinomial model of event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 756–777.

Smith, R. E., & Bayen, U. J. (2005). The effects of working memory resource availability on prospective memory: A formal modeling approach. *Experimental Psychology*, *52*, 243–256.

Smith, R. E., Hunt, R. R., McVay, J. C., & McConnell, M. D. (2007). The cost of event-based prospective memory: Salient target events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 734–741.

West, R. (2011). The temporal dynamics of prospective memory: A review of the ERP and prospective memory literature. *Neuropsychologia*, *49*, 2233–2245.

West, R., & Bowry, R. (2005). Effects of aging and working memory demands on prospective memory. *Psychophysiology*, *42*, 698–712.

West, R., Bowry, R., & Krompinger, J. (2006). The effects of working memory demands on the neural correlates of prospective memory. *Neuropsychologia*, *44*, 197–207.

West, R., & Krompinger, J. (2005). Neural correlates of prospective and retrospective memory. *Neuropsychologia*, *43*, 418–433.

West, R., Krompinger, J., & Bowry, R. (2005). Disruptions of preparatory attention contribute to failures of prospective memory. *Psychonomic Bulletin & Review*, *12*, 502–507.

West, R., McNerney, M. W., & Krauss, I. (2007). Impaired strategic monitoring as the locus of a focal prospective memory deficit. *Neurocase*, *13*, 115–126.

West, R., McNerney, M. W., & Travers, S. (2007). Gone but not forgotten: The effects of cancelled intentions on the neural correlates of prospective memory. *International Journal of Psychophysiology*, *64*, 215–225.

West, R., Scolaro, A. J., & Bailey, K. (2011). When goals collide: The interaction between prospective memory and task switching. *Canadian Journal of Experimental Psychology*, *65*, 38–47.

Zöllig, J., West, R., Martin, M., Altgassen, M., Lemke, U., & Kliegel, M. (2007). Neural correlates of prospective memory across the lifespan. *Neuropsychologia*, *45*, 3299–3314.