

2 Empirie

In vier Untersuchungen werden die vorgestellten Hypothesen differenziert, operationalisiert und geprüft. Dabei ist empirische Teil zweigeteilt. Im ersten Teil stehen die Hypothesen zur Planungszeit und Speicherzahl im Mittelpunkt, im zweiten kommen die Hypothesen zu den metakognitiven Parametern hinzu. Die erste Studie beider Teile ist jeweils eine Pilotstudie mit kleinerer Versuchspersonenzahl, deren Ergebnisse in die Planung der jeweils folgenden Studie einfließen.

Jede Studie wird zunächst für sich diskutiert; ein Vergleich aller Studien im Rahmen des SWIEGS-Modells findet erst in der Schlußdiskussion statt.

2.1 *Erster empirischer Teil: Planungszeit und Speicherzahl*

Im ersten empirischen Teil steht vor allem die Planungszeit im Mittelpunkt. Basis des SWIEGS-Modells ist die Annahme verschiedener, unterschiedlich aufwendiger Prozesse bei der Speicherwahl je nach Geläufigkeit der zu bearbeitenden Aufgabe. Vor der Prüfung differenzierter Zusammenhänge ist es deswegen erforderlich, den Nachweis zu erbringen, daß tatsächlich Unterschiede im Zeitbedarf für die Speicherwahl vorhandensind.

2.1.1 Studie 1: Planungszeit, Speicherzahl und Umgebungsvariation

Studie 1 ist eine Pilotstudie, in der zwei basale Hypothesen einer ersten empirischen Prüfung unterzogen werden: Je ungeläufiger die Aufgabe, um so länger die Planungszeit, um einen passenden Speicher zu finden (H1); die Aufgabengruppe mit sehr geläufigen Aufgaben wird mit weniger unterschiedlichen Speichern bearbeitet als die Aufgabengruppen mit weniger geläufigen Aufgaben (H6).

Erfragt wird das Verhalten im Alltag. Es wäre daher möglich, daß die Laborsituation Effekte zerstört, die in der anregenderen Alltagsumgebung durchaus vorhanden sind. Eine solche Wirkung der Umgebung ist aus Untersuchungen zum Alltagsgedächtnis durchaus bekannt (z.B. 'strategic time monitoring', Ceci & Bronfenbrenner, 1985). Zur Planung weiterer Untersuchungen ist es deswegen erforderlich zu wissen, ob die Umwelt, in der der Versuch durchgeführt wird, diesen Einfluß auf die vorhergesagten Effekte hat. Deswegen wird die Umgebung, in der der Versuch durchgeführt wird, als unabhängig Variable in das Design aufgenommen.

Material

In einer Voruntersuchung sollten 30 Studenten und Studentinnen typische Repräsentationsaufgaben aus ihrem Alltag und die Art ihrer Lösung benennen. Aus diesem Pool wurden die 17 am häufigsten genannten Aufgaben ausgewählt. In zwei Stufen wurden nun geläufige Elemente der jeweiligen Aufgabe durch ungeläufige ersetzt. Eine Aufgabe lautete beispielsweise: „Morgen ein Buch in die Bibliothek zurückbringen“. Im studentischen Alltag ist diese Aufgabe sehr geläufig; die in der Voruntersuchung am häufigsten genannte Lösung war, daß das Buch an eine Stelle an der Ausgangstür plaziert wird, wo es beim Verlassen der Wohnung am nächsten morgen auffallen muß. Ein konstituierendes Element dieses Aufgabentyps ist, daß ihre Erledigung nicht zu weit in der Zukunft liegt; ein zweites, daß der Gegenstand, mit dem die Aufgabe durchgeführt wird, präsent ist. Durch Ersetzen dieser für diesen Aufgabentyp geläufigen Elemente wird die Aufgabe ungeläufig: „Du mußt in acht Wochen ein Buch von der Bibliothek abholen“. Der Zeitraum bis zum Erledigen der Aufgabe ist hier gestreckt und das Objekt, mit dem die Aufgabe durchgeführt wird, ist nicht mehr greifbar; die Aufgabe kann dadurch auch nicht mehr mit dem gewohnten Speicher gelöst werden. In einem zweiten Schritt wurden diese beiden Elemente der Aufgabe durch noch ungeläufigere ersetzt: „Du mußt 1996 ein Buch vorbestellen“. Da die Untersuchung im Frühjahr 1995 durchgeführt wurde, lag der Termin in einem Zeitraum, der nicht mehr vom aktuellen Kalender erfaßt wurde. Das Objekt, mit dem die Aufgabe durchgeführt werden muß, ist noch weniger greifbar als in der Vorversion, da es um den allerersten Kontakt - das Bestellen des Buches - geht. In der Version, in der das Buch abzuholen war, mußte dieser Schritt schon geschehen sein; der Titel des Buches ist dadurch schon extern gespeichert - nämlich in der Bibliothek - was die Anzahl möglicher Lösungen erhöht. War in der Vorversion die Lösung über einen geläufigen Speicher wie den Terminkalender noch problemlos möglich, erfordert die Lösung der letzten Version weitere Überlegungen.

Es wurde besonders darauf geachtet, daß die Aufgaben nicht aus einem anderen Kontext heraus wieder geläufig wurden. So ist das Abholen eines Buches aus der Bibliothek oder das Bestellen eines Buches in der Bücherei durchaus eine geläufige Aufgabe. Trotzdem können gewohnte Speicherungen dieser Aufgaben nicht greifen: ist ein Buch aus der Bibliothek abzuholen, informiert die Bibliothek in der Regel den Benutzer. Will man ein Buch bestellen, so kann man das unmittelbar tun. Ziel der Verfremdung war, daß die Aufgaben nach der ersten Verfremdung nicht mehr mit dem ursprünglich genannten Speicher zu lösen waren und nach der zweiten Verfremdung mit keinem geläufigen Speicher überhaupt. Hier ist noch einmal wichtig zu betonen, daß die Ungeläufigkeit der Aufgaben kein Aspekt der Aufgaben an sich ist, sondern ein Resultat der Erfahrungen mit Repräsentationsaufgaben. Ungeläufig sollten diese Aufgaben für die Teilpopulation der Versuchspersonen - Studenten und Studentinnen - sein; eine andere Teilpopulation würde eine andere Konstruktion erfordern. So ist die

letzte Version der beschriebenen Aufgabe - das Bestellen eines Buches im nächsten Jahr - für Buchhändler sicherlich nicht ungeläufig. Hier müßten andere Elemente der Aufgabe ausgetauscht werden, um denselben Effekt zu erzielen.

In entsprechender Weise wurden die übrigen 16 Aufgaben verfremdet, so daß ein Pool von 17 Aufgaben in drei Geläufigkeitsstufen, also insgesamt 51 Aufgaben, vorlag (siehe Anhang 5.1.1).

Versuchsdurchführung

Die Aufgaben wurden in zufälliger Abfolge auf Karteikarten präsentiert. Wenn die Probanden eine Lösung wußten, sollten sie das durch ein „Okay“ dem Versuchsleiter mitteilen. Der Versuchsleiter maß die Zeit vom Beginn der Präsentation bis zu dieser Äußerung mit einer Stoppuhr. Danach nannte die Versuchsperson den Speicher und der Versuchsleiter notierte ihn. Durch die Teilung in zwei Phasen sollte gewährleistet werden, daß die Probanden den Prozeß der Planung tatsächlich abgeschlossen hatten, wenn sie den Speicher nannten. Abschließend schätzten die Probanden die Geläufigkeit der Aufgabe auf einer Skala von eins (sehr geläufig) bis fünf (sehr ungeläufig) ein. Diese Skala lag als Erinnerungsstütze während des gesamten Versuchs vor dem Probanden.

Die Versuchspersonen wurden ausführlich über ihre Aufgabe instruiert. Dabei wurde besonders betont, daß es nicht um eine möglichst schnelle Lösung der Repräsentationsaufgaben geht, sondern vielmehr darum, die Lösung zu finden, die sie tatsächlich nehmen würden, wenn ihnen diese Aufgabe in ihrem Alltag begegnen würde. Schließlich wurde das Procedere mit drei Übungsaufgaben trainiert, um insbesondere die Trennung des Endes der Planungszeit durch ein „Okay“ von der Nennung des gewählten Speichers zu üben.

Versuchspersonen und Design

Zehn Männer und zehn Frauen nahmen an der Untersuchung teil. Alle Probanden waren Studenten und Studentinnen der Psychologie, die für die Teilnahme einen Versuchspersonenschein erhielten. Die Hälfte der Probanden führte den Versuch in ihrer häuslichen Umgebung durch, die andere Hälfte im psychologischen Labor. In beiden Gruppen waren Männer und Frauen gleichverteilt; innerhalb dieser Beschränkung erfolgte die Zuweisung zu einer der beiden Gruppen per Zufall. Das Design war also 2*3 gemischt faktoriell mit der Umgebung - Labor vs. zu Hause - als unabhängiger Variable und den drei Geläufigkeitsstufen der Repräsentationsaufgaben als Meßwiederholungsvariable.

Auswertungsprinzip der Studie 1-4

Das Prinzip der Auswertung ist für alle Studien gleich. Signifikanzniveau ist die übliche fünfprozentige Zufallswahrscheinlichkeit; zur besseren Orientierung werden aber alle Ergebnisse bis

zum 10%-Niveau erwähnt. Zur statistischen Absicherung des Effektes der unabhängigen Variablen und der Meßwiederholungsvariablen werden jeweils eine ein oder zweifaktorielle Varianzanalyse mit Meßwiederholung berechnet. Dazu diente die Prozedur MANOVA des Statistikpaketes SYSTAT in der Version 5.0. Die Berechnung einer Varianzanalyse bei einem Meßwiederholungsdesign hat als Voraussetzung die Homogenität der Varianz-Kovarianz Matrix (z.B. Bortz, 1979, S. 437ff). Da diese Voraussetzung bei realen Datensätzen fast nie erfüllt ist, berechnet man statt dessen weitere, Meßwiederholungsanalysen, die diese Voraussetzung nicht erfordern (z.B. Wilks' Lambda oder Pillai Trace, vgl. Systat Manual, 1987, p.581). Kommen diese Analysen zum gleichen Ergebnis wie der F-Test, kann man davon ausgehen, daß die Verletzung der Voraussetzung nicht zu einer Verzerrung der Ergebnisse geführt hat (Systat Manual, 1987, p. 582). Der Einfachheit halber werden signifikante Ergebnisse der F-Tests nur dann aufgeführt, wenn sie mit den zusätzlichen Analysen übereinstimmen; diese selbst werden nicht aufgeführt. Soweit die weitere Abklärung signifikanter Haupteffekte oder Interaktionen es erfordert, werden weitere F-Tests zur Berechnung der isolierten Wirkung eines einzelnen Faktors sowie zweiseitige t-Tests zum Paarvergleich bei unabhängigen Stichproben bzw. bei Meßwiederholungsvariablen durchgeführt. Die Gefahr, durch die Berechnung mehrerer t-Test per Zufall ein signifikantes Ergebnis zu bekommen (erhöhter α - Fehler) wäre durch Testverfahren mit einer entsprechenden α - Fehler Adjustierung (z.B. Scheffé - Test) möglich. Dieser zusätzliche Test ist aber nur erforderlich, wenn eine Interpretation wesentlich auf der Signifikanz eines bestimmten Paarvergleichs ruht (vgl. z.B. Czienskowski, 1996). In den hier diskutierten Studien dienen t-Tests aber ausschließlich der Erhellung von komplexeren Zusammenhangsmustern; das Ergebnis eines einzelnen t-Tests wird also nirgends zentral. Deswegen werden über t-Tests hinaus keine weitergehenden Testverfahren eingesetzt.

Ergebnisse

Die Ergebnisse zu Studie 1 gliedern sich nach drei abhängigen Variablen: Die Einschätzung der Geläufigkeit der Aufgaben, die zur Lösung der Aufgabe erforderliche Zeit (Planungszeit) und schließlich die gewählten Speicher insgesamt sowie die Anzahl unterschiedlicher Speicher, die zur Lösung einer Aufgabengruppe von jedem einzelnen Probanden eingesetzt wurden.

Die Einschätzung der Geläufigkeit der Aufgaben sinkt mit der Aufgabengruppe (siehe Abb. 4). Dabei werden die Aufgaben in der natürlichen Umgebung extremer eingeschätzt: Geläufige Aufgaben wirken geläufiger und ungeläufige Aufgaben ungeläufiger als im Labor.

Die Varianzanalyse zeigte einen signifikanten Haupteffekte für die Aufgabengruppe $F(2,36)=58.88$, $p<.001$ sowie die Signifikanz der Interaktion zwischen Aufgabengruppe und

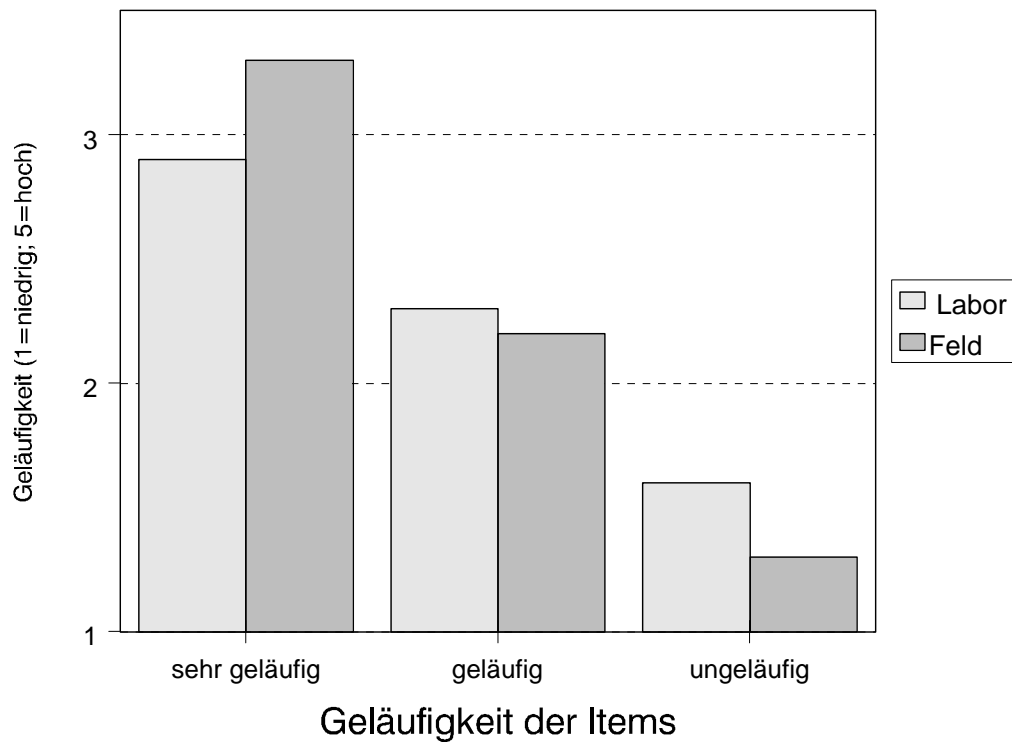


Abb. 4 Die Geläufigkeitseinschätzung der Aufgaben im der Labor- und in der Feldbedingung.

Umgebung, $F(2,36)=3.77$, $p<.05$. Der Haupteffekt der Aufgabengruppe bleibt auch bei der Analyse einzelner Effekte für beide Umgebungen signifikant: $F(2,18)= 10.73$, $p=.001$ für die Laborbedingung und $F(2,18),=112.25$, $p<.001$ für die häusliche Umgebung. T-Tests für Paarvergleiche unabhängiger Variablen zwischen Aufgaben gleicher Aufgabengruppen unter der Labor vs. unter der häuslichen Umgebung ergeben keinerlei signifikanten Unterschiede in der Einschätzung der Aufgaben. Dagegen zeigen alle t-Tests für den Paarvergleich von Meßwiederholungsvariablen sowohl in der gewohnten Umgebung als auch im Labor signifikante oder hochsignifikante Unterschiede in der Einschätzung der drei Aufgabengruppen (vgl. Tab. 2).

Einschätzung der Geläufigkeit der Aufgaben (Mittelwertsvergleiche über t-Tests)

<i>Bedingung:</i>	<i>Gewohnte Umgebung</i>	<i>Labor</i>
Geläufigkeitsgruppe:		
hoch vs.mittel	$t(9)=10.34^{***}$	$t(9)=3,27^*$
mittel vs.gering	$t(9)=6.84^{***}$	$t(9)=2.53^*$
gering vs. hoch	$t(9)=14.22^{***}$	$t(9)=3.51^{**}$

Tab. 2 Ergebnisse der t-Tests der Paarvergleiche für die Geläufigkeitseinschätzungen der drei Aufgaben-gruppen.

* für $p<.05$, **für $p<.01$, ***für $p<.001$.

Die Planungszeiten sind in Abb. 5 dargestellt. Die Planungszeiten sind am kürzesten für die geläufigsten Aufgaben und steigen mit sinkender Geläufigkeit. Dieser Gradient ist in der natürlichen Umwelt steiler als im Labor. Während die Lösung geläufiger Aufgaben in beiden Umwelten gleich lange benötigt, sind die Lösungszeiten mit sinkender Geläufigkeit der Aufgabe in der natürlichen Umwelt deutlich länger als im Labor.

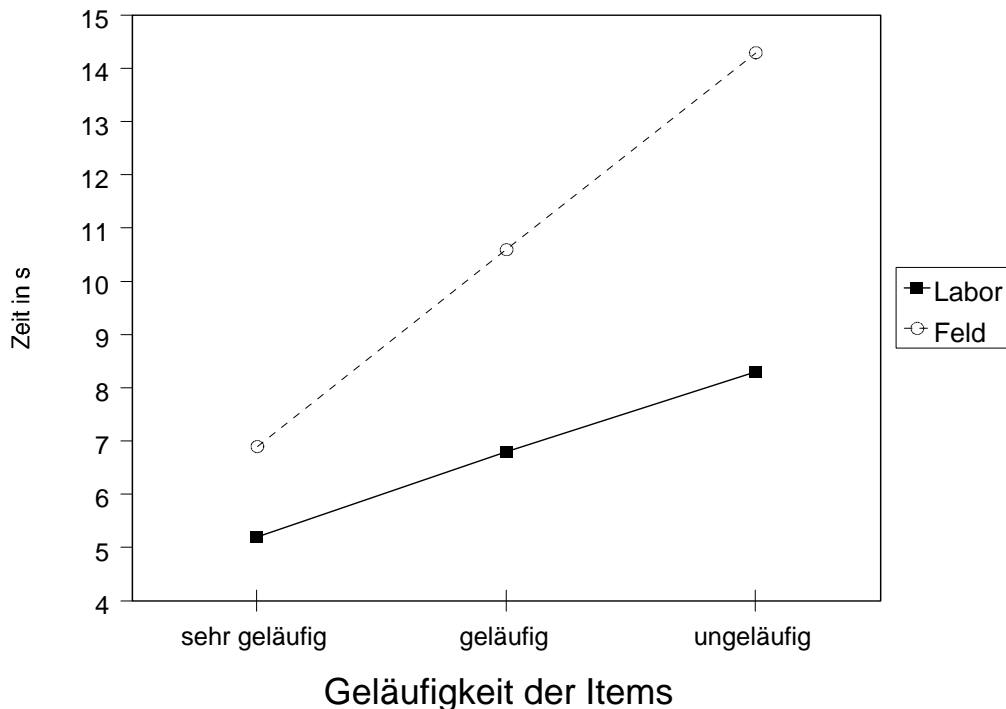


Abb. 5 Die Zeit, bis eine passender Speicher genannt wurde, getrennt für die die Geläufigkeit der Aufgaben und Labor- vs. Feldbedingung.

Die Varianzanalyse zeigt signifikante Haupteffekte sowohl für die unabhängige Variable Umgebung, $F(1,18)=6.28$, $p<.05$ als auch für die Meßwiederholungsvariable Geläufigkeit, $F(2,6)=22.66$, $p<.001$ sowie eine signifikante Interaktion beider Variablen, $F(2,36)=3.76$, $p<.05$. Der Haupteffekt der Aufgabengruppe bleibt auch bei der Analyse einzelner Effekte - für beide Umgebungen getrennt betrachtet - signifikant: $F(2,18)= 4.71$, $p<.05$, für die Laborbedingung und $F(2,18)=19.423$, $p<.001$ für die häusliche Umgebung. T-Tests für den Mittelwertsvergleich unabhängige Variablen zeigen, daß sich die Planungszeiten in beiden Umwelten nicht bei sehr geläufigen, $t(18)= -1,48$, $p>.05$, wohl aber bei geläufigen, $t(18)=-2,46$, $p<.05$, und ungeläufigen Aufgaben, $t(18)=-2,49$, $p<.05$, unterscheiden. T-Tests für Meßwiederholungsvariablen zeigen in der natürlichen Umgebung einen signifikanten Unterschied zwischen allen Planungszeiten: $t(9)=4.46$, $p<.05$, zwischen sehr geläufigen und geläufigen, $t(9) =3.19$, $p<.01$, zwischen geläufigen und ungeläufigen und $t(9) =4.99$, $p<.01$, zwischen sehr geläufigen und ungeläufigen Aufgaben. Dagegen verfehlen die Zeitunterschiede im Labor teilweise knapp das Signifikanzniveau: $t(9) =2.1$, $p=.065$, zwischen sehr geläufigen und geläufigen,

$t(9)=4.46$, $p<.05$, zwischen geläufigen und sehr ungeläufigen und $t(9) =1.98$, $p=.085$, zwischen sehr geläufigen und ungeläufigen Aufgaben. Der korrelative Zusammenhang zwischen der eingeschätzten Geläufigkeit der Aufgaben und der Planungszeit beträgt $r=.8$, $p<.001$.

Zur Lösung der Repräsentationsaufgaben wurden insgesamt 41 verschiedene Speicher genannt (siehe Tab. 3).

Eingesetzte Speicher

Speicher Anzahl der Nennungen: absolut / in %				Speicher Anzahl der Nennungen: absolut / in %			
<u>%</u>				22. Zettel in Küche	6	0,5%	
1. Gedächtnis	419	34,9%		23. Küchenkalender	5	0,4%	
2. Terminkalender	129	10,8%		24. Notizbuch	5	0,4%	
3. Notizzettel auf Schreibtisch	54	4,5%		25. Stadtplan	5	0,4%	
4. Wandkalender	50	4,2%		26. Telefonblock	5	0,4%	
5. Pinnwand	50	4,2%		27. Zettel in Rucksack	4	0,3%	
6. Wecker	34	2,8%		28. Telefonischer Weckdienst	3	0,3%	
7. Zettel in Terminkalender	32	2,7%		29. Zettel im Handschuhfach	3	0,3%	
8. Allzweckbuch	22	1,8%		30. Zettel in Geldbörse	3	0,3%	
9. Gegenstand plazieren	20	1,7%		31. Geburtstagskalender	3	0,3%	
10. Andere Person	19	1,6%		32. Zettelablage	3	0,3%	
11. Ordner für Uni	16	1,3%		33. Zettel auf Kommode	2	0,2%	
12. Jahreskalender	13	1,1%		34. Zettel auf Fußboden	2	0,2%	
13. Zettel an der Wohnungstür	13	1,1%		35. Zettel an Fernseher	1	0,1%	
14. Kalender Einband hinten	13	1,1%		36. Adreßbuch in Tasche	1	0,1%	
15. Anderer Terminkalender	12	1,0%		37. Kalender und Wecker	1	0,1%	
16. Allgemeiner Ordner	11	0,9%		38. Abizeitung	1	0,1%	
17. Einkaufsliste	11	0,9%		39. Zettel an Ofen	1	0,1%	
18. Computer	10	0,8%		40. Karteikasten	1	0,1%	
19. Zettel an Schreibtisch und Tür	8	0,7%		41. Tagebuch	1	0,1%	
20. Zettel über Schreibtisch	7	0,6%		Keine Angaben	14	1,2%	
21. Adreßbuch	7	0,6%					
				Gesamt	1020	100%	

Tab. 3 Zur Lösung der Repräsentationsaufgaben eingesetzte Speicher; Prozentangaben auf eine Dezimalstelle hinter dem Komma gerundet. Die Auflistung ist, bis auf Angleichung von Synonymen, unverändert.

~~Die Speicher werden sehr unterschiedlich oft genannt. Über 45% der Aufgaben werden über nur zwei Speicher (Gedächtnis und Terminkalender) gelöst. Dreizehn weitere Speicher werden noch in mehr als einem Prozent der Fälle genannt und 26 Speicher in unter einem Prozent. Der Notizzettel taucht explizit in insgesamt 14 Speichermöglichkeiten auf. Faßt man diese unter dem Oberbegriff Notizzettel zusammen, wäre dieser Speicher mit insgesamt 138 Nennungen oder 13,5% der zweithäufigst genannte Speicher. In diesem Fall würden über 59% der Aufgaben über nur drei verschiedene Speicher gelöst.~~

Für jeden Probanden wurde errechnet, mit wie vielen unterschiedlichen Speichern er die Repräsentationsaufgaben der drei Geläufigkeitsstufen gelöst hat. Maximal wären siebzehn unterschiedliche Speicher möglich: wenn nämlich für jede einzelne Aufgabe ein anderer Speicher genannt würde. Weniger als siebzehn Speicher ergeben sich, wenn ein Speicher mehrfach als Lösung für verschiedene Aufgaben auftaucht. Je ungeläufiger die Aufgabengruppe, desto mehr unterschiedliche Speicher wurden zu ihrer Lösung eingesetzt. Dieser Effekt ist in der gewohnten Umgebung stärker als in der Laborumwelt (siehe Abb. 6)

Die zweifaktorielle Varianzanalyse mit Meßwiederholung auf dem Faktor Geläufigkeit zeigt, daß der Faktor der Umgebung das Signifikanzniveau knapp verfehlt, $F(1,18)=3.17$, $p=.09$, aber sowohl der

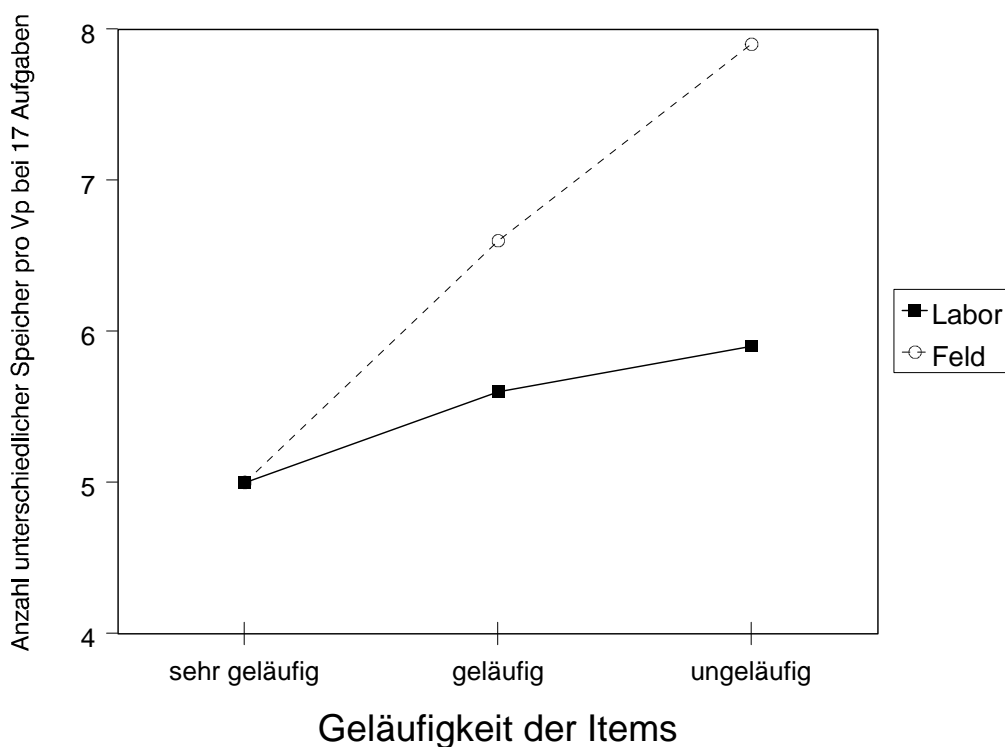


Abb. 6 Die Anzahl unterschiedlicher zur Lösung der Aufgaben genannten Speicher, getrennt nach der Geläufigkeit der Aufgabe. Maximal wären 17 unterschiedliche Speicher möglich.

Haupteffekt der Geläufigkeit signifikant wird, $F(2,36)=12.22$, $p<.001$, als auch die Interaktion zwischen Geläufigkeit und Umgebung, $F(2,36)=3.36$, $p<.05$. Der Haupteffekt der Geläufigkeit bleibt - auch für beide Umgebungen getrennt betrachtet - signifikant: $F(2,18)= 4.71$, $p<.05$, für die Laborbedingung und $F(2,18),=19.423$, $p<.001$ für die häusliche Umgebung. Getrennt für jede der drei Geläufigkeiten, zeigen t-Tests für unabhängige Stichproben nur für sehr ungeläufige Aufgaben eine signifikante Wirkung der Umgebung, $t(18)=-2.59$, $p<.05$, nicht aber für geläufige und sehr geläufige Aufgaben. Während sich in der Laborumwelt die Anzahl unterschiedlicher Speicher für die drei Aufgabengruppen durch t-Tests zum Paarvergleich bei Meßwiederholung nicht differenzieren lassen, erreichen alle Paarvergleiche für

die gewohnte Umwelt das Signifikanzniveau: $t(9)=4.31$, $p<.01$ für den Vergleich von geläufigen mit sehr geläufigen Aufgaben; $t(9)=4.33$, $p<.01$ für den Vergleich von geläufigen mit ungeläufigen Aufgaben und $t(9)=9.22$, $p<.001$ für den Vergleich von sehr geläufigen mit ungeläufigen Aufgaben.

Diskussion

Zwei Fragen standen im Mittelpunkt der ersten Studie: ob sich die Modellannahmen bezüglich der Bearbeitungszeit und der Anzahl der eingesetzten Speicher bestätigen lassen und ob bzw. wie die Variation der Umgebung die vorhergesagten Effekte verändert.

Nach dem SWIEGS-Modell entscheidet die Geläufigkeit einer Repräsentationsaufgabe, durch welchen der drei Prozesse ein Speicher gewählt wird. Die konstruierten Aufgaben müssen sich also aus Sicht der Probanden hinsichtlich ihrer Geläufigkeit deutlich unterscheiden, wenn das Versuchsmaterial zur Überprüfung des Modells taugen soll. Dies wurde durch Einschätzung der Geläufigkeit der einzelnen Aufgaben überprüft und bestätigt: die Einschätzung der Geläufigkeit der Aufgabengruppe folgt den Annahmen zur Geläufigkeit bei der Konstruktion der Aufgaben, wobei die Unterschiede zwischen den drei Aufgabengruppen sowohl in der Laborumwelt als auch in der gewohnten Umgebung hochsignifikant wurden.

Das Modell nimmt drei Prozesse zur Wahl eines Speichers an, die sich im kognitiven Aufwand unterscheiden: Die Anwendung eines Repräsentationsscriptes bedarf nur der Klärung einer einzigen Frage, die Evaluation schon der Erhebung und des Vergleichs von mindestens fünf Parametern und die Konstruktion darüber hinaus zusätzlicher Suchprozesse oder Modifikationen der Prüfvoraussetzungen. Da die Geläufigkeit der Aufgabe bestimmt, welcher Prozeß abläuft, sollte die Zeit zur Bearbeitung der Aufgaben von der Geläufigkeit der Aufgabe abhängen (H1). Dies wurde bestätigt: die Bearbeitungszeit steigt in beiden Bedingungen signifikant mit sinkender Geläufigkeit der Aufgabengruppe. Auch unabhängig von der Aufgabengruppe korreliert die Bearbeitungszeit hoch mit der eingeschätzten Geläufigkeit der einzelnen Items.

Der steigende Zeitbedarf hängt nach dem Modell mit dem komplexeren Prozeß der Speicherwahl zusammen. Im Vergleich von sehr geläufigen Aufgaben mit geläufigen heißt das, daß an die Stelle der Anwendung eines Scriptes mit nur einem Speicher die Prüfung mehrerer unterschiedlicher Speicher tritt. Deswegen müßten geläufige und ungeläufige Aufgaben mit mehr unterschiedlichen Speichern gelöst werden als sehr geläufige Aufgaben (H6). Genau das ist auch der Fall: die Aufgabengruppe mit sehr geläufigen Aufgaben wird mit deutlich weniger unterschiedlichen Speichern gelöst als die der beiden weniger geläufigen Gruppen.

Betrachtet man die Ergebnisse zu H1 und H6 gemeinsam, taucht die Frage auf, ob diese Ergebnisse nicht mit einem viel einfacheren Modell zu erklären sind: Seltene Aufgaben brauchen seltene Speicher zur Repräsentation; es dauert aber einfach länger, an seltenere Speicher zu denken, als an öfter gebrauchte. Denn der Unterschied in der Einsatzhäufigkeit von Speichern ist zumindest bei den hier gestellten Aufgaben kraß: Das Gedächtnis und der Terminkalender werden fast genauso oft genannt wie die übrigen 39 genannten Speicher zusammen. Faßt man alle Nennungen von Notizzetteln zusammen, lösen gar drei Speicher fast 2/3 der Aufgabe und 38 weitere das restliche Drittel. Das SWIEGS-Modell nimmt aber an, daß generell Entscheidung für häufig eingesetzte Speicher schneller ablaufen als für seltener eingesetzten Speicher (H3). Denkbar wäre folglich, daß sehr geläufige Aufgaben mit den häufigen Speichern, geläufige mit regelmäßig eingesetzten Speichern und ungeläufige mit seltenen Speicher gelöst werden. Der Zeitunterschied käme folglich nur durch das Ergebnis - die Häufigkeit des gewählten Speichers - und nicht durch unterschiedliche Prozesse der Speicherwahl zustande. Diese Kritik ist widerlegt, wenn man nachweisen kann, daß die Planungszeit unabhängig vom letztendlich gewählten Speicher über die Geläufigkeit der Aufgabengruppe steigt (H2). Die vorliegenden Daten reichen zur inferenzstatistischen Überprüfung dieser Hypothese allerdings nicht aus; hierzu muß die Anzahl der Probanden und Aufgaben erhöht werden.

Die beschriebenen Effekte sind in der Laborsituation signifikant schwächer als in der alltäglichen Umgebung, bleiben aber bestehen. Dies zeigt sich schon bei der Einschätzung der Aufgabengeläufigkeit: die gleichen Aufgaben werden in der Alltagsumgebung extremer eingeschätzt als in der Laborumgebung, also die sehr geläufigen Aufgaben geläufiger und die ungeläufigen Aufgaben ungeläufiger. Das setzt sich auch bei der Planungszeit und bei der Speicherzahl fort. Mit sinkender Geläufigkeit steigt die Planungszeit in der alltäglichen Umgebung stärker als in der Laborumgebung; die Anzahl eingesetzter Speicher steigt in der alltäglichen Umgebung von sehr geläufigen zu geläufigen Aufgaben stärker als in der Laborumgebung. Trotzdem bleiben beide Effekte auch für die Laborsituation alleine betrachtet signifikant.

Differenzierter sieht der Effekt der Umwelt für den Vergleich zwischen geläufigen und ungeläufigen Aufgaben aus. Während nämlich im Labor die Anzahl unterschiedlicher Speicher von geläufigen zu ungeläufigen Aufgaben kaum noch steigt, steigt sie signifikant in der häuslichen Umgebung. Nach dem SWIEGS-Modell gibt es für ungeläufige Aufgaben zwei Möglichkeiten der Konstruktion: entweder werden neue Speicher, oder, wenn das mißlingt, neue Speicherbedingungen konstruiert, d.h. die Prüfparameter werden verändert und die bekannten Speicher werden erneut geprüft. Deswegen ist es nicht möglich, aus dem SWIEGS-Modell vorherzusagen, ob bei ungeläufigen Aufgaben mehr unterschiedliche Speicher als bei den geläufigen Aufgaben eingesetzt werden oder nicht.

Die Ergebnisse geben aber Anlaß zu einer interessanten Überlegung. Da die Bearbeitungszeit von den geläufigen zu den ungeläufigen Aufgaben in beiden Umgebungen deutlich steigt, würde man davon ausgehen, daß in beiden Umgebungen eine Konstruktion stattfindet. Innerhalb des SWIEGS-Modells diskutiert lassen sich die unterschiedlichen Ergebnisse in Bezug auf die Speicherzahl so interpretieren, daß in der Laborumgebung eher die Prüfparameter verändert und bekannte Speicher noch einmal getestet werden; in der häuslichen Umgebung dagegen eher neue Speicher konstruiert werden. Dies wäre als Wirkung der reizarmen Umgebung des psychologischen Laboratoriums im Vergleich zur anregenden häuslichen Umgebung unmittelbar plausibel. Entsprechend dem 'time-monitoring' (Ceci & Bronfenbrenner, 1985; vgl. Einleitung zu Studie 1) könnte die Konstruktion neuer Speicher also tatsächlich ein Effekt sein, der nur in der alltäglichen Umgebung auftritt, in der Laborumgebung aber verhindert wird. Hier könnten weitere Untersuchungen ansetzen; da das Ergebnis aber für das Modell aber nicht kritisch ist, wird dieser Gedanke in der vorliegenden Arbeit nicht weiter verfolgt.

Die Variation der Umgebung liefert aber auch eine weitere Stützung für das SWIEGS-Modell: Sowohl bei der Planungszeit als auch bei der Anzahl unterschiedlicher Speicher unterscheiden sich die Ergebnisse - im Vergleich über beide Umwelten - für sehr geläufige Aufgaben nicht voneinander, wohl aber die für geläufige und ungeläufige Aufgaben. So würde auch der Einwand entkräftet, es handle sich bei der Feldbedingung einfach um eine unspezifische Motivationserhöhung. Tatsächlich sind die Umgebungseffekte spezifisch für geläufige und ungeläufige Aufgaben. Bei sehr geläufigen Aufgaben reicht nach dem Modell die Identifikation des Aufgabentyps aus, um ein passendes Repräsentationsscript und damit auch unmittelbar einen Speicher zuzuordnen. Dieser Vorgang läuft weitgehend automatisch - ein Einfluß der Umgebung erscheint unwahrscheinlich. Bei geläufigen und ungeläufigen Aufgaben dagegen geht es um die Evaluation oder gar Konstruktion von Speichern bzw. Speicherbedingungen. Diese Prozesse laufen kontrollierter - die Umgebung kann bei beiden Prozessen anregend wirken.

Vier Schwachpunkte weist die Diskussion bis hierhin auf. Weder der Einfluß der Lesezeit noch der der Valenz der Aufgaben auf die Bearbeitungszeit ist geklärt. Es erwies sich bei der Itemkonstruktion als schwer möglich, die einzelnen Aufgaben im Hinblick auf Wortzahl oder Propositionenzahl zu standardisieren. Beides hat aber einen Einfluß auf die Lesezeit (vgl. Kintsch & Keenan, 1973). Da ungewöhnliche Aufgaben aber auch komplexer formuliert sind, könnte sich die Bearbeitungszeit als ein Artefakt der Lesezeit herausstellen. Gleiches gilt für die Valenz der einzelnen Aufgaben: eine Standardisierung erscheint hier prinzipiell ausgeschlossen. Zusammenhänge zwischen Valenz und Planungszeit nimmt aber auch das SWIEGS-Modell an: So soll die Planungszeit mit der Valenz der Aufgaben steigen (H7). Sollten die Probanden geläufige und ungeläufige Aufgaben für

subjektiv wichtiger halten als sehr geläufige Aufgaben, könnte nicht mehr entschieden werden, ob der Effekt der Planungszeit ein Effekt der Geläufigkeit oder aber ein Effekt der Valenz ist. Die Sensibilität der Planungszeit für die Umgebung, in der die Untersuchung durchgeführt wird, nimmt beiden Einwänden allerdings die Schlagkraft. Es gibt keinen Grund, warum die Lesezeit für gewöhnliche und ungewöhnliche Aufgaben in der gewohnten Umgebung länger sein sollte als in der Laborumgebung, während sie für sehr geläufige Aufgaben in beiden Umgebungen gleich lang ist. Gleiches gilt, wenn man die Variation der Planungszeit als einen Effekt der Valenz der Aufgaben erklären will: Auch hier ist der Einfluß der Umgebung auf die Planungszeit nicht plausibel. Trotzdem muß die Rolle der Lesezeit abgeklärt werden.

Ein weiteres Problem ist, daß die Zeitmessung nicht verdeckt erfolgen konnte. Zwar wurde in der Instruktion der Probanden betont, daß es nicht darum geht, die Aufgabe möglichst schnell zu lösen und im postexperimentellen Interview fand sich kein Hinweis auf einen Zeitdruck. Trotzdem kann nicht ausgeschlossen werden, daß der sichtbare Einsatz der Stoppuhr die Probanden unter Druck gesetzt hat und die Aussagen des postexperimentellen Interviews auf 'compliance' (Levy & Loftus, 1994) zurückzuführen sind. Die Probanden würden sich dann daran erinnern, daß sie die Stoppuhr eigentlich nicht unter Druck setzen sollte, und nichts darüber sagen, ob tatsächlich ein Zeitdruck vorhanden war oder nicht.

Die Messung der Geläufigkeit erfolgte in Studie 1 unmittelbar, nachdem die Probanden die Aufgabe gelöst hatten. Das könnte aber den angenommenen Zusammenhang zwischen Planungszeit und Geläufigkeit überhaupt erst induzieren: die Probanden beantworten nicht mehr, ob die Aufgabe, verglichen mit ihrer Alltagserfahrung, geläufig ist, sondern, wie einfach oder schwer es war, einen passenden Speicher zu finden. Diese Kritik kann die zentrale Interpretation der Ergebnisse nicht grundsätzlich erschüttern: die drei Geläufigkeitsgruppen der Aufgaben wurden a priori und theoretisch geleitet konstruiert. Fraglich ist aber die Unabhängigkeit des treatment-check - die Einschätzung der Aufgaben - von der Hypothesenprüfung - der Messung der Planungszeit. Die Geläufigkeit der Aufgaben müßte getrennt von der Bearbeitung eingeschätzt werden, um sicherzustellen, daß Geläufigkeit von den Probanden nicht unmittelbar als 'habe leicht einen passenden Speicher gefunden' interpretiert wird. Theoretisch wird dieser Zusammenhang zwar angenommen; empirisch den Probanden in Studie 1 aber geradezu in den Mund gelegt.

Weitere Studien zur Klärung dieser Fragen können durchaus im Labor durchgeführt werden: die Effekte im Feld unterscheiden sich - mit Ausnahme des hypothesenunkritischen Effektes auf die Speicherzahl - nicht von denen im Labor. Allerdings muß im Labor mit einer Verminderung der

Effektstärke gerechnet werden. Untersuchungen im Labor mit diesem experimentellen Ansatz erfordern also eine größere Versuchspersonenzahl.

2.1.2 Studie 2: Lesezeit, Itemkonstruktion und Beurteilung der Aufgaben

Studie 2 repliziert Studie 1, wobei neben der genaueren Messung und der größeren Datenbasis fünf wesentliche Präzisierungen hinzukommen.

- Erstens wird die Lesezeit für die Aufgaben erhoben, so daß die Planungszeit um die Lesezeit korrigiert werden kann;
- zweitens erfolgt Zeitmessung verdeckt, um einen unspezifischen Effekt durch Zeitdruck auszuschließen werden;
- drittens wird die Valenz der Items erhoben worden, um auszuschließen, daß die Valenz der Aufgaben für die Planungszeiterhöhung verantwortlich ist;
- viertens wird die Erhebung der Geläufigkeit aus dem Kontext der Speicherwahl herausgenommen, um einen von der Aufgabenlösung unabhängigen treatment-check zu gewährleisten;
- fünftens wird die Itemkonstruktion modifiziert. Studie 1 ergab, daß für sehr geläufige Aufgaben weniger unterschiedliche Speicher eingesetzt werden als für geläufige Aufgaben (H6). Das SWIEGS-Modell sagt diesen Effekt mit der Begründung vorher, daß zur Lösung sehr geläufiger Aufgaben Repräsentationsscripte eingesetzt wurden. Die einzelnen Aufgaben werden also zu einem Aufgabentyp zusammengefaßt, der auf eine durch das Script bestimmte Art behandelt wird. Dieser Effekt sollte nicht vom Typ der Aufgabe abhängen, sondern ausschließlich von der Geläufigkeit der Aufgabe. Um diese Forderung empirisch zu überprüfen, ist es notwendig, Aufgabentyp und Geläufigkeit unabhängig voneinander zu variieren. Das Versuchsmaterial wurde deswegen für Studie 2 neu konstruiert, wobei sich die Aufgaben in fünf thematische Gruppen gliedern. Innerhalb jeder dieser Gruppen ist jede Geläufigkeitsstufe vertreten.

Die Hypothesen in Studie 2 entsprechen denen von Studie 1 bezüglich Planungszeit (H1) und Speicherzahl (H6). Beide Hypothesen konnten in Studie 1 bestätigt werden und sollen in Studie 2 unter Ausschaltung von Alternativerklärungen repliziert werden. Die Erhöhung der Versuchspersonenzahl und der Aufgabenzahl soll die Prüfung weiterer Hypothesen zur Planungszeit ermöglichen. Die Erhöhung der Planungszeit bei sinkender Geläufigkeit der Aufgaben soll unabhängig von dem gewählten Speicher sein (H2). Werden bei Aufgaben gleicher Geläufigkeit unterschiedliche Speicher gewählt, so sollte die Planungszeit für den seltener gewählten Speicher länger sein (H3). Dieser

Unterschied in der Planungszeit sollte um so geringer sein, je geläufiger die bearbeitete Aufgabe ist (H4).

Nach den Ergebnissen von Studie 1 ist es möglich, das Forschungsparadigma auch in seiner weniger aufwendigen und kontrollierbareren Form - in der Laborsituation - einzusetzen. Hinweise für spezifische Umgebungseffekte betrafen keine der zu prüfenden Hypothesen. Allerdings waren alle Effekte in der Laborsituation schwächer als in der alltäglichen Umgebung; deswegen muß die Versuchspersonenzahl für Studie 2 erhöht werden.

Material und Versuchsdurchführung

Fünf Itemgruppen mit unterschiedlichem thematischen Schwerpunkt wurden konstruiert: Termine, zu erledigen, wiederkehrende Aufgaben, Informationsspeicher und einkaufen. Diese Titel sind als Schwerpunkte zu verstehen; trennscharf sind die Gruppen nicht. Termine hatten als Schwerpunkt einen zukünftigen Zeitpunkt, zu dem etwas erinnert werden mußte, z.B. 'Ein Termin beim Friseur am kommenden Donnerstag um 11.00'. Wiederkehrende Aufgaben waren solche, die in einem bestimmten Turnus immer wieder durchzuführen waren, z.B. 'Täglich Blumen gießen'. Aufgaben des Typs zu erledigen betrafen Gegenstände, mit denen zu einem bestimmten Zeitpunkt eine Handlung durchgeführt werden mußte, z.B. 'Morgen eine Überweisung zur Bank bringen'. Einkaufen ist eindeutig, z.B. 'Morgen frische Oliven kaufen'. Informationsspeicher schließlich betraf die Frage, wie bestimmte Informationen gespeichert werden, ohne daß direkt eine Handlung intendiert ist, z.B. 'Die Telefonnummer eines Bekannten'. Zu jedem Thema wurden vier Aufgaben konstruiert. Wie in Studie 1 wurden nun bei jeder Aufgabe geläufige Elemente in zwei Stufen durch ungeläufige ersetzt, so daß jede Aufgabe in drei Geläufigkeitsstufen vorlag (vgl. Studie 1). Insgesamt bestand das Itempool so aus 60 Aufgaben, die vollständig im Anhang 5.1.2 aufgeführt sind.

Die Präsentation der Items erfolgte auf einem Computerbildschirm. Er zeigte eine aufgedeckte Karteikarte, in deren Zentrum die jeweilige Aufgabe stand. Die Kontrollgruppe erhielt die Aufgabe, das jeweilige Item zu lesen und danach auf seine Geläufigkeit und Valenz hin einzuschätzen. Durch diesen Kontext sollte sichergestellt werden, daß die Probanden die Aufgaben tatsächlich vollständig lasen und verstanden. Die Experimentalgruppe sollte die Aufgaben lösen, indem sie einen passenden Speicher nannte.

Dazu wurden die Items den Probanden in zufälliger Reihenfolge präsentiert. Die Kontrollgruppe erhielt die Aufforderung, ein Item zunächst zu lesen und zu verstehen. Danach konnten die Probanden durch die Leertaste eine fünfstufige Skala für die Einschätzung der Geläufigkeit des Items öffnen. Durch diese Trennung von Lese- und Ratingphase sollten die Probanden dazu angehalten werden,

Aufgaben vor der Beurteilung wirklich vollständig zu lesen. Zudem konnte so die Zeit zum Lesen eines Items erhoben, ohne daß die Probanden den Eindruck einer Zeitmessung hatten. Nach der Einschätzung der Geläufigkeit erschien automatisch eine zweite fünfstufige Skala für die Valenz der Items auf dem Monitor. Die Probanden sollten hier einschätzen, wie wichtig diese Aufgabe für sie wäre, wenn ihnen diese Aufgaben in ihrem Alltag begegnen würde. Nach dieser Einschätzung folgte automatisch das nächste Item.

Die Experimentalgruppe erhielt wie in Studie 1 den Auftrag, einen Speicher zu den Aufgaben zu nennen, den Sie auch in Ihrem Alltag benutzen würden. Zur Realisierung der verdeckten Zeitmessung war an den Computer ein Mikrofon angeschlossen; auf dem Bildschirm war das Symbol für ein Tonbandgerät zu sehen. Die Probanden erhielten nun den Auftrag, ihre Antwort in das Mikrofon zu sprechen. Sie wurden instruiert, die Aufnahme durch den Druck einer beliebigen Taste zu starten. Nach dem Tastendruck begannen sich die Tonbandspulen auf dem Monitor zu drehen. Durch eine bestimmte Taste konnte die 'Aufnahme' wieder beendet werden; das Tonbandgerät auf dem Monitor stoppte und das nächste Item erschien. Durch dieses Verfahren konnte die Planungszeit der Probanden als Zeit von der Präsentation des Items bis zum Tastendruck verdeckt gemessen werden; tatsächlich erfolgte keinerlei Aufnahme. Der Versuchsleiter notierte die Antwort der Probanden auf einem vorbereiteten Bogen, was mit auswertungstechnischen Gründen erklärt wurde.

Experimental- und Kontrollgruppe erhielten drei Übungsitems, um sich mit der Aufgabenstellung vertraut zu machen.

Versuchspersonen und Design

Vierzig Studenten und Studentinnen der Psychologie nahmen an dem Versuch teil; ihre Aufteilung in Experimental- und Kontrollgruppe erfolgte per Zufall. Die Altersspanne reichte von 19 bis 31 Jahren; 24 Probanden waren Frauen, 16 Männer. Für die Teilnahme erhielten die Probanden einen Versuchspersonenschein.

Unabhängige Variablen sind die Aufgabenstellung mit zwei Ausprägungen: Items lesen und beurteilen in der Kontrollgruppe bzw. passende Speicher nennen in der Experimentalgruppe sowie die drei Geläufigkeitsgruppen der Items, die als Meßwiederholungsvariable konzipiert ist. Das Design ist also 2*3 gemischt faktoriell. Abhängige Variablen sind wieder die Zeiten sowie die Einschätzungen und Lösungen der Aufgaben.

Ergebnisse

Die Ergebnisse werden im folgenden zunächst für die Einschätzung der Items, dann für die Verarbeitungszeit und schließlich für die eingesetzten Speicher präsentiert.

Die drei Geläufigkeitsgrade differenzieren die Items sowohl in Bezug auf deren geschätzte Geläufigkeit als auch deren geschätzte Valenz. Beide Größen nehmen über die drei Geläufigkeitsgruppen ab (siehe Abb. 7).

Varianzanalysen mit dem Meßwiederholungsfaktor Geläufigkeit zeigen einen hochsignifikanten Einfluß der Geläufigkeitsgruppe sowohl für die geschätzte Geläufigkeit, $F(2,38)=182.14$, $p<.001$, als auch für die geschätzte Valenz, $F(2,38)=68.80$, $p<.001$ mit einem linearen Trend: $F(1,19)=241.60$,

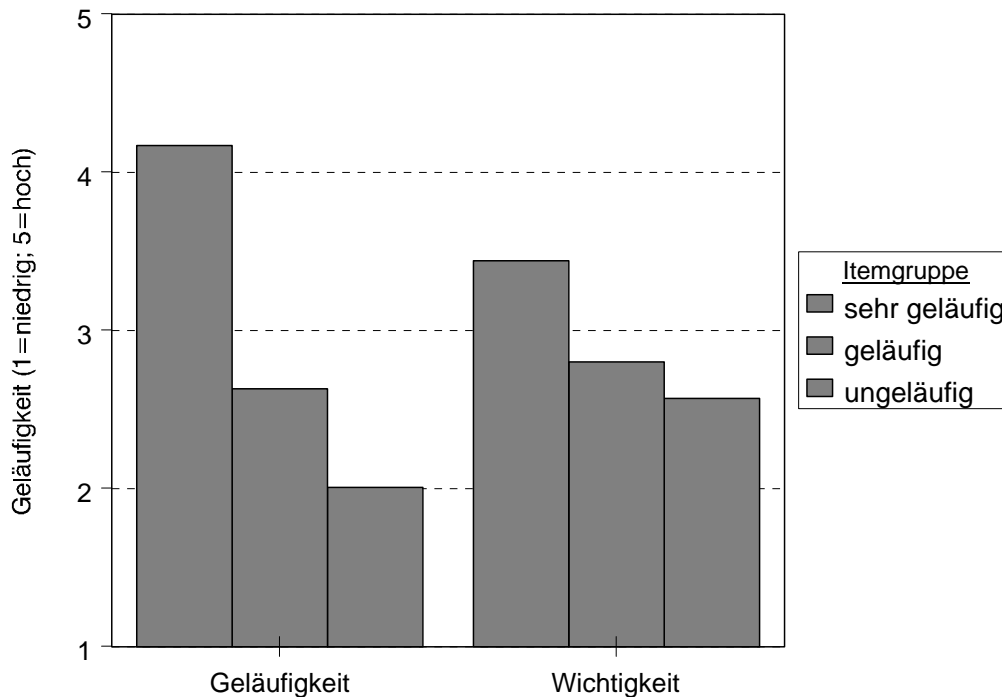


Abb. 7 Die Einschätzung der drei Itemgruppen nach Geläufigkeit und Valenz der einzelnen Items.

$p<.001$, für die Geläufigkeit und $F(1,19)=102.49$, $p<.001$ für die Valenz. T - Tests differenzieren die einzelnen Mittelwerte alle hochsignifikant; für alle sechs Vergleiche gilt $t(19)>4.2$, $p<.001$.

Kein Proband bemerkte die verdeckte Messung der Zeiten; auch die 'Aufnahme' in der Experimentalgruppe wurde als solche akzeptiert. Die Zeiten für die Verarbeitung der Aufgaben steigen kontinuierlich mit sinkender Geläufigkeit. Dieser Gradient ist steiler, wenn die Aufgabe gelöst werden soll, als wenn sie nur gelesen werden muß. So ist die Differenz zwischen beiden Zeiten bei sehr geläufigen Items unter einer Sekunde und steigt auf 3.5s bei ungeläufigen Items (siehe Abb. 8).

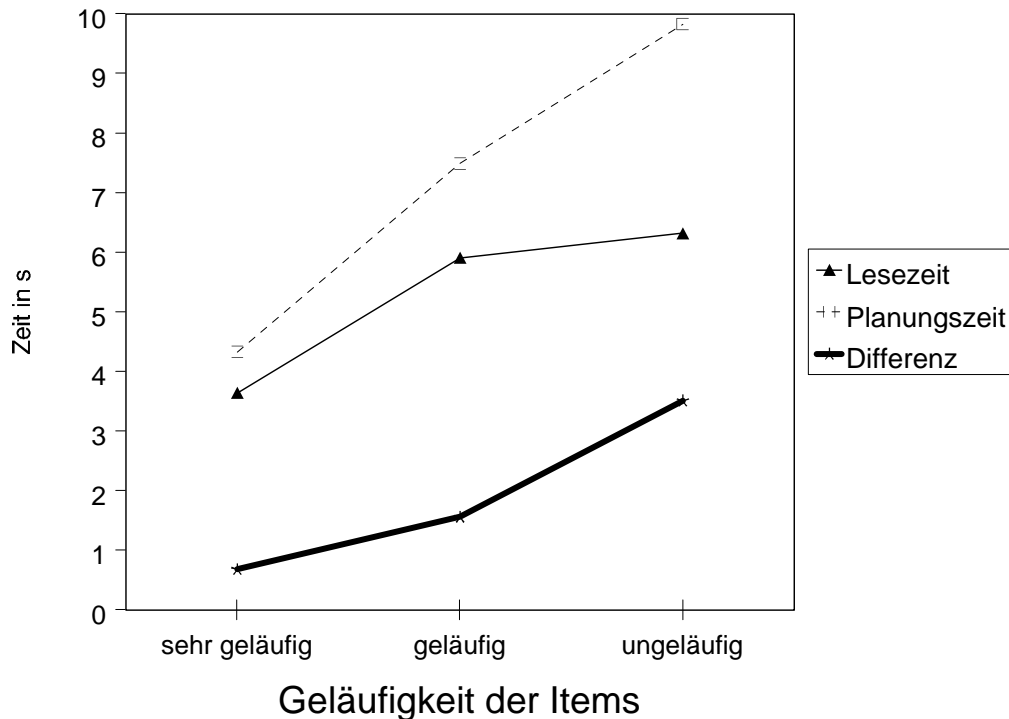


Abb. 8 Die Lesezeit, Bearbeitungszeit und Differenz zwischen beiden, getrennt nach der Geläufigkeit der bearbeiteten Aufgaben.

Die Varianzanalyse mit Aufgabenstellung als u.V. und Geläufigkeit als Meßwiederholungsvariable zeigt hoch signifikante Haupteffekte, $F(1,38)=110.04$, $p<.001$ für Aufgabenstellung und $F(2,76)=99.91$, $p<.001$ für Geläufigkeit sowie für die Interaktion beider Parameter, $F(2,76)=12.07$, $p<.001$. T - Tests für abhängige Stichproben trennen die Mittelwerte der Planungs bzw. Lesezeiten über die Geläufigkeitsstufen signifikant, bis auf die Lesezeit für Aufgaben mit mittlerem vs. Aufgaben mit geringer Geläufigkeit; hier wird das Signifikanzniveau mit $t(19)=-1.94$, $p=.067$ knapp verfehlt. Mittelwertsvergleiche für unabhängige Stichproben zum Vergleich der beiden Aufgabenstellungen ergeben eine signifikante Trennung nur für geläufige, $t(38)=-2.09$, $p<.05$, und ungeläufige $t(38)=-3.39$, $p<.01$, nicht aber für sehr geläufige Aufgaben, $t(38)=-1.53$, $p>.05$.

Die durchschnittliche Lesezeit der einzelnen Items wurde für jede Versuchsperson der Experimentalgruppe von ihrer Bearbeitungszeit des jeweiligen Items subtrahiert (vgl. Abb. 8). Eine Varianzanalyse mit dem Faktor Geläufigkeit zeigt hoch signifikante Unterschiede dieser Differenz zwischen Lesezeit und Planungszeit für die drei Geläufigkeitsstufen, $F(2,18)=14.88$, $p=.001$. T - Tests für abhängige Stichproben trennen die Differenz für alle Geläufigkeitsstufen signifikant: $t(19)=-2.12$, $p<.05$, für hohe vs. mittlere Geläufigkeit, $t(19)=-3.93$, $p=.001$, für mittlere vs. geringe Geläufigkeit und $t(19)=-4.91$, $p<.001$ für hohe vs. geringe Geläufigkeit.

Mit den in der Lesebedingung gewonnenen Kennwerten der Aufgaben wurde der korrelative Zusammenhang zwischen Bedeutung und Geläufigkeit eines Items auf der einen und seiner Verarbeitungszeit auf der anderen Seite analysiert. Planungszeit, Lesezeit und die Differenz zwischen beiden korrelieren hoch signifikant ($p < .001$) mit der Geläufigkeit der Items, $r = -.69$, $r = -.52$ und $r = -.53$. Mit der Bedeutung des Items korrelieren nur Planungszeit und die Differenz zwischen Planungszeit und Lesezeit der Items signifikant, $r = -.38$ und $r = -.43$, $p < .05$, nicht aber die Lesezeit $r = -.14$. Valenz und Geläufigkeit einer Aufgabe korrelieren ebenfalls stark positiv miteinander, $r = .7$, $p < .001$.

Zur Lösung der Repräsentationsaufgaben wurden 40 verschiedene Speicher genannt (siehe Tab. 4).

Eingesetzte Speicher

Speicher	Anzahl der Nennungen:	absolut	in %	Speicher	Anzahl der Nennungen:	absolut	in %
1. Gedächtnis		467	38,9%	21. Eselsbrücken		4	0,3%
2. Kalender		272	22,6%	22. Tagebuch		4	0,3%
3. Zettel an auffälligem Ort / Pinnwand	116	9,7%		23. Zettel verstecken		3	0,3%
4. Zettel (ohne Ortsangabe)	54	4,5%		24. Langzeitkalender (>1Jahr)		3	0,3%
5. Adreßbuch	49	4,1%		25. Uniblock		2	0,2%
6. Zettel an kritischem Objekt / Ort	36	3%		26. Telefonbuch		2	0,2%
7. Gegenstand plazieren	30	2,5%		27. anstreichen, plazieren		2	0,2%
8. Kalender, übertragen	19	1,6%		28. codierte im Telefonbuch		2	0,2%
9. aufschreiben (ohne Ortsangabe)	16	1,3%		29. Schreibtischunterlage		2	0,2%
10. Mitschrift	12	1%		30. in Holzmöbel einritzen		2	0,2%
11. andere Person	13	1,1%		31. Strichliste		2	0,2%
12. Terminkalender vorne/hinten/Notizteil	13	1,1%		32. Zettel in Ordner ablegen		2	0,2%
13. teilweise sofort ausführen	10	0,8%		33. mit rot in Notizteil Kalender		2	0,2%
14. Einkaufszettel	10	0,8%		34. Wecker		1	0,1%
15. Geburtstagskalender	10	0,8%		35. anstreichen und aufheben		1	0,1%
16. Notizbuch	9	0,8%		36. codiert notieren		1	0,1%
17. modifizierter Kalender	7	0,6%		37. Fach für geheime Dinge		1	0,1%
18. modifiziertes Adreßbuch	6	0,5%		38. anstreichen		1	0,1%
19. Ordner	5	0,4%		39. Infomaterial		1	0,1%
20. Zettel im Schrank	6	0,5%		40. Radio		1	0,1%
				Keine Angaben		14	1,2%
				<hr/>			
				Gesamt		1200	100%

Tab. 4 Zur Lösung der Repräsentationsaufgaben eingesetzte Speicher; Prozentangaben auf eine Dezimalstelle hinter dem Komma gerundet. Speicher Nr.3,6,7 und 13 sind zusammengefaßt, die übrigen unverändert (vgl. Studie 1).

Dabei wurden die Angaben der Probanden entsprechend der Auswertung bei Studie 1 nach Möglichkeit wörtlich übernommen. Bei den Speichern Nr. 3, 6, 7 und 13 wurden Synonyme oder offensichtlich gleiche Funktionen der Speicher zusammengefaßt. Bei Speicher Nr. 3, Notizzettel an auffälligem Ort, wurden die unterschiedlichsten Orte genannt, von der Haustür über den Kühlschrank

bis zum Klo. Ähnlich verhielt es sich bei Speicher Nr. 6, Notizzettel an kritischem Objekt. Hier wurde der Notizzettel direkt an dem Objekt befestigt, mit dem eine Aufgabe in der Zukunft verbunden ist, beispielsweise die Hausapotheke, die in zwei Jahren zu überprüfen ist, oder der Wasserfilter, der in einem halben Jahr ausgetauscht werden muß. Bei beiden Gruppen wurde der Einzelfall also nicht noch einmal getrennt aufgeführt, sondern über seine Funktion klassifiziert. Ähnlich bei Speicher Nr. 7, 'Objekt plazieren'. Hier wird das kritische Objekt, mit dem etwas gemacht werden soll, so plaziert, das es auffällt und damit an die Aufgabe erinnert. Ein Brief, der zur Post mitgenommen werden muß und deswegen vor der Ausgangstür liegt, ein Buch, daß zurückgegeben werden muß und deswegen auf dem Frühstückstisch landet oder eine Gießkanne, die mitten im Raum steht und so das Blumengießen anmahnt, gehören hierzu. Schließlich betrifft Speicher Nr. 13, teilweise sofort ausführen, alle Speicherarten, wo ein Teil der geforderten Handlung sofort ausgeführt wird. Beispielsweise einen Brief, der zur Post gebracht werden muß, sofort in die Jackentasche zu stecken oder ein Script, was zur Uni zu bringen ist, sofort in die Aktenmappe einzupacken.

Die verschiedenen Speicher werden sehr unterschiedlich häufig eingesetzt. Fast 72% der Aufgaben werden über drei Speicher gelöst, 24% über weiter 13 Speicher und die verbleibenden 4% der Aufgaben mit 23 Speichern. Für jede der drei Geläufigkeiten der Aufgaben wurde berechnet, wie viele unterschiedliche Speicher die Probanden pro Aufgabengruppe (Termine, zu erledigen, einkaufen usw.) eingesetzt haben. Maximal wären vier unterschiedliche Speicher möglich - wenn ein Proband für jede Aufgabe eines bestimmten Typs und einer bestimmten Geläufigkeitsstufe einen anderen Speicher gewählt hätte. Tatsächlich wurden bei sehr geläufigen Aufgaben durchschnittlich 1.99, bei geläufigen 2.7 und bei ungeläufigen 2.65 unterschiedliche Speicher eingesetzt. Eine Anova mit der Geläufigkeit als Meßwiederholungsvariable zeigt die hohe Signifikanz dieses Effektes, $F(2,38)=35.58$, $p<.001$. T-Tests für Meßwiederholungsvariablen belegen, daß dieser Effekt auf der unterschiedlichen Speicherzahl zwischen sehr geläufigen und geläufigen, $t(19)=-6.38$, $p<.001$, bzw. sehr geläufigen und ungeläufigen Aufgaben, $t(19)=-6.124$, $p<.001$, nicht aber durch die zwischen geläufigen und ungeläufigen Aufgaben zustande kommt, $t(19)=1.22$, $p>.05$.

Um den Zusammenhang zwischen den tatsächlich gewählten Speichern und der Planungszeit für den Einsatz dieser Speicher zu analysieren, wurden zwei komplementäre Wege eingeschlagen. Zum einen wurden die Planungszeit für die nach der Häufigkeit der gewählten Speicher gruppierten Aufgaben berechnet; zum anderen wurde die Planungszeit für die drei am häufigsten eingesetzten Speicher berechnet. Beide Wege kranken daran, daß nur die Probanden in die Varianzanalyse eingehen können, die die jeweiligen Speicher bzw. Speichergruppen tatsächlich in allen drei Geläufigkeitsstufen von Aufgaben eingesetzt haben (d.h. alle neun Zellen müssen besetzt sein). Diese Voraussetzung führt

in beiden Fällen dazu, daß nur ein Teil der Probanden in die statistische Analyse einbezogen werden kann.

Um den Zusammenhang zwischen der Häufigkeit, mit der ein Speicher eingesetzt wird, und der Planungszeit für den Einsatz diese Speichers zu analysieren, wurden die Speicher in drei Gruppen unterteilt. Die Gruppe der häufig eingesetzten Speicher betrifft die Speicher, die bei den 1200 Nennungen über 100 mal genannt wurden (das Gedächtnis, der Terminkalender und der Zettel am auffälligem Ort); die Gruppe der regelmäßig eingesetzten Speicher betraf diejenigen, die über 10 mal und unter 100 mal eingesetzt wurden (das betraf 14 Speicher); die Gruppe der selten eingesetzten Speicher bildeten die Speicher, die unter 10 mal eingesetzt wurden (das betraf 25 Speicher, vgl. Tab. 4). Für diese drei Gruppen wurde jeweils die durchschnittliche Planungszeit für ihren Einsatz getrennt nach der Geläufigkeit der Aufgabe berechnet (siehe Abb. 9).

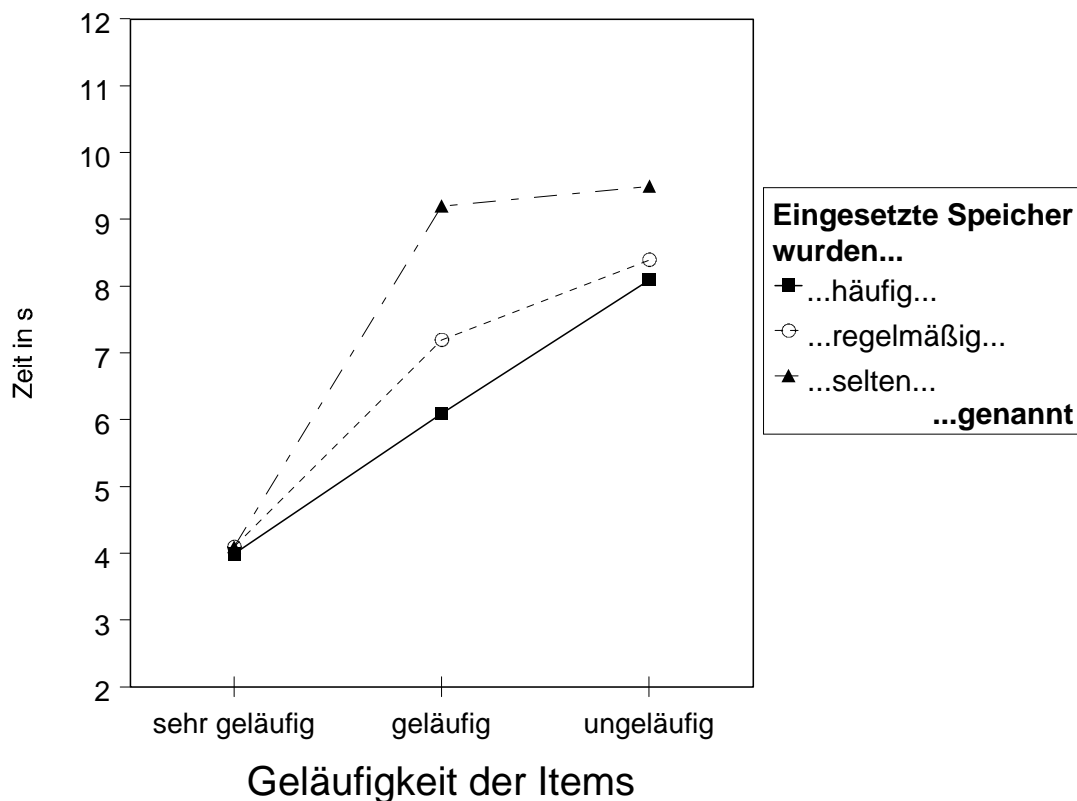


Abb. 9 Die Zeit, bis ein Speicher zur Lösung der Aufgabe genannt wird, differenziert nach der Aufgabengeläufigkeit und der Häufigkeit, mit der der jeweilige Speicher insgesamt genannt wurde. Die Mittelwerte beziehen sich auf die 16 Probanden, die in die Varianzanalyse aufgenommen wurden.

Die Planungszeit steigt für alle Speichergruppen über sinkende Geläufigkeit der Aufgaben. Dabei ist die Planungszeit über alle Aufgaben um so höher, je seltener die jeweiligen Speicher eingesetzt wurden. In die Varianzanalyse mit den beiden Meßwiederholungsvariablen Geläufigkeit der Aufgabe und Häufigkeitsgruppe des eingesetzten Speichers gehen nur 14 Probanden ein, da bei sechs Probanden

nicht alle Zellen besetzt sind. Trotzdem sind beide Haupteffekt signifikant, $F(2,26)=3,75$, $p<.05$ für Speichergruppen und $F(2,26)=24,28$, $p<.001$ für die Geläufigkeit. Die Interaktion wird nicht signifikant, $F(4,52)=1.49$, $p>.05$. T-Tests für den Mittelwertsvergleich abhängiger Stichproben ergeben hochsignifikante Unterschiede für den Paarvergleich aller drei Geläufigkeitsstufen, $t(13)=-6.61$, $p<.001$ sehr geläufige vs. geläufige, $t(13)=-5.44$, $p<.001$ für sehr geläufige vs. ungeläufige und $t(16)=-2.51$, $p<.05$ geläufige vs. ungeläufige Aufgaben. Bei den Speichergruppen erreicht in der Planungszeit nur der Unterschied in der Planungszeit zwischen häufig und selten eingesetzten Speichern das Signifikanzniveau, $t(13)=-2,71$, $p<.05$, nicht aber der Vergleich dieser beider Gruppen mit der mittleren Häufigkeit.

Die Interaktion zwischen Geläufigkeit und Speichergruppe ist den Mittelwerten angedeutet (vgl. Abb. 9), erreicht aber das Signifikanzniveau nicht. Dafür zeichnet auch der hohe Ausfall an Probanden verantwortlich. Deswegen werden trotz der fehlenden Interaktion die t-Tests zum Paarvergleich der Speichergruppen bei den drei Geläufigkeitsstufen gerechnet. Bei sehr geläufigen Aufgaben unterscheiden sich die Mittelwerte nicht, bei den beiden ungeläufigeren Aufgaben aber werden zumindest die Unterschiede zwischen der häufigen und der seltenen sowie zwischen der häufigen und regelmäßigen Speichergruppe signifikant. Der Unterschied zwischen der regelmäßigen und seltenen Speichergruppe verpaßt bei geläufigen Aufgaben mit $t(16)=-1.7$, $p=.1$, bei ungeläufigen Aufgaben noch deutlicher das Signifikanzniveau (vgl. Tab. 5).

Planungszeiten differenziert für häufig, regelmäßig und selten verwandte Speicher(Mittelwertsvergleiche über t-Tests)

<i>Speichergruppe:</i> Geläufigkeitsgruppe:	<i>häufig vs.</i> <i>regelmäßig</i>	<i>häufig vs.</i> <i>selten</i>	<i>regelmäßig vs.</i> <i>selten</i>
sehr geläufig	$t(19)=-0.3$	$t(19)=0.4$	$t(19)=0.1$
geläufig	$t(17)=5.1***$	$t(17)=5.2***$	$t(16)=-1.7$
ungeläufig	$t(14)=4.3**$	$t(13)=-3.5*$	$t(17)=-0.4$

Tab. 5 Ergebnisse der t-Tests der Paarvergleiche für die Planungszeiten über die Speichergruppen getrennt nach den drei Geläufigkeitsstufen. * für $p<.05$, **für $p<.01$, ***für $p<.001$

Die drei am häufigsten eingesetzte Speicher waren das Gedächtnis (467 von 1200 Entscheidungen), der Terminkalender (275 von 1200 Entscheidungen) und der Notizzettel, an einem auffälligen Ort plaziert (119 von 1200 Entscheidungen). Für diese drei Speicher wurde für jede Versuchsperson jeweils nach Geläufigkeitsgruppe der Aufgaben getrennt die Planungszeit berechnet; also die Zeit, die die Probanden brauchten, um sich für den jeweiligen Speicher zu entscheiden.

Für alle drei Speicher wird die Planungszeit um so länger, je geringer die Geläufigkeit der Items. Die Planungszeiten sind für den Notizzettel für alle Geläufigkeitsstufen am längsten, während der Terminkalender nur für bekannte und weniger bekannte Aufgaben längere Planungszeit als das Gedächtnis benötigt. Bei unbekanntem Aufgaben kehrt sich dieses Verhältnis um: die Planungszeit ist hier für das Gedächtnis länger (siehe Abb. 10).

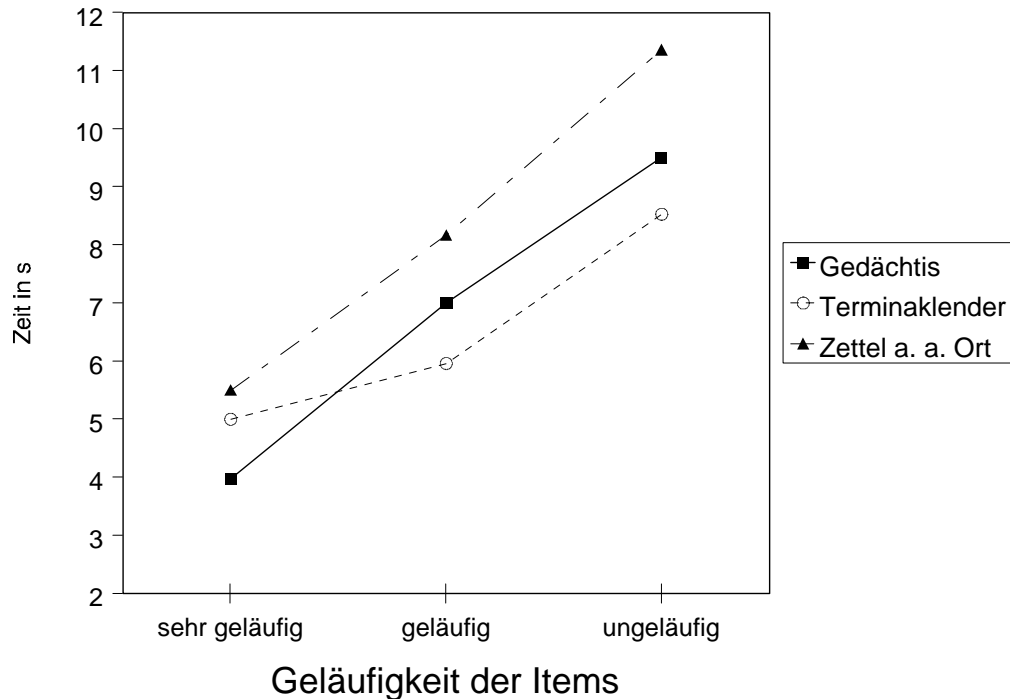


Abb. 10 Die Planungszeiten über alle Geläufigkeitsstufen getrennt für die drei meistgenannten Speicher. Die Zeiten für das Gedächtnis beziehen sich auf 19, für den Terminkalender auf 17 und für Zettel an auffälligem Ort auf 7 Probanden.

Da nur bei drei Versuchspersonen alle neun Zellen gefüllt sind, wird auf eine vollständige MANOVA verzichtet. Stattdessen werden für die drei Speicher jeweils getrennte Varianzanalysen gerechnet. In diese Varianzanalysen konnten für das Gedächtnis als Speicher 19, für den Terminkalender 17 und für den Zettel am auffälligen Ort sieben Versuchspersonen einbezogen werden.

In allen drei Fällen war die Wirkung der Geläufigkeit signifikant: $F(2,36)=20.08, p<.001$ für Gedächtnis, $F(2,32)=8.00, p<.01$ für Terminkalender und $F(2,12)=5.20, p<.05$ für Notizzettel an auffälligem Ort. Die Signifikanz der T - Tests zum Vergleich der Mittelwerte hängt hauptsächlich von der Anzahl der Probanden ab, die darin eingehen. Für das Gedächtnis unterscheiden die Geläufigkeitsstufen der Aufgaben alle Mittelwerte signifikant, für den Terminkalender und für Notizzettel an auffälligem Ort wird der Unterschied zwischen geläufigen und ungeläufigen Aufgaben nicht mehr signifikant. Vergleicht man die Planungszeit für die drei Speicher jeweils auf einer Geläufigkeitsstufe der Aufgabe, so wird bei sehr geläufigen Aufgaben nur die unterschiedliche

Planungszeit von Gedächtnis und Terminkalender signifikant, $t(16)=-3.055$, $p<.01$ Bei geläufigen Aufgaben verpaßt der Unterschied zwischen Gedächtnis und Notizzettel an auffälligem Ort mit $t(13)=-1.91$, $p=.077$ nur knapp das Signifikanzniveau, bei ungeläufigen Aufgaben unterscheiden sich die Speicher nicht mehr. Über alle Geläufigkeitsstufen hinweg erreicht kein Unterschied in der Planungszeit für die drei Speicher das Signifikanzniveau, obwohl er in den Mittelwerten vorliegt: Gedächtnis 10.1s, Terminkalender 10.2s und Zettel am auffälligen Ort 12.0s.

Um den mit der Ungeläufigkeit der Speicher und Aufgaben steigenden Datenausfall in den Griff zu bekommen, wurde die Analyse mit den drei für jede Versuchsperson individuell häufigsten Speichern wiederholt. Für jede Versuchsperson wurde dazu die Anzahl der von ihr eingesetzten Speicher ausgezählt. Die Planungszeit wurde nur für die Aufgaben berechnet, bei denen einer der drei vom Probanden am häufigsten genannten Speicher als Lösung genannt wurde. Der Datenausfall wurde aber auch durch diese aufwendigere Analyse kaum reduziert; die Ergebnisse änderten sich insgesamt im Vergleich zur Berechnung über die drei von allen Probanden am häufigsten genannten Speicher nur unwesentlich. Deswegen wird auf ihre Darstellung hier verzichtet.

Diskussion

Studie 2 sollte Alternativerklärungen für die Ergebnisse von Studie 1 ausschließen sowie einige zusätzliche Hypothesen testen.

Auch in Studie 2 folgte die Einschätzung der Geläufigkeit der Aufgaben den Zielsetzungen bei der Konstruktion der Items. Anders als in Studie 1 war die Einschätzung der Items aber nicht mit ihrer Bearbeitung verknüpft. Der hohe korrelative Zusammenhang zwischen Geläufigkeitseinschätzung und Bearbeitungszeit eines Items kann damit nicht mehr darauf zurückgeführt werden, daß die Probanden ihre eigene Bearbeitungszeit als Grundlage für die Einschätzung der Geläufigkeit nehmen.

Eine zentrale Hypothese von Studie 2 war die steigende Planungszeit für passende Speicher bei sinkender Geläufigkeit der Aufgaben. Wie in Studie 1 konnte diese Hypothese mit hoher Signifikanz bestätigt werden: je ungeläufiger eine Aufgabe, um so länger dauert es, einen passenden Speicher zu finden. Anders als in Studie 1 können Alternativerklärungen für diese Beobachtung ausgeschlossen werden. Erstens wurde die Zeit verdeckt erhoben, ohne daß dies das Ergebnis verändert hätte; ein Einfluß von Zeitdruck ist damit ausgeschlossen. Zweitens kann die subjektive Valenz einer Aufgabe als Ursache für steigende Planungszeit bei sinkender Aufgabengeläufigkeit ausgeschlossen werden. Ausgangspunkt dieser Frage war, daß höhere Valenz einer Aufgabe auch zu längeren Planungszeiten führen müßte (vgl. 1.4.2.3, H7). Würde nun Aufgabervalenz mit sinkender Geläufigkeit der Aufgabe steigen, könnte die steigende Planungszeit nicht mehr eindeutig auf die sinkende Aufgabengeläufigkeit

zurückgeführt werden. Valenz und Geläufigkeit korrelieren aber tatsächlich hoch positiv miteinander: die ungeläufigen Aufgaben wurden als weniger bedeutsam eingeschätzt als die geläufigeren. Wenn aber Aufgaben um so länger bearbeitet werden, je höher ihre Valenz ist, müßte die Valenz alleine zu sinkenden Planungszeiten bei sinkender Geläufigkeit der Aufgaben führen. Die Hypothesen zur Planungszeit sagen aber steigende Planungszeit mit sinkender Aufgabengeläufigkeit voraus (H1). Die mit der Geläufigkeit der Aufgaben sinkende Valenz bietet somit keine Alternativerklärung, sondern erschwert die empirische Verifikation der Hypothese.

Drittens erlaubt die Erhebung der Lesezeit für jedes Item eine klarere Aufteilung der Bearbeitungszeit der Items. Zwar steigt auch die Lesezeit mit der sinkenden Geläufigkeit der Items. Dieser Effekt verwundert nicht; sind doch ungewöhnliche Items häufig in der sprachlichen Darstellung komplexer. Die Unterschiede in der Planungszeit lassen sich dadurch aber nicht erklären: selbst die Differenz zwischen Planungs- und Lesezeit zeigt noch einen hochsignifikanten Effekt der Itemgeläufigkeit. Ungeläufige Items erfordern längere Lesezeit; aber auch nachdem sie gelesen worden sind, dauert es länger, einen passenden Speicher zu finden als bei geläufigeren Items.

Trotzdem bleibt fraglich, ob die Differenzbildung zwischen Lesezeit und Bearbeitungszeit eines Items sinnvoll ist. Voraussetzung dieser Analyse ist, daß sich die Bearbeitungszeit ergibt als Summe aus der Lesezeit für das Item und der Planungszeit zur Entscheidung für einen passenden Speicher. Für sehr geläufige Items erscheint das - nach dem SWIEGS-Modell - aber fraglich. Wenn ein Repräsentationscript angewandt wird, reicht nämlich die einfache Klassifikation des Items aus. Es würde also beispielsweise reichen zu erkennen, daß es sich um einen Termin handelt; die Art des Termins ist dabei unwichtig.

Eine differenziertere Analyse der Zeiten für jedes einzelne der sehr geläufigen Items zeigt tatsächlich, daß für fünf der 20 Items die Bearbeitungszeit kürzer ist als die reine Lesezeit. Es dauert also länger, das Item zu lesen und zu verstehen, als einen passenden Speicher für es zu finden (siehe Abb. 11: Item Nr. 1, 46, 55 und 4). Gleiches gilt für immerhin noch drei Items der geläufigen Gruppe und für eines der ungeläufigen. Auch wenn also die Korrektur der Bearbeitungszeit um die Lesezeit von der Versuchsplanung her als notwendig erscheint, wird sie vor dem Hintergrund des untersuchten Modells fraglich. Einen passenden Speicher für eine Aufgabe zu finden, scheint nicht etwas zu sein, was nach dem Lesen und Verstehen der Aufgabe einsetzt. Die Aufgabe scheint von vornherein auf eine andere Art gelesen zu werden, die das vollständige Verarbeiten der Aufgabe nicht notwendig macht. Diese Überlegungen sind zur Verifikation steigender Planungszeiten mit sinkender Aufgabengeläufigkeit (H1) nicht erforderlich, da auch die Differenz der Zeiten mit sinkender Geläufigkeit hochsignifikant

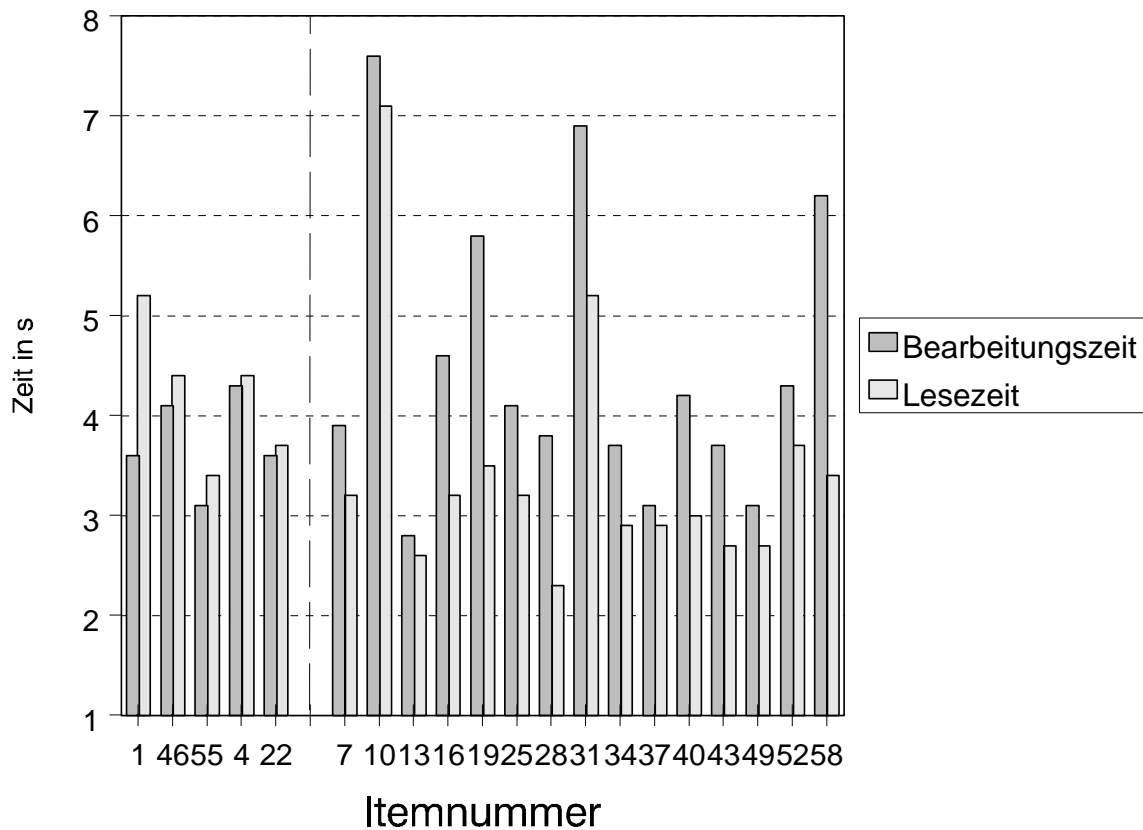


Abb. 11 Die Zeit für die Suche nach einem passenden Speicher (Bearbeitungszeit) kontrastiert mit der Zeit, das Item in Vorbereitung auf die Einschätzung des Items zu lesen (Lesezeit). Für fünf Items ist die Lesezeit länger als die Bearbeitungszeit.

steigt. Für weitere Studien ist die Differenzbildung damit aber nicht unproblematisch, da die Lesezeit einen anderen Prozeß zu betreffen scheint als den der Speicherwahl.

Die Anzahl unterschiedlicher Speicher zur Lösung eines Aufgabentyps ist für geläufige und ungeläufige Aufgaben signifikant höher als für sehr geläufige Aufgaben. Durch die systematische Variation der Aufgabentypen über die Geläufigkeitsstufen hinweg kann ausgeschlossen werden, daß Unterschiedlichkeit der Aufgabentypen für diesen Effekt verantwortlich waren. Nach den Überlegungen in der Diskussion zu Studie 1 war zu erwarten, daß im Labor kein Unterschied in der Zahl der Speicher zwischen geläufigen und ungeläufigen Aufgaben zu finden sein würde. Das war auch der Fall: die Laborumgebung scheint eher zur Modifikation der kritischen Parameter als zu einer kreativen Lösung mit neuen Speichern anzuregen.

Drei zusätzliche Hypothesen zum Zusammenhang zwischen Planungszeit und Speicherwahl konnten in Studie 2 geprüft werden. Erstens soll die Planungszeit unabhängig vom tatsächlich gewählten Speicher mit sinkender Geläufigkeit steigen (H2). Zur Testung wurden die Planungszeiten noch einmal danach differenziert, welcher Speicher gewählt worden war. In der ersten Analyse wurden

alle Speicher nach ihrer Häufigkeit in drei Gruppen unterteilt. In der zweiten Analyse wurden nur die Aufgaben aufgenommen, in denen einer der drei am häufigsten genannten Speicher gewählt worden war. Die Speichergruppe bzw. die Speicher wurden als zusätzlicher Faktor in die Analyse der Planungszeit aufgenommen. Beide Analysen belegen die Unabhängigkeit der mit sinkender Aufgabengeläufigkeit steigenden Planungszeit von der Speicherwahl. Dieser Effekt war für alle sechs Teilgruppen hochsignifikant. Die Erhöhung der Planungszeit bei sinkender Aufgabengeläufigkeit kann also nicht dadurch zustande kommen, daß ungeläufige Aufgaben mit ungeläufigen Speichern gelöst werden. Es muß tatsächlich die Unterschiedlichkeit im Prozeß der Speicherwahl verantwortlich sein, der unterschiedlich viel Zeit erfordert, auch, wenn schlußendlich der gleiche Speicher gewählt wird.

Zweitens sollte die Planungszeit bei gleicher Aufgabengeläufigkeit um so länger sein, je seltener ein Speicher insgesamt eingesetzt wurde (H3) Dieser Effekt sollte drittens um so größer sein, je ungeläufiger die Aufgabe (H4). Beide Hypothesen werden mit der beschriebenen Analyse der Planungszeit mit den Speichergruppen als zusätzlichem Faktor getestet. Die Planungszeit hängt dabei, wie von H3 gefordert, von der Häufigkeit des Speichereinsatzes ab: sie steigt von den häufig über die regelmäßig bis zu den selten eingesetzten Speichern. Allerdings erreicht nur der Unterschied zwischen den häufig und selten eingesetzten Speichern das Signifikanzniveau. Die Differenzierung dieser Hypothese durch die Aufgabengeläufigkeit (H4) ist als Tendenz vorhanden, erreicht aber nicht die zur Bestätigung der Hypothese notwendige Stärke. In den Mittelwerten ist der Effekt deutlich: die Planungszeit unterscheidet sich zwischen den drei Speichergruppen bei sehr geläufigen Aufgaben nicht, wohl aber bei geläufigen und ungeläufigen Aufgaben. Entsprechend trennen t-Tests die Planungszeit zumindest zwischen häufigen und seltenen Speicher bei geläufigen und ungeläufigen Aufgaben signifikant. Die zur Bestätigung der Hypothese erforderliche Interaktion zwischen Geläufigkeit und Speichergruppe erreicht das Signifikanzniveau jedoch nicht. Dabei ist aber zu berücksichtigen, daß die Möglichkeit einer signifikanten Interaktion verringert war, da die Auswertungsmethode zum Ausfall etlicher Versuchspersonen führte. So erscheint es sinnvoll, Hypothese 4 zunächst als plausibel, aber nicht bestätigt, für weitere Untersuchungen offen zu lassen.

2.2 Zweiter empirischer Teil: metakognitive Parameter

Die Ergebnisse im ersten empirischen Teil belegen die Hypothesen zur Planungszeit, zur Speicherzahl und zum Zusammenhang zwischen beidem. Deutlich ist, daß je nach Geläufigkeit einer Repräsentationsaufgabe Prozesse ablaufen, die sich in ihrem Zeitbedarf und in der Qualität ihres Ergebnisses - der Anzahl unterschiedlicher Speicher - unterscheiden. Über das, was in den unterschiedlichen Planungszeiten abläuft, sind nur indirekte Rückschlüsse möglich; tatsächlich lassen

sich die bisherigen Ergebnisse mit einem sehr viel einfacheren Modell als dem im theoretischen Teil entwickelten erklären.

Im zweiten empirischen Teil steht deswegen die Rolle der metakognitiven Parameter im Mittelpunkt. Sie liefern einen direkteren Zugang zu dem, was in den unterschiedlichen Planungszeiten abläuft. Grundfrage der folgenden Untersuchungen ist, ob Entscheidungen von metakognitiven Parametern geleitet werden. Ziel ist nachzuweisen, daß der Zusammenhang zwischen metakognitiven Parametern und einem gewählten Speicher eine Funktion der Geläufigkeit der Aufgabe ist. Sind Entscheidungen durch solche bestimmt oder gar vorhersagbar, müssen die metakognitiven Parameter bei dem in der Planungszeit ablaufenden Prozeß eine Rolle gespielt haben.

Die erste Studie ist eine Pilotstudie, in der erstmals versucht wird, die Frage nach dem Zusammenhang von metakognitiven Parametern und Speicherwahl als eine Funktion der Aufgabe zu verstehen. Die zweite Studie folgt diesem Gedanken, benutzt aber darüber hinaus konkret die im SWIEGS-Modell angenommenen metakognitiven Parameter, um damit die Speicherwahl vorherzusagen.

2.2.1 Studie 3: Termine vs. Texte: eine Pilotstudie

Dobbs und Rule fanden in ihrer Studie von 1987 keinerlei Zusammenhang zwischen der Einschätzung des eigenen Gedächtnisses und der Einsatzhäufigkeit von externen Speichern. Im theoretischen Teil wurde dieses Ergebnis in zwei Richtungen problematisiert. Erstens dürfte die Einschätzung des Gedächtnisses nicht der einzige bedeutsame Parameter sein. Zweitens ist nach genauerer theoretischer Analyse ein Zusammenhang überhaupt nur dann zu erwarten, wenn die Aufgaben ungewöhnlich sind, für sie also kein Repräsentationsscript besteht (vgl. 1.3.2.1).

Dieser zweite Kritikpunkt soll einer ersten empirischen Prüfung unterzogen werden. Grundthese: der Zusammenhang zwischen metakognitiven Einschätzungen und Speicherwahl hängt von der Aufgabenart ab. Eine sehr geläufige Aufgabe soll durch den Einsatz von Scripten, und das heißt ohne Einfluß der metakognitiven Parameter, gelöst werden. Für ungeläufige Aufgaben dagegen sind keine Repräsentationsscripte vorhanden. Folglich sollten bei diesen Aufgaben metakognitive Parameter die Grundlage für die Entscheidung sein (H5).

Ein zweiter Ansatzpunkt der Studie ist neben dem allgemeinspsychologischen, manipulativen einen differentiellen Aspekt zu testen. Auch aus unterschiedlichen Persönlichkeitseigenschaften müßten unterschiedliche Umgehensweisen mit externen Speichern folgen. Exemplarisch wird die Ängstlichkeit gewählt, da hier deutliche Effekte zu erwarten sind. Hochängstliche haben eine schlechtere Selbsteinschätzung und ein höheres Sicherheitsbedürfnis als Niedrigängstliche (Schönpflug, 1989a). Für

die Evaluation und Konstruktion eines Speichers heißt das, daß sich das größere Sicherheitsbedürfnis Ängstlicher in schärferen Prüfgrößen niederschlagen sollte. Ängstliche sollten also eine höhere minimale Erfolgswahrscheinlichkeit haben und dafür zu höherem Aufwand bereit sein als Niedrigängstliche. Für die Anwendung eines Repräsentationsscriptes dürfte sich ein erhöhtes Sicherheitsbedürfnis nur indirekt bemerkbar machen. Hochängstliche müßten ihre eigene Gedächtnisleistung generell schlechter einschätzen als Niedrigängstliche. Daher sollten sie mehr Repräsentationsscripte mit einem externen Speicher gebildet haben als Niedrigängstliche. Dagegen sollte die Frage, auf welche Art ein Speicher gewählt wird, unabhängig von der Ängstlichkeit sein. Ängstlichkeit sollte das 'was?' nicht aber das 'wie?' der Entscheidung beeinflussen. Bei beiden Ängstlichkeitsgruppen dürften die metakognitiven Parameter nur dann die Entscheidung für einen Speicher leiten, wenn die Aufgabe ungeläufig ist. Während Ängstlichkeit also einen Einfluß auf die Entscheidungsparameter und die vorhandenen Scripte haben sollte, sollte die strukturelle Komponente der Entscheidung unbeeinflusst bleiben.

Drei Aufgabentypen werden eingesetzt: Termine, kurze Texte und Texte, die eine Information beschreiben. Die Aufgabenstellung, auf die die Probanden sich vorbereiten sollten, war entweder das Einhalten des jeweiligen Termins oder das merken der jeweiligen Information bis zu einem bestimmten Zeitpunkt. Termine einzuhalten ist eine geläufige Aufgaben, für die ein Script vorliegen sollte. Tatsächlich weist darauf ja nicht nur die Alltagsbeobachtung (vgl. 1.3.3.1), sondern auch die Ergebnisse zu den Terminaufgaben in Studie 1 und Studie 2 hin. Texte zu merken sollte dagegen ungeläufig sein, da ein Kontext, der einen Sinnzusammenhang stiftet und bestimmte Scripte aktivieren könnte (z.B. „Diesen Text zu merken ist Teil der Vordiplomsprüfung“), fehlt. Der Einsatz von Scripten ist damit nach den Annahmen des SWIEGS-Modells nicht möglich. Das sollte bei der dritten Aufgabengruppe verstärkt gelten, da bei Texten, die die zu merkende Information nur beschreiben, noch weniger über die eigentliche Zielinformation bekannt ist. Folglich sollten bei kurzen Texten und beschriebener Information metakognitive Parameter die Entscheidung für oder gegen einen Speicher leiten; nicht aber bei Terminen.

Studie 3 testet als Pilotstudie eine vereinfachte Form dieser Hypothese. Wenn alle Aufgaben in einer leichteren und einer schwierigeren Form vorliegen, sollte sich das in unterschiedlicher Einschätzung der Erfolgswahrscheinlichkeit einer Speicherung dieser Aufgabe im Gedächtnis niederschlagen. Dieses vorausgesetzt, ist die kritische Frage, ob diese unterschiedliche Einschätzung auch zu unterschiedlichen Speicherwahlen führen, ob also das Gedächtnis tatsächlich weniger oft gewählt wird, wenn die Einschätzung seines Erfolges geringer ist. Aus den theoretischen Überlegungen folgt die Hypothese, daß bei den ungeläufigeren Aufgaben - den Texten und der beschriebene Information - das auch der Fall sein müßte, nicht aber bei sehr geläufigen Aufgaben - den Terminen.

Wie oft das Gedächtnis als Speicher eingesetzt wird, müßte also unabhängig von der Einschätzung des Erfolges sein, den der Einsatz des Gedächtnisses zum Merken dieses Termins bringt.

Material

Von jedem der drei Aufgabentypen wurden 16 Aufgaben konstruiert. Der erste Aufgabentyp, die Termine, waren entsprechend der Alltagserfahrung konstruiert, z.B. 'Du hast in 2 Tagen, 11 Uhr einen Termin beim Zahnarzt'. Der zweite Aufgabentyp bestand aus kurzen Texten unterschiedlichsten Inhalts wie Zitaten, Definitionen, Theaterstücken usw. von durchschnittlich 30 Worten, z.B. „Die Freiheit existiert, und auch der Wille existiert, aber eine Willensfreiheit existiert nicht, denn ein Wille, der sich auf seine Freiheit richtet, stößt ins Leere. (Thomas Mann)“. Beim dritten Aufgabentyp waren die Zielinformationen nur beschrieben, z.B. „Stelle Dir vor, wir würden Dir eine Abhandlung über Maler geben, und du sollst dir Informationen über Monet, Picasso und Van Gogh aneignen.“ Jede Aufgabe wurde in zwei Versionen präsentiert. In der leichteren Version lag der Termin in zwei Tagen bzw. war die jeweilige Information bis in zwei Tagen zu merken. In der schwierigen Version betrug die Zeitspanne zwei Wochen: der Termin lag also in zwei Wochen und die Information war bis in zwei Wochen zu merken. Das Versuchsmaterial findet sich vollständig im Anhang 5.1.3.

Versuchspersonen, Design und Durchführung

Am Versuch nahmen 9 Studenten und 13 Studentinnen der Psychologie als Versuchspersonen teil, deren Teilnahme mit einem Versuchspersonenschein honoriert wurde. Die unabhängige Variable ist die Ängstlichkeitsgruppe, zu der der jeweilige Proband gehört, mit zwei Ausprägungen. Die Art der Aufgabe mit den Faktorstufen Termin, kurzer Text und beschriebene Information sowie die Zeitdauer mit den Faktorstufen zwei Tage und zwei Wochen sind als Meßwiederholungsvariablen operationalisiert. Das Design ist folglich 2*2*3 gemischt faktoriell. Abhängige Variablen sind die Einschätzung der Erfolgswahrscheinlichkeit, eine Aufgabe im eigenen Gedächtnis zu behalten sowie die Entscheidung für oder gegen das eigene Gedächtnis als Speicher.

Den Probanden wurden in einer ersten Phase die 48 Aufgaben auf Karteikarten in zufälliger Reihenfolge präsentiert. In der ersten Versuchsphase schätzen die Probanden für jede Aufgabe ein, wie groß der Erfolg in der Bewältigung der Aufgabe wäre, wenn sie die jeweilige Aufgabe nur mit Hilfe des eigenen Gedächtnisses merken würden. Diese Einschätzung erfolgte auf einer Prozentskala, wobei entsprechend 100% absolute Sicherheit, die Aufgabe erfolgreich durch das Gedächtnis bewältigen zu können, und 0% absolute Sicherheit, dies nicht zu können, bedeutet. Die Skala lag den Probanden als Erinnerungshilfe vor. In einer Zwischenphase füllten die Probanden den Trait-Teil der deutschen Ausgabe (Laux, Glanzmann, Schaffner & Spielberger, 1981) des STAI (Spielberger, Gorsuch &

Lushene, 1970) aus. In der dritten Versuchsphase erhielten die Probanden alle Aufgaben ein zweites Mal in der gleichen Reihenfolge. Die Probanden sollten sich entscheiden, ob sie diese Aufgabe nur mit dem Gedächtnis oder mit einem externen Speicher lösen würden.

Die Probanden wurden ausführlich über die Aufgabe instruiert; drei Trainingsaufgaben vor den beiden Einschätzungen sollten zudem Sicherheit im Umgang mit der Skala bzw. der Entscheidungssituation gewährleisten.

Ergebnisse

Mit den Trait-Werten des STAI wurden über einen Mediansplit zwei Versuchspersonengruppen gebildet. Die Gruppe der niedriger Ängstlichen hatte ein Minimum von 28 und ein Maximum von 38 Skalenpunkten mit einem Mittelwert von 33.2, die Gruppe der höher Ängstlichen hatte Werte von 40 bis 54 Skalenpunkten mit einem Mittelwert von 44.9.

Die Einschätzung der Gedächtnisleistung und die Entscheidungen für oder gegen einen externen Speicher sind in Abb. 12 dargestellt. Die Einschätzung der Wahrscheinlichkeit einer erfolgreichen Bearbeitung der Aufgabe mit dem Gedächtnis ist bei einer Zeitdauer von zwei Tagen zwischen dem Einprägen und Rückruf deutlich höher als bei einer Dauer von zwei Wochen (vgl. die schwarzen Balken in Abb. 12). Dies gilt für alle drei Aufgabengruppen und beide Versuchspersonengruppen. Dabei wird der Erfolg für Termine höher eingeschätzt als für Texte und beschriebene Information. Eine Varianzanalyse mit der Ängstlichkeit als unabhängiger Variable sowie Aufgabenart und Zeitspanne als Meßwiederholungsvariable zeigt hochsignifikante Haupteffekte für die Aufgabenart, $F(2,40)=45.1$, $p<.001$, und die Zeitspanne, $F(1,20)=40.4$, $p<.001$, sowie eine signifikante Interaktion zwischen Aufgabenart und Ängstlichkeit, $F(2,40)=4.0$, $p<.05$. Wie t-Tests für Meßwiederholungsvariablen zeigen, wird der signifikante Unterschied in der Einschätzung der Gedächtnisleistung im Vergleich beider Zeitspannen von allen Aufgabengruppen getragen: $t(21)=5.38$ für Termine, $t(21)=-5.45$ für kurze Texte und $t(21)=-4.12$ für beschriebene Information; jeweils $p<.001$. T-Tests mit der Ängstlichkeit als unabhängiger Variable zur Klärung der Interaktion zeigen, daß sich bei kurzen Texten und beschriebener Information die Einschätzung des Erfolges mit dem Gedächtnis als Speicher zwischen beiden Ängstlichkeitsgruppen nicht unterscheidet, $t(20)=.8$, $p>.1$ bzw. $t(20)=1.1$, $p>.1$. Bei den Terminen dagegen wird das Signifikanzniveau annähernd erreicht, $t(20)=-1.86$, $p=.08$. Höher Ängstliche schätzen ihre Gedächtnisleistungen bei den beiden Textarten also nicht anders, bei den Terminen aber schlechter ein als Niedrigängstliche.

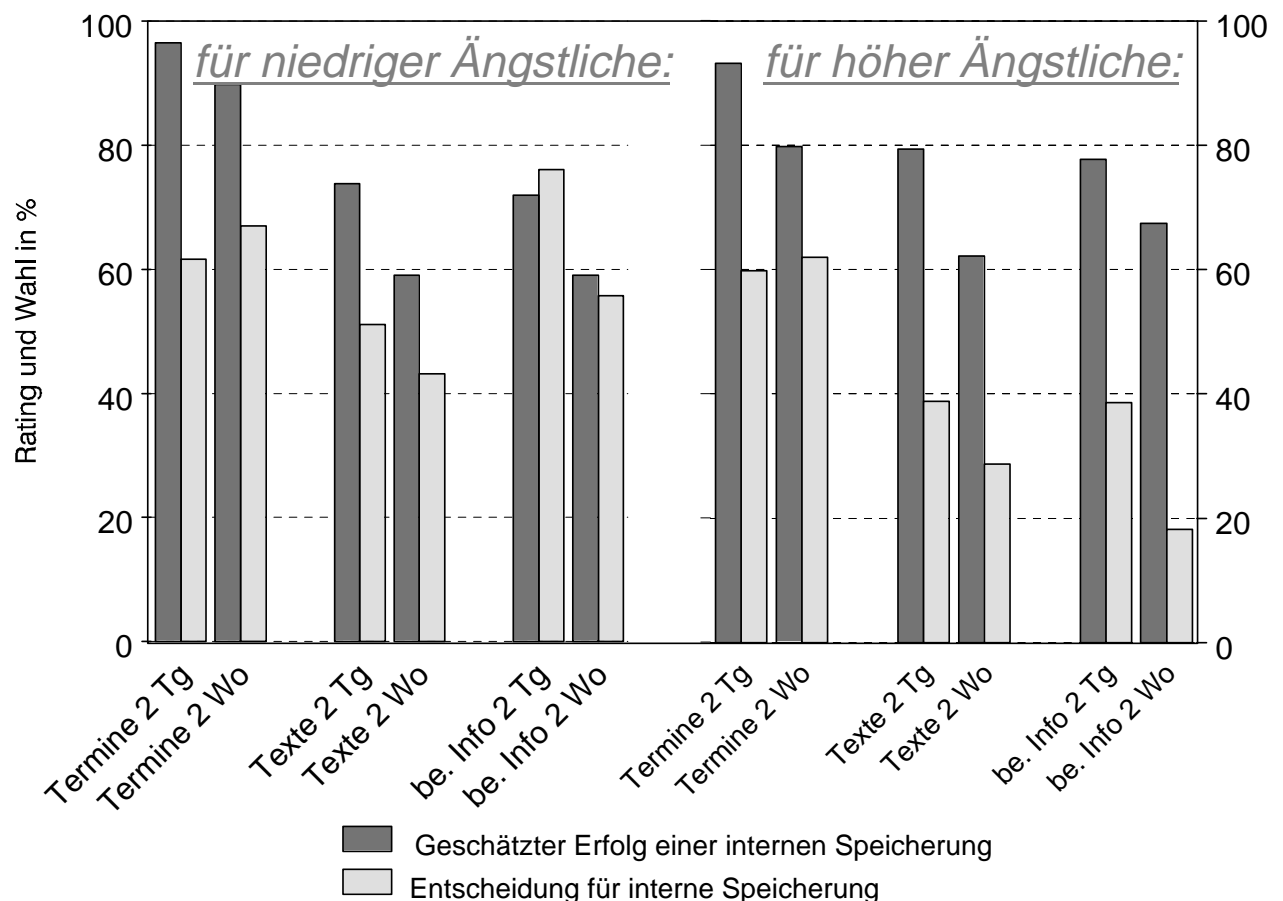


Abb. 12 Die Einschätzung der Wahrscheinlichkeit, die jeweilige Aufgabe mit Hilfe des Gedächtnisses erfolgreich zu lösen kontrastiert mit der Häufigkeit der Entscheidung für das Gedächtnis als Speicher; getrennt für niedriger vs. höher Ängstliche, die drei Aufgabentypen Termine vs. kurze Texte vs. beschriebene Information (be. Info) und die beiden Zeitspannen zwei Tage (2 Tg) vs. 2 Wochen (2Wo).

Bei der Entscheidung zwischen dem Gedächtnis und einem externen Speicher entscheiden sich höher Ängstliche seltener für das eigene Gedächtnis als niedriger Ängstliche. Die Zeitspanne bis zum Erinnern wirkt sich deutlich bei beiden Textarten aus: bei 14-tägiger Spanne wird das Gedächtnis viel seltener gewählt. Bei den Terminen dagegen hat die Zeitspanne keinen Einfluß auf die Entscheidung zwischen externen und internen Speicher: Termine werden sowohl wenn sie in zwei Tagen als auch wenn sie in zwei Wochen liegen gleich häufig im Gedächtnis gespeichert. Eine Varianzanalyse mit der Ängstlichkeit als unabhängiger Variable sowie Aufgabenart und Zeitspanne als Meßwiederholungsvariable zeigt signifikante Haupteffekte für die Ängstlichkeit, $F(1,20)=10.9, p<.01$ und die Zeitspanne, $F(1,20)=7.7, p=.01$. Die einzige signifikante Interaktion ist die zwischen Aufgabenart und Zeitspanne, $F(2,40)=3.6, p<.05$. T-Tests zur Klärung der Interaktion zeigen, daß für kurze Texte und beschriebene Information die Wahl des Gedächtnisses bei 14-tägiger Zeitspanne im Vergleich zur 2-tägigen Zeitspanne signifikant seltener ist, $t(21)=2.2, p<.05$, bzw. $t(21)=3.05, p<.01$. Bei Terminen dagegen

führt die unterschiedliche Zeitspanne nicht zu unterschiedlicher Häufigkeit bei der externen Speicherung, $t(21)=.024$, $p>.1$.

Diskussion

Die Ergebnisse bestätigen die Hypothesen zur Einschätzung des Gedächtnisses und zur Wahl des Gedächtnisses. Aufgaben mit einer Zeitspanne von 14 Tagen sollten weniger erfolgreich mit dem Gedächtnis zu merken sein als solche mit einer Zeitspanne von zwei Tagen. Entsprechend unterschiedlich wurde die Erfolgswahrscheinlichkeit für ein Lösung der Aufgaben mit dem Gedächtnis eingeschätzt. Diese wird für Termine zwar generell höher eingeschätzt als für die beiden anderen Aufgabenarten. Der Unterschied zwischen den Versionen mit zweitägiger bzw. zweiwöchiger Dauer ist aber bei allen drei Aufgabentypen signifikant. Dem steht die Speicherwahl gegenüber. Sowohl bei den kurzen Texten als auch bei der beschriebenen Information folgt die Entscheidung der Logik, die man erwarten würde: bei den Aufgaben, bei denen die Probanden das Gedächtnis als schlechter einschätzen, entscheiden sie sich auch seltener für das Gedächtnis als Speicher. Anders bei den Terminen: die Häufigkeit, mit der das Gedächtnis als Speicher gewählt wird, ist bei beiden Versionen der Aufgaben gleich groß. Obwohl für 14-tägige Termine die Wahrscheinlichkeit, diese mit Hilfe des Gedächtnisses einzuhalten, kleiner eingeschätzt wird als für Termine in zwei Tagen, wird das Gedächtnis nicht seltener eingesetzt. Dieses überraschende Ergebnis ist innerhalb des zu prüfenden SWIEGS-Modells logisch: bei Terminen läuft die Speicherwahl scriptgesteuert; metakognitive Einschätzungen spielen dabei keine Rolle.

Dieser Zusammenhang gilt sowohl für niedriger als auch für höher Ängstliche. Ängstliche schätzen ihr Gedächtnis bei Terminen schlechter ein als Nichtängstliche. Daß dieser Effekt nur bei den Terminen auftrat mag daher rühren, daß die anderen Aufgaben zu künstlich waren, um einen Effekt der Ängstlichkeit zu erzeugen. Sie wählen bei allen Aufgaben seltener ihr Gedächtnis als Speicher als weniger Ängstliche. Trotzdem unterscheidet Ängstlichkeit nicht die Entscheidungsstruktur: die unterschiedliche Aufgabenschwierigkeit führt, unabhängig von der Ängstlichkeit der Probanden, bei Terminen nicht zu unterschiedlicher Häufigkeit der Wahl des Gedächtnisses als Speicher.

Bei Terminen schätzen ängstlichere Probanden ihr Gedächtnis schlechter ein als weniger ängstliche, bei beiden anderen Aufgabentypen hingegen nicht. Trotzdem wählen Ängstliche ihr Gedächtnis deutlich seltener als Speicher. Ein plausibler Erklärungsansatz ist, daß Ängstliche strengere Prüfparameter anwenden als weniger Ängstliche, um ihrem höheren Sicherheitsbedürfnis gerecht zu werden. Damit Ängstliche mit einer Lösung zufrieden sind, müssen Aufgaben mit größerer Wahrscheinlichkeit erfolgreich erinnert werden als dies bei weniger Ängstlichen der Fall ist.

Die erste Bestätigung der zentralen Hypothesen zum Zusammenhang zwischen metakognitiven Parametern und Speicherwahl ermutigt zu differenzierteren Studien. Denn der zu verifizierende Zusammenhang zwischen Geläufigkeit eines Aufgabentyps und dem Prozeß der Speicherwahl ist hier nur durch die drei unterschiedlichen Aufgabentypen operationalisiert worden. Studie 3 läßt für sich gesehen nur Aussagen über die Speicherwahl bei Terminen zu; für eine Generalisierung auf weitere Aufgaben und das eigentlich interessierende theoretische Konstrukt, die Aufgabengeläufigkeit, fehlt die empirische Basis.

Aber auch die Interpretation in Bezug auf die Termine läßt eine Alternativerklärung zu. Wenn den Probanden Termine, die in 14 Tagen liegen, weniger wichtig erscheinen als solche, die in zwei Tagen liegen, wäre es auch plausibel, daß die 14-Tage-Termine nicht seltener im Gedächtnis gemerkt werden. Geringere Valenz der Aufgaben würde zu einer geringeren Erforderlichen Erfolgswahrscheinlichkeit (EEW) und Akzeptiertem Aufwand der Speicherung (AAS) führen. Die schlechtere Merkbarkeit würde zwar zu einer schlechteren Einschätzung der Erfolgswahrscheinlichkeit führen - in der Speicherwahl würde sich das durch die veränderten Prüfparameter aber nicht niederschlagen. Gegen diese Überlegung spricht, daß sie mit wesentlich mehr Parametern arbeitet als die eigentliche Hypothese. Trotzdem muß diese Erklärungsmöglichkeit in den Folgestudien ausgeschlossen sein.

2.2.2 Studie 4: Metakognitive Parameter im Test

Im Zentrum der Studien 1 und 2 standen die Planungszeit und die tatsächlich gewählten Speicher. Studie 3 stellte die Vorhersagbarkeit der Speicherwahl aus metakognitiven Parametern ins Zentrum. Theoretische Annahme dabei war, daß sich in den verschiedenen Planungszeiten ein unterschiedlicher Umgang mit den metakognitiven Parametern ausdrückt. Nach dem SWIEGS-Modell kommt größere Planungszeit dadurch zustande, daß die metakognitiven Parameter in der Entscheidung intensiver geprüft werden. Studie 4 verknüpft deswegen die Ansätze der Studien 1 bis 3, indem sowohl die Planungszeit als auch die Vorhersage der Speicherwahl aus den metakognitiven Parametern als abhängige Variablen operationalisiert werden.

Studie 3 konnte belegen, daß für Termine - als Vertreter sehr geläufiger Aufgaben - kein Zusammenhang zwischen der Einschätzung, wie gut das Gedächtnis als Speicher fungieren würde und der Entscheidung für oder gegen das Gedächtnis als Speicher besteht. Bei kurzen Texten und beschriebener Information - als Vertreter ungeläufiger Aufgaben - führt unterschiedliche Einschätzung der Merkbarkeit der Aufgaben mit dem Gedächtnis dagegen zu einem signifikanten Unterschied in der Häufigkeit, mit der das Gedächtnis als Speicher gewählt wird.

Studie 4 prüft diese Zusammenhänge in generalisierter Form. Der Zusammenhang zwischen metakognitiven Parametern und Wahl eines Speichers soll nach dem SWIEGS-Modell nicht von den Aufgabentypen abhängen, sondern von der Geläufigkeit einer Aufgabe (H5). Die Geläufigkeit entscheidet über den Zusammenhang zwischen metakognitiven Parametern und der Wahl des Speichers. In Studie 3 war dieser Parameter nur durch die drei unterschiedliche Aufgabentypen operationalisiert. Studie 4 erweitert dies erheblich, indem fünf unterschiedliche Aufgabentypen als Versuchsmaterial verwandt werden. Jeder Aufgabentyp erscheint auf jeder Geläufigkeitsstufe; so sind Aufgabentyp und Aufgabengeläufigkeit entkoppelt.

Um Geläufigkeit der Aufgaben auch als unabhängige Variable varriieren zu können, werden die Aufgaben in zwei unterschiedlichen Kontexten gestellt. Der reale Kontext entspricht dem von Studie 1 und Studie 2; die Probanden sollen sagen, wie sie die Aufgaben in ihrem Alltag repräsentieren würden. Der virtuelle Kontext sucht Vorerfahrungen weitgehend auszuschalten. Die Aufgaben werden als kurze Texte behandelt, auf deren Wiedererkennen die Probanden sich vorbereiten sollen. Entsprechend der ungeläufigen Aufgaben bei Studie 3 wird keinerlei erklärender Kontext zu diesem Textlernen gegeben. Externe Speicherung ist dabei nur über einen Karteikartenstapel möglich, dessen Nutzung bestimmten Regeln unterliegt. So werden Vorerfahrungen mit dem Speicher ebenfalls ausgeschlossen. In der realen Form der Aufgaben sollte die Geläufigkeit des jeweiligen Aufgabentyps wie in Studie 1 und 2 die Planungszeit und wie in Studie 3 den Zusammenhang zwischen Aufgabe und metakognitiven Parametern steuern. In der virtuellen Form dagegen sollte die Geläufigkeit aller Aufgaben gleich sein. Planungszeit und Vorhersagbarkeit aus den metakognitiven Parametern sollten sich folglich für die verschiedenen Aufgaben nicht mehr unterscheiden. Da die Geläufigkeit aller Aufgaben gering ist, sollten sie wie ungeläufige Aufgaben behandelt werden - also mit langer Planungszeit und hoher Vorhersagbarkeit der Speicherwahl.

Um die Erhebung der metakognitiven Parameter praktikabel zu gestalten, mußte der Umfang der Erhebung reduziert werden. Der vom SWIEGS-Modell vorhergesagte Zusammenhang zwischen metakognitiven Parametern und der Wahl eines Speichers betrifft neben der in Studie 3 getesteten Wahrscheinlichkeit des Erfolges auch den Aufwand der Speicherung. Beide Parameter müßten bei jeder Aufgabe für jeden dem Probanden geläufigen Speicher erhoben werden (bei 60 Aufgaben und 20 Speichern wären das 2400 Einschätzungen). Da das die Probanden völlig überfordern würde, schätzten die Probanden für jede Aufgabe Aufwand und Erfolgswahrscheinlichkeit für zwei wahrscheinliche Speicher ein. Über einen Entscheidungsalgorithmus konnte dann derjenige Speicher bestimmt werden, der nach den Parametern gewählt werden müßte. Der Vergleich dieses Ergebnisses mit dem tatsächlich

gewählten Speicher dient als aussagekräftiger Parameter für den Einfluß der Einschätzungen auf die Speicherwahl.

Eine weitere Erweiterung betrifft die Valenz einer Aufgabe. In Studie 3 wurde sie nicht kontrolliert; tatsächlich könnten die Ergebnisse zu den Terminen aber auch erklärt werden, wenn längerfristige Termine den Probanden weniger bedeutsam erscheinen würden (vgl. Diskussion der Studie 3). Aufgaben so zu konstruieren, daß sie bei unterschiedlicher Geläufigkeit gleiche Valenz haben, ist aber praktisch kaum realisierbar (vgl. Diskussion in Studie 1). Deswegen wird die Valenz als unabhängige Variable eingeführt, so daß ihr Einfluß unabhängig von der Aufgabengeläufigkeit getestet werden kann. Praktisches Problem dabei ist, daß zu hohe Valenz einer Aufgabe diese tatsächlich zu einem anderen Aufgabentyp machen könnte. So sind ein Arzttermin, der über Leben und Tod entscheidet oder ein erstes Rendezvous, nachdem man sich leidenschaftlich verliebt hat, keine Termine wie andere auch. Um die Valenz der Aufgaben zu erhöhen, ohne dadurch andere Aufgabentypen zu konstruieren, werden die Aufgaben in den Kontext einer hypothetische Glücksspielsituation gestellt (vgl. Meacham & Singer, 1977). Einen Gewinn erwarte dabei denjenigen, der mehr Aufgaben erinnert hat als seine Mitspieler. So sollte die Valenz aller Aufgaben gleichmäßig und moderat erhöht werden.

Diese Vorgehen kontrolliert nicht nur die Wirkung der Aufgabenvalenz, sondern erlaubt auch die Testung der Hypothesen zur Rolle der Aufgabenvalenz im SWIEGS-Modell. Höhere Valenz einer Aufgabe sollte zu schärferen Prüfparametern für die Evaluation oder Konstruktion eines Speichers führen. Das SWIEGS-Modell sagt vorher, daß sowohl die Erforderliche Erfolgswahrscheinlichkeit der Speicherung (EEW) als auch der für die Speicherung Akzeptierte Aufwand (AAS) mit der Valenz der Aufgabe steigen. Empirisch müßten sich höhere Prüfkriterien in drei Bereichen niederschlagen. Da höhere Prüfkriterien zu einem aufwendigeren Vergleichsprozeß führen, müßte erstens auch die Planungszeit mit der Valenz der Aufgaben steigen (H7). Zweitens müßte der intensivere Suchprozeß zu differenzierteren Lösungen führen - die Anzahl der für einen Aufgabentyp eingesetzten unterschiedlicher Speicher also bei erhöhter Valenz der Aufgaben größer sein als bei normaler Valenz (H9). Drittens sollte die Bewertung des dann gewählten Speichers in Bezug auf Erfolgswahrscheinlichkeit und Aufwand den Prüfkriterien folgen: Höhere Valenz einer Aufgabe sollte also bei weniger geläufigen Aufgaben zu Lösungen mit höherer Erfolgswahrscheinlichkeit und höherem Aufwand führen (H11).

Speicherwahlen über die Anwendung eines Repräsentationsscripts sollten ohne Einfluß der Prüfparameter erfolgen. Folglich hat auch die erhöhte Valenz keinen Einfluß auf die Speicherwahl. Die drei Hypothesen zur Aufgabenvalenz müssen also nach dem SWIEGS-Modell hinsichtlich der Aufgabengeläufigkeit differenziert werden: Je ungeläufiger eine Aufgabe, desto größer müßte sie

Steigerung der Planungszeit (H8), die Erhöhung der Speicherzahl (H10) und die veränderte Einschätzung des gewählten Speichers (H12) bei erhöhter Valenz sein.

Material

Als Versuchsmaterial dienten die in Studie 2 verwandten 60 Items.

Versuchspersonen und Design

Versuchspersonen im Alter zwischen 21 und 37 Jahren nahmen an der Untersuchung teil. Die Hälfte der Versuchspersonen wurde bezahlt, die andere Hälfte bestand aus Studenten und Studentinnen der Psychologie, die einen Versuchspersonenschein erhielten. In jeder Bedingungsversion war die gleiche Anzahl von bezahlten und unbezahlten Probanden; bis auf diese Einschränkung erfolgte die Zuordnung von Proband zu Bedingung per Zufall.

Unabhängige Variable ist die Valenz der Aufgaben mit den beiden Stufen normal vs. erhöht und die Aufgabenstellung mit den beiden Stufen real vs. virtuell. Die Geläufigkeit der Aufgaben mit den Stufen sehr geläufig vs. geläufig vs. ungeläufig ist als Meßwiederholungsvariable operationalisiert. Das Design ist also $2 \times 2 \times 3$ gemischt faktoriell.

Abhängige Variablen sind die zur Wahl des Speichers erforderliche Zeit, die gewählten Speicher, die Einschätzungen der Aufgaben in den Speichern sowie davon abgeleitete Variablen.

Versuchsdurchführung

Den Probanden wurden die Items in zufälliger Reihenfolge auf einer am Monitor dargestellten Karteikarte präsentiert. In der 'realen' Bedingung waren die Probanden instruiert, sich zu entscheiden, welchen Speicher sie für eine Aufgabe wählen würden, wenn sie diese Aufgabe in ihrem Alltag bearbeiten würden. Für die verdeckte Zeitmessung wurde wieder die in Studie 2 entwickelte Technik der scheinbaren Mikrofonaufnahme angewandt. Nach der ersten Bearbeitung aller 60 Aufgaben ging der Versuchsleiter zusammen mit dem Probanden alle die Aufgaben noch einmal durch, bei denen der Proband sich für das Gedächtnis als Speicher entschieden hatte. Der Proband sollte den externen Speicher nennen, mit dem er die jeweilige Aufgabe am wahrscheinlichsten lösen würde, wenn er für die Aufgabe explizit einen externen Speicher nutzen müßte. So war für die dritte Phase für jede Aufgabe ein möglicher externer Speicher vorhanden.

In der dritten Phase wurden noch einmal alle 60 Aufgabe in der gleichen Reihenfolge wie in der ersten Phase auf realen Karteikarten präsentiert. Diesmal sollte die Erfolgswahrscheinlichkeit und der Aufwand der Speicherung der Aufgabe eingeschätzt werden. Zum einen in Bezug auf das Gedächtnis, zum anderen in Bezug auf den in der ersten oder zweiten Versuchsphase zu dieser Aufgabe genannten

externen Speicher. Vor dem Probanden lagen zwei Skalen, eine mit der Erfolgswahrscheinlichkeit von 0% bis 100 %, die andere mit dem Aufwand von 1 (sehr gering) bis 7 (sehr groß). Über und unter den Skalen stand je eine Frage zu den beiden Speichern, einmal für das Gedächtnis, einmal für den externen Speicher formuliert. Der Versuchsleiter legte eine Karteikarte mit der jeweiligen Aufgabe in die Mitte der Skalen und nannte dem Probanden den externen Speicher, den der Proband für die Aufgabe in der ersten oder zweiten Phase des Versuches genannt hatte. Darauf schätzte der Proband vier Parameter ein: den Aufwand der Speicherung und ihren Erfolg, jeweils sowohl bezüglich der Speicherung der Aufgabe im Gedächtnis als auch bezüglich der Speicherung im externen Speicher.

In der virtuellen Bedingung bearbeiteten die Probanden die gleichen Aufgaben mit anderem Kontext. Die Probanden wurden instruiert, sich auf einen Gedächtnistest vorzubereiten, bei dem sie die 60 Aufgaben in einer zweiten Phase noch einmal präsentiert bekommen würden, teilweise in der Originalform, teilweise in einer leicht veränderten Fassung. Sie sollten sich darauf vorbereiten, dann innerhalb von 12 s zu entscheiden, ob ein Text in der Original- oder einer veränderten Fassung vorläge. Zu überlegen sei, wie sie sich auf diese Aufgabe vorbereiten würden. Bei jeder Aufgabe sollten sie entscheiden, ob sie diese Aufgabe vor dem Hintergrund der genannten Aufgabenstellung im Kopf behalten oder auf eine Karteikarte aufschreiben würden. Ein Karteikartenstapel lag zur Hand; den Probanden wurde sein Einsatz in der Wiedererkennensphase ausführlich erläutert. In Vorversuchen hatte sich herausgestellt, daß ein komplexe Aufgaben- und Speicherform notwendig war, um zu verhindern, daß den Probanden bei einer externen Speicherung 100% Erfolgswahrscheinlichkeit sicher schienen. Deswegen war die Nutzung des Karteikartenstapels in der Wiedererkennensphase an eine Reihe von Regeln gebunden. Die Karteikarten sollten dabei nicht auf dem Tisch ausgebreitet werden dürfen und würden vor der Wiedererkennensphase noch einmal gemischt werden. So wurde die Nutzung des Karteikartenstapels mit der Schwierigkeit verbunden, die extern gespeicherte Information schnell genug wiederzufinden - ein Problem, daß fast immer mit dem Einsatz von externen Speichern verbunden ist (vgl. Esser, 1992). Entsprechend der 'realen' Bedingung mußten in einer zweiten Phase alle Aufgaben noch einmal in Bezug auf Erfolgswahrscheinlichkeit und Aufwand der Speicherung im Gedächtnis bzw. mit Hilfe des Karteikartenstapels beurteilt werden.

Beide Versuchsbedingungen lagen in einer Version mit normaler Valenz und in einer Version mit erhöhter Valenz vor. Bei erhöhter Valenz wurde die Probanden instruiert, sich vorzustellen, daß sie einen Jackpot von DM 10.000,- gewinnen würden, wenn sie am meisten Aufgaben erfolgreich bearbeiteten. Da sich in Vorversuchen gezeigt hatte, daß diese Instruktion leicht vergessen wird, blinkte das Wort 'Jackpot' zur Erinnerung vor jeder neuen Aufgabe auf dem Bildschirm auf.

Die Probanden wurden ausführlich in den verschiedenen Phasen des Versuches instruiert; Probeaufgaben, die beliebig oft wiederholt werden konnten, sollten jeweils die Sicherheit im Umgang mit dem Rechner bzw. den Skalen garantieren. Bei der Einschätzung der Items auf den beiden Skalen wurden für die ersten vier Items die jeweiligen Fragen vom Versuchsleiter laut formuliert. Diese Maßnahmen reichten für eine problemlose Bewältigung aller Aufgaben durch die Probanden aus. Die Versuchsdauer betrug je nach Aufgabentyp und Proband zwischen zweieinhalb und dreieinhalb Stunden.

Ergebnisse

Die Ergebnisdarstellung gliedert sich in sechs Punkte:

- Die zur Wahl eines Speichers benötigte Zeit,
- Die Vorhersagbarkeit der Wahlen aus den metakognitiven Parametern,
- Die Bewertung der eigenen Entscheidung,
- Die Anzahl der je Aufgabengruppe zur Lösung herangezogenen unterschiedlichen Speicher, sowie
- Die Planungszeit in Abhängigkeit von der Speicherwahl

Besonderheiten der Auswertung werden, wenn erforderlich, innerhalb der einzelnen Punkte behandelt.

Die zur Wahl eines Speichers benötigte Zeit

Die Zeiten, die die Probanden brauchen, um sich für einen Speicher zu entscheiden, sind dargestellt in Abb. 13. Bei realer Aufgabenstellung steigt die Planungszeit über die drei Geläufigkeitsstufen der Aufgaben, wobei dieser Gradient bei erhöhter Valenz steiler wird als bei normaler Valenz. Bei virtueller Aufgabenstellung dagegen steigt die Planungszeit nur schwach von sehr geläufigen zu geläufigen Aufgaben; von geläufigen zu ungeläufigen Aufgaben bleibt sie gleich. Höhere Valenz führt hier nicht zu einer Verlängerung, sondern sogar zu einer Verkürzung der Planungszeit; dieser Effekt ist für alle Geläufigkeitsstufen gleich.

Eine Varianzanalyse mit Aufgabenstellung und Valenz als unabhängige Variable sowie Geläufigkeit als Meßwiederholungsvariable ergibt signifikante Haupteffekte für Aufgabenstellung, $F(1,76)=79,28$, $p<.001$, sowie Geläufigkeit der Aufgaben, $F(2,152)=247.31$, $p<.001$. Aufgabenstellung interagiert signifikant mit Geläufigkeit, $F(2,152)=81.55$, $p<.001$, und Valenz, $F(1,76)=11.86$, $p=.001$. Auch die Interaktion zwischen Valenz und Geläufigkeit ist signifikant, $F(2,152)=3.50$, $p<.05$, ebenso die Dreifachinteraktion von Valenz, Aufgabenstellung und Geläufigkeit, $F(2,152)=6.45$, $p<.01$.

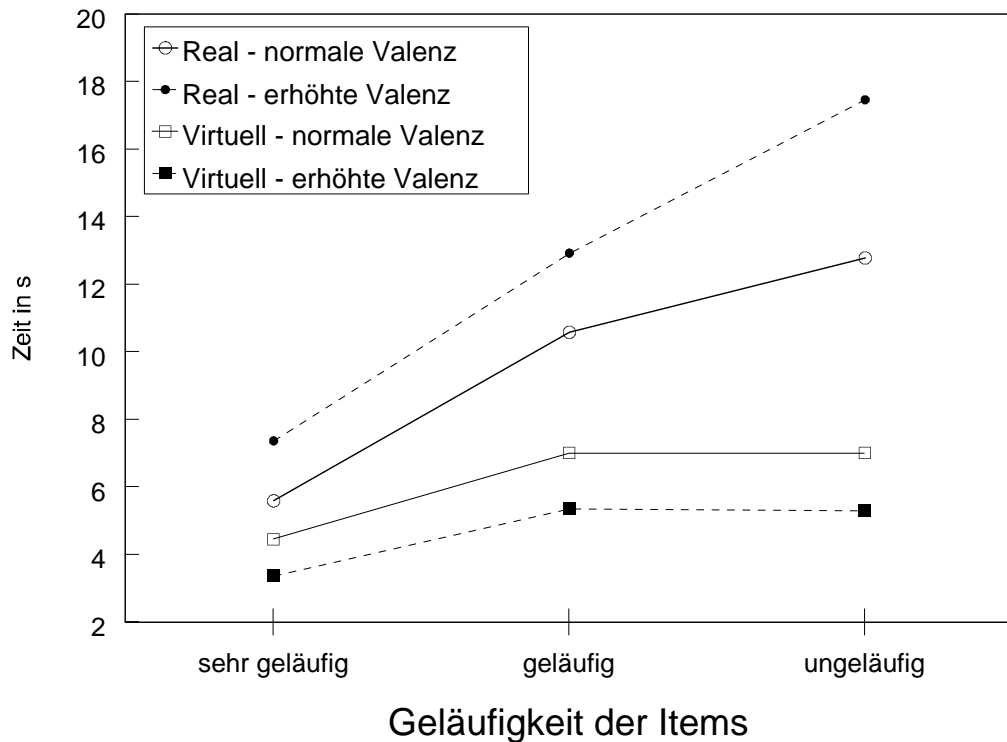


Abb. 13 Die Zeit von der Präsentation der Aufgabe bis zu Nennung des Speichers - getrennt für die vier Probandengruppen und die Geläufigkeit der Aufgaben.

Zur Klärung der Interaktion werden t-Tests zum Paarvergleich berechnet. Entlang der Geläufigkeitsstufen bestätigen sich die beschriebenen Mittelwertsunterschiede. Während sich die Planungszeiten bei realer Aufgabenstellung für alle drei Geläufigkeitsstufen signifikant voneinander unterscheiden, unterscheiden sich bei virtueller Aufgabenstellung die geläufigen und ungeläufigen Aufgaben nicht signifikant. Diese Beziehung gilt sowohl mit normaler als auch erhöhter Valenz (siehe Tab. 6, oberer Teil).

Die Erhöhung der Planungszeit durch Erhöhung der Aufgabenvalenz bei realer Aufgabenstellung ist für jede Geläufigkeitsstufe signifikant; nur bei mittlerer Geläufigkeit wird das Signifikanzniveau mit, $t(38)=-1.88$, $p=.068$ knapp verpaßt. Die Verringerung der Planungszeit durch erhöhte Valenz bei virtueller Aufgabenstellung dagegen erreicht bei allen drei Geläufigkeitsstufen das Signifikanzniveau (siehe Tab. 6, unterer Teil). Alle Mittelwertsvergleiche entlang der Aufgabenstellungen real vs. virtuell sind bei gleicher Valenz und gleichen Geläufigkeitsstufe signifikant (aus Platzgründen nicht aufgeführt).

Die signifikante Dreifachinteraktion kommt folglich durch die je nach Aufgabentyp unterschiedliche Wirkung der Aufgabenvalenz auf die Geläufigkeitsstufen zustande. Bei realer Aufgabenstellung steigt die Planungszeit bei erhöhter Valenz um so mehr, je ungeläufiger die

Planungszeiten bei realer vs. virtueller Aufgabenstellung sowie normaler vs. erhöhter Valenz (Mittelwertsvergleiche über t-Tests)

Vergleich zwischen Geläufigkeitsstufen (t-Tests für Meßwiederholungsvariablen)						
	Valenz normal			Valenz erhöht		
<i>Aufgabenstellung:</i>	sehr geläufig vs. geläufig	geläufig vs. ungeläufig	sehr geläufig vs. ungeläufig	sehr geläufig vs. geläufig	geläufig vs. ungeläufig	sehr geläufig vs. ungeläufig
<i>real</i>	t(19)=-10.85***	t(19)=-3.78**	t(19)=-10.92***	t(19)=-10.53***	t(19)=-6.05***	t(19)=-10.98***
<i>virtuell</i>	t(19)=-9.96***	t(19)=-0.01	t(19)=-8.43***	t(19)=-7.93***	t(19)=-0.3	t(19)=-9.56***
Vergleich zwischen Valenz normal vs. erhöht (t-Tests für unabhängige Variablen)						
	Aufgabenstellung real			Aufgabenstellung virtuell		
<i>Geläufigkeitsstufe:</i>	sehr geläufig	geläufig	ungeläufig	sehr geläufig	geläufig	ungeläufig
	t(38)=-2.3*	t(38)=-1.88	t(38)=-2.79**	t(38)=2.71**	t(38)=2.55*	t(38)=2.72**

Tab. 6 T-Tests zum Vergleich der Mittelwerte der Planungszeit. * für $p < .05$; ** für $p < .01$; *** für $p < .001$

Aufgabe. Bei virtueller Aufgabenstellung dagegen führt erhöhte Valenz zu einer Verkürzung der Planungszeit, die für alle drei Geläufigkeitsstufen gleich groß ist.

Die Vorhersagbarkeit der Wahlen aus den metakognitiven Parametern

Um den Einfluß der von den Probanden geschätzten Parameter auf die Entscheidung für einen Speicher auswerten zu können, wurde zunächst der Speicher berechnet, den der Proband wählen müßte, wenn er seiner Einschätzung folgen würde. Dieser wurde mit dem vom Probanden tatsächlich gewählten Speicher verglichen; stimmte beide überein, wurde die Entscheidung als ‘rational’, stimmten sie nicht überein als ‘nicht rational’ (im Sinne von: dem Rationalen der Parameter folgend oder nicht) bewertet.

Für jede Aufgabe schätzte der Proband Aufwand und Erfolgswahrscheinlichkeit zweier Speicheralternativen. Diese vier Parameter können in vier Beziehungen zu einer Speicherwahl stehen:

- 'eindeutig': Aufwand und Erfolg weisen in die gleiche Richtung, d.h. einer der beiden Speicher wird sowohl als erfolgreicher als auch als weniger aufwendig eingeschätzt;

- 'eindeutig' Aufwandoder Erfolg weisen auf einen Speicher, wobei der jeweils andere Parameter für beide Speicher gleich eingeschätzt wurde; z.B. der Aufwand ist für beide Speicher gleich groß, aber einer der beiden wird als erfolgreicher eingeschätzt;
- 'zwar - aber': während der Aufwand für den einen Speicher spricht, spricht der Erfolg für den anderen; ein Speicher wird also zwar als weniger aufwendig in der Anwendung, aber auch als weniger erfolgreich eingeschätzt;
- 'unentschieden': Aufwand und Erfolg werden für beide Speicher gleich eingeschätzt.

Je eindeutiger eine Parameterkombination auf einen Speicher hinweist, um so häufiger taucht sie auf. Die Häufigkeit der Kombinationen halbiert sich jeweils von 'eineindeutig' (55%) zu eindeutig (24%), von 'eindeutig' zu 'zwar - aber' (13%) und von 'zwar - aber' zu 'unentschieden' (7.5%). Je nach Parameterkombination läßt sich der diesen Parametern folgende Speicher unterschiedlich festlegen. Sowohl die eindeutigen als auch die eineindeutigen Kombinationen weisen unmittelbar auf einen Speicher hin; wird dieser auch tatsächlich von der Versuchsperson gewählt, gilt die Entscheidung als 'rational': die Wahl eines Speichers stimmt mit der Bewertung dieses Speichers überein. Umgekehrt gilt eine Entscheidung als 'nicht rational', wenn der Speicher gewählt wird, der schlechter eingeschätzt wurde. Die Kombination 'unentschieden' ist immer nicht rational, da die Entscheidung für einen Speicher nicht durch die Parameter bestimmt sein kann. Differenzierterer Analyse bedarf die Kombination 'zwar - aber'. Hier weist ja der Vergleich des geschätzten Aufwandes auf einen anderen Speicher hin als der Vergleich der geschätzten Erfolgswahrscheinlichkeit beider Speicher. Trotzdem ist es möglich, rationale von nicht rationalen Urteilen zu unterscheiden, indem man die Konsistenz der Urteile überprüft. Lassen sich die Urteile mit Regeln beschreiben, die Aufwand und Erfolg von zwei Speicheralternativen in Beziehung setzen, so gelten die Urteile als rational. Annahme dabei ist, daß die Schätzung des Aufwandes auf Intervallskalenniveau stattfindet. Dann läßt sich im Vergleich zweier Speicher für jede Aufwandsdifferenz eine Differenz in der Erfolgswahrscheinlichkeit errechnen, die mindestens erreicht werden muß, um den jeweiligen Speicher zu wählen. Beispielsweise mag ein Proband eine Aufwandssteigerung um 2 Punkte in Kauf nehmen, um eine Steigerung der Erfolgswahrscheinlichkeit von 20% oder mehr zu erreichen; nicht aber für eine Steigerung der Erfolgswahrscheinlichkeit von 10% oder weniger. Weitere Beispiele und eine differenzierte Beschreibung dieser Auswertung finden sich in Anhang 5.1.2. Diese Auswertung ist nicht ganz unproblematisch, da sie nicht - wie bei den anderen Urteilkombinationen - unabhängig von der Kenntnis der tatsächlichen Speicherwahl erfolgen kann. Sie spielt allerdings insgesamt keine große Rolle, da sie nur bei 13% der Urteile erforderlich war und die zu überprüfenden Hypothesen in keinem systematischen Zusammenhang mit dieser Urteilkombination stehen. Trotzdem wurde sie zusätzlich

von einem zweiten unabhängigen Beurteiler durchgeführt; 94% der 610 Urteile stimmten überein, Differenzen ließen sich einvernehmlich klären.

Bei realer Aufgabenstellung steigt die Vorhersagbarkeit der Speicherwahl mit sinkender Geläufigkeit der Aufgaben, bei virtueller Aufgabenstellung bleibt sie dagegen über die Geläufigkeitsstufen gleich (siehe Abb. 14).

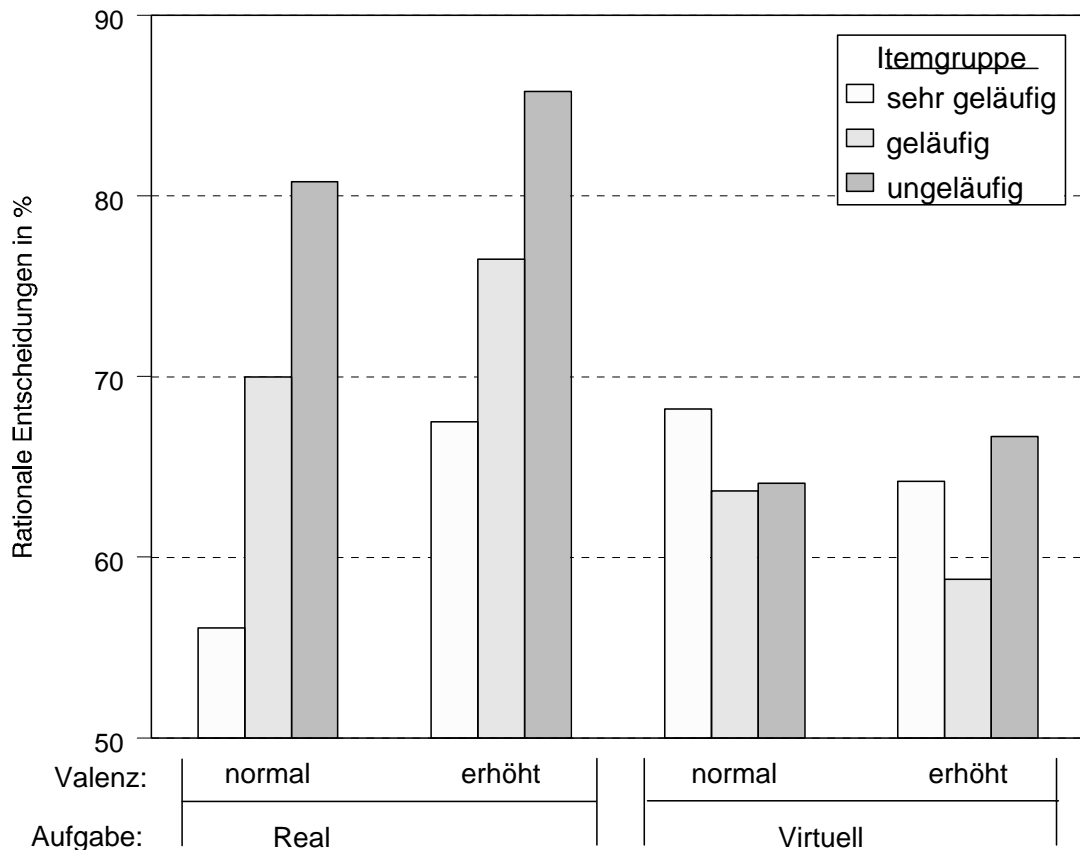


Abb. 14 Der Anteil der Speicherwahlen, bei denen der aus den Parametern geschlossene Speicher mit dem gewählten übereinstimmt (=rational). Da nur zwei Speicher zur Wahl standen, entspricht die Grundlinie der Zufallswahrscheinlichkeit von 50%.

Eine Varianzanalyse des Anteil rationaler Urteile mit der Aufgabenstellung und Valenz als unabhängige Variablen sowie der Geläufigkeit der Aufgabe als Meßwiederholungsfaktor zeigt signifikante Haupteffekte für die Aufgabenstellung $F(1,76)=5,73, p<.05$ und die Geläufigkeit $F(2,152)=9.14, p<.001$ sowie für die Interaktion von Geläufigkeit und Aufgabenstellung $F(2,152)=11.12, p<.001$. Obwohl erhöhte Valenz zu einer Verbesserung der Vorhersageleistung bei der realen Aufgabenstellung (von 69,4% auf 76,5%) und zu einer leichten Verschlechterung bei der virtuellen Aufgabenstellung führt (von 65,3% auf 62,2%), wird die Interaktion von Valenz und Aufgabenstellung nicht signifikant, $F(1,76)=1,85, p>.1$. Auch der Haupteffekt der Valenz sowie alle

weiteren Interaktionen mit ihr erreichen das Signifikanzniveau nicht. Eine Partialanalyse für die reale Aufgabenstellung zeigt einen signifikanten Effekt der Geläufigkeit des Aufgabentyps, $F(2,78)=28.78$, $p<.001$, bei virtueller Aufgabenstellung zeigt die Aufgabengeläufigkeit keinerlei Effekte auf den Anteil rationaler Urteile, $F(2,78)=.81$, $p>.05$.

T-Tests zum Vergleich der einzelnen Mittelwerte stützen dieses Bild. Während sich in der realen Aufgabenstellung alle Mittelwerte der verschiedenen Geläufigkeitsstufen für beide Bedingungen hochsignifikant unterscheiden, unterscheiden sich die Mittelwerte der virtuellen Aufgabenstellung nicht voneinander (siehe Tab. 7).

Vorhersage der Speicherwahl aus den Parametern (Mittelwertsvergleiche über t-Tests)

Vergleich zwischen Geläufigkeitsstufen (t-Tests für Meßwiederholungsvariablen)						
	Valenz normal			Valenz erhöht		
<i>Aufgabenstellung:</i>	sehr geläufig vs. geläufig	geläufig vs. ungeläufig	sehr geläufig vs. ungeläufig	sehr geläufig vs. geläufig	geläufig vs. ungeläufig	sehr geläufig vs. ungeläufig
<i>real</i>	t(19)=-3.76***	t(19)=-3.05**	t(19)=-4.77***	t(19)=-2.33*	t(19)=-3.41***	t(19)=-3.94***
<i>virtuell</i>	t(19)=1.06	t(19)=-0.26	t(19)=-0.84	t(19)=-0.62	t(19)=-0.48	t(19)=-0.2

Tab. 7 T-Tests zum Vergleich der Mittelwerte des Anteils rationaler Urteile. * für $p<.05$; ** für $p<.01$; *** für $p<.001$

Die Interaktion von Aufgabenart und Geläufigkeit ergibt sich folglich, weil die Geläufigkeit der Aufgaben bei realer Aufgabenstellung die Vorhersagbarkeit der Speicherwahl bestimmt: je ungeläufiger eine Aufgabe, um so besser ist die Speicherwahl durch die metakognitiven Parameter vorhersagbar. Bei virtueller Aufgabenstellung dagegen unterscheidet sich die Vorhersagbarkeit der Speicherwahl zwischen Aufgaben unterschiedlicher Geläufigkeiten nicht.

Die Bewertung der eigenen Entscheidung

Die Probanden gaben in der dritten Versuchsphase für zwei zur Lösung der jeweiligen Aufgabe mögliche Speicher eine Einschätzung des Aufwandes und des Erfolges dieser Speicherung ab. Die folgende Auswertung betrachtet nun die Einschätzung derjenigen Alternative, für die sich der Proband in der ersten Versuchsphase entschieden hatte.

Die Bewertung dieser vom Probanden gewählten Lösung fällt insgesamt bei realer Aufgabenstellung deutlich besser aus als bei virtueller. Dabei wird die Bewertung mit sinkender Geläufigkeit der Aufgaben schlechter. Der geschätzte Aufwand steigt mit sinkender Geläufigkeit der

Aufgabe für reale Aufgabenstellung kontinuierlich, während er für virtuelle Aufgabenstellung nur von der sehr geläufigen zur geläufigen, nicht aber von der geläufigen zur ungeläufigen Aufgabengruppe steigt. Die Valenz der Aufgabe hat dabei keinerlei Einfluß auf den geschätzten Aufwand der Speicherung (siehe Abb. 15, linker Teil).

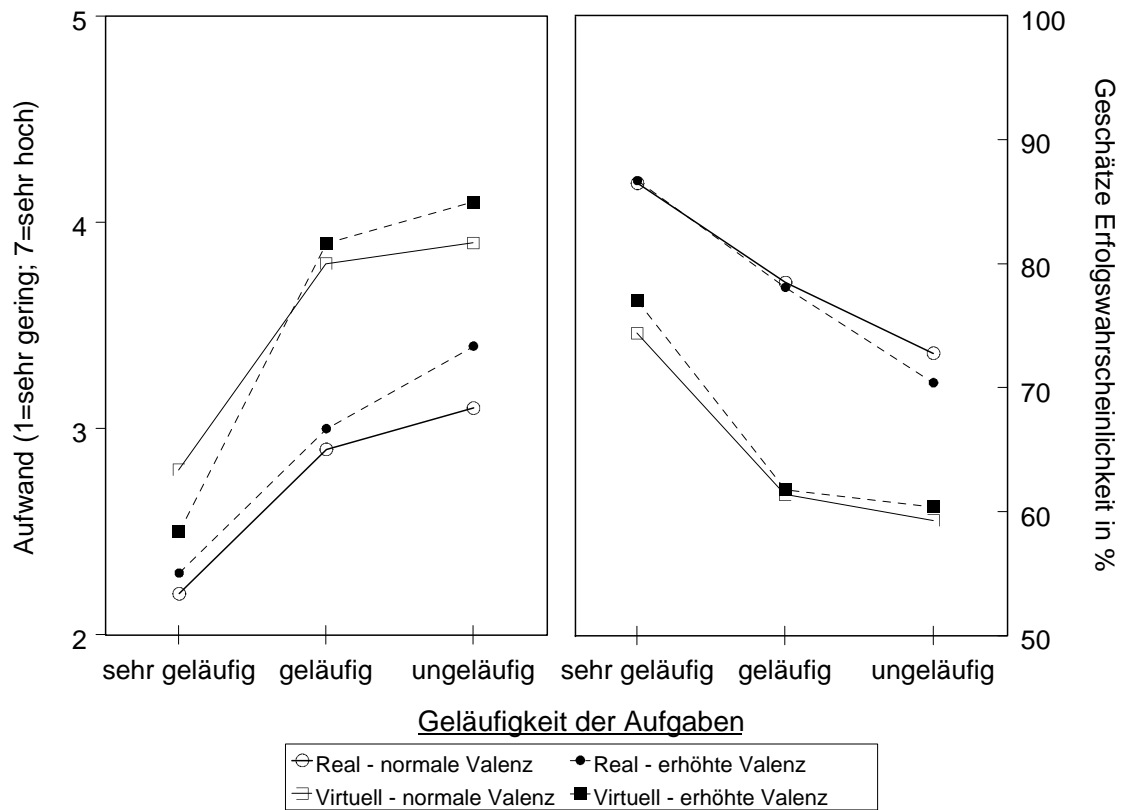


Abb. 15 Die Einschätzungen der Erfolgswahrscheinlichkeit und des Aufwandes desjenigen Speichers, der von den Probanden für die jeweilige Aufgabe gewählt wurde.

Für die geschätzte Erfolgswahrscheinlichkeit ergibt sich ein paralleles Bild. Hier sinkt die Erfolgswahrscheinlichkeit kontinuierlich über alle drei Geläufigkeitsstufen der Aufgabe bei realer Aufgabenstellung. Bei virtueller Aufgabenstellung sinkt die geschätzte Erfolgswahrscheinlichkeit nur von der sehr geläufigen zur geläufigen, nicht aber von der geläufigen zur ungeläufigen Aufgabengruppe. Auch hier bleibt die Valenz ohne Einfluß (siehe Abb. 15, rechter Teil).

Die Varianzanalyse mit Aufgabenstellung und Valenz als unabhängige Variable sowie Geläufigkeit der Aufgaben als Meßwiederholungsfaktor zeigt entsprechend gleiche Zusammenhänge für beide unabhängige Variablen. Beim geschätzten Aufwand einen signifikanten Haupteffekt für Aufgabenstellung, $F(1,76)=17.34, p<.001$ und Geläufigkeit, $F(2,152)=195.92, p<.001$ sowie eine signifikante Interaktion zwischen Geläufigkeit und Aufgabenstellung, $F(2,152)=9.7, p<.001$. Ebenso bei der geschätzten Erfolgswahrscheinlichkeit: ein signifikanter Haupteffekte für Aufgabenstellung,

$F(1,76)=22.2$, $P<.001$ und Geläufigkeit, $F(2,152)=102.9$, $p<.001$ sowie eine signifikante Interaktion zwischen Aufgabenstellung und Geläufigkeit, $F(2,152)=3.7$, $p<.05$. Die unabhängige Variable der Valenz sowie alle Interaktionen mit ihr zeigen in keiner der beiden Analysen einen signifikanten Einfluß.

T-Tests zur Klärung der Interaktion bestätigen die Beschreibung der Mittelwerte. Alle Mittelwerte zwischen verschiedenen Geläufigkeitsstufen der Aufgaben werden im Paarvergleich auf einem hohen Signifikanzniveau getrennt. Ausnahme davon: bei virtueller Aufgabenstellung unterscheiden sich geläufige und ungeläufige Aufgaben nicht. Beim Aufwand verringert sich hier das Signifikanzniveau; bei der Erfolgswahrscheinlichkeit ist kein Unterschied mehr nachweisbar (vgl. Tab. 8).

Aufwand und Erfolg (Mittelwertsvergleiche über t-Tests)

Vergleich zwischen Geläufigkeitsstufen (t-Tests für Meßwiederholungsvariablen)						
	geschätzter Aufwand			geschätzte Erfolgswahrscheinlichkeit		
<i>Aufgabenstellung:</i>	sehr geläufig vs. geläufig	geläufig vs. ungeläufig	sehr geläufig vs. ungeläufig	sehr geläufig vs. geläufig	geläufig vs. ungeläufig	sehr geläufig vs. ungeläufig
<i>real</i>	t(39)= -8.87***	t(39)= -8.99***	t(39)= -4.75***	t(39)= 6.8***	t(39)= 4.89***	t(39)= 7.35***
<i>virtuell</i>	t(39)= -11.74**	t(39)= -2.59*	t(39)= -12.71***	t(39)= 9.37***	t(39)= 1.59	t(39)= 8.41***
Vergleich zwischen beiden Aufgabenstellungen (t-Tests für unabhängige Variablen)						
	geschätzter Aufwand			geschätzte Erfolgswahrscheinlichkeit		
<i>Geläufigkeitsstufe:</i>	sehr geläufig	geläufig	ungeläufig	sehr geläufig	geläufig	ungeläufig
	t(78)= -2.28*	t(78)= -5.88***	t(78)= -3.71***	t(38)= 2.71**	t(38)= 2.55*	t(38)= 2.72**

Tab. 8 t-Tests zum Vergleich der Mittelwerte des von den Probanden eingeschätzten Aufwandes bzw. der eingeschätzten Erfolgswahrscheinlichkeit der Speicherung, für die sie sich entschieden hatten. * für $p<.05$; ** für $p<.01$; *** für $p<.001$

Während bei realer Aufgabenstellung also eine durchgängige Verschlechterung der Bewertung der Lösung erfolgt - d.h. der Erfolg wird geringer, der Aufwand höher eingeschätzt - werden bei der virtueller Aufgabenstellung nur die Lösungen der sehr geläufigen Aufgabenstellung als günstiger eingeschätzt als die der anderen beiden. Zwischen geläufigen und ungeläufigen Aufgaben verringert sich der Unterschied (beim Aufwand) bzw. verschwindet völlig (bei der Erfolgswahrscheinlichkeit).

Die Anzahl der je Aufgabengruppe zur Lösung herangezogenen unterschiedlicher Speicher

Die folgenden Analysen über die jeweils gewählten Speicher beziehen sich, wie oben beschrieben, nur auf die 40 Probanden in der realen Aufgabenstellung. Achtzig verschiedene Speicher wurden insgesamt eingesetzt. Die Einsatzhäufigkeit wies ein starkes Gefälle auf: drei Speicher wurden über 200 mal gewählt, zwölf Speicher zwischen 20 und 200 mal und 65 Speicher unter 20 mal. Insgesamt wurde die drei am häufigsten genannten Speicher bei 62% der Aufgaben eingesetzt, die zweite Gruppe bei 27% und die dritte bei 11%. Die gewählten Speicher werden im folgenden in zwei Richtungen analysiert: zum einen in Bezug auf die Anzahl unterschiedlicher Speicher pro Aufgabentyp und zum anderen in Bezug auf den Zusammenhang zwischen gewähltem Speicher und Planungszeit.

Die 60 Aufgaben sind unterteilt in drei Geläufigkeitsstufen und fünf Themen. Pro Geläufigkeitsstufe und Thema gibt es vier Aufgaben; folglich könnte jede dieser Teilgruppen mit maximal vier unterschiedlichen Speichern gelöst werden - wenn nämlich für jede Aufgabe ein anderer Speicher gewählt würde. Mit wie vielen unterschiedlichen Speichern diese Gruppen tatsächlich gelöst worden sind, ist in Abb. 16 dargestellt.

Für sehr geläufige Aufgaben sind deutlich weniger unterschiedliche Speicher eingesetzt worden

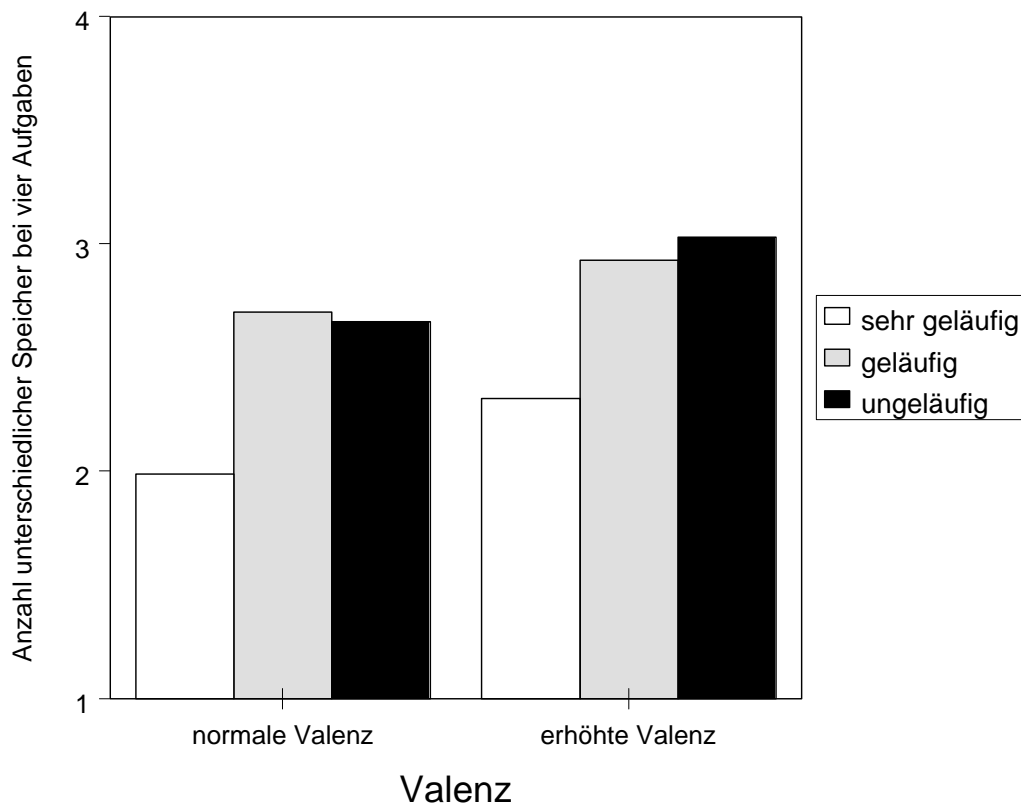


Abb. 16 Anzahl der pro Themengruppe der jeweiligen Aufgabe, Geläufigkeitsstufe und Valenz eingesetzter Speicher. Maximal wären vier unterschiedliche Speicher möglich.

als für geläufige und ungeläufige Aufgaben; zwischen den letzten beiden besteht kein Unterschied. Bei erhöhter Valenz werden mehr unterschiedliche Speicher eingesetzt; dies gilt gleichermaßen für alle drei Geläufigkeitsstufen. Eine Varianzanalyse mit der Geläufigkeit als Meßwiederholungs- und Valenz als unabhängiger Variable zeigt dementsprechende signifikante Haupteffekte, aber keine signifikante Interaktion: $F(1,39)=6.71$, $p=.01$, für Valenz und $F(2,78)=50.34$, $p<.001$ für Geläufigkeit. T-Tests belegen, daß der Haupteffekt der Geläufigkeit nur durch die deutlich geringere Speicherzahl bei sehr geläufigen Aufgaben verursacht wurde. Diese unterscheidet sich deutlich gegenüber geläufigen, $t(39)=-7.71$, $p<.001$, und ungeläufigen Aufgaben, $t(39)=-8.69$, $p<.001$; die Speicherzahl für geläufige und ungeläufige Aufgaben unterscheidet sich aber nicht $t(39)=-0.4$, $p>.05$.

Da die Aufgaben in fünf Themengruppen unterteilt sind, läßt sich der Effekt der Anzahl unterschiedlicher Speicher noch weiter für jeden Aufgabentyp differenzieren. Es zeigt sich, daß dieser Effekt tatsächlich von allen Aufgabengruppen getragen wird. T-Tests für abhängige Stichproben im Vergleich über die einzelnen Themengruppen ergeben durchweg sehr signifikante Unterschiede zwischen sehr geläufigen und geläufigen bzw. sehr geläufigen und ungeläufigen Aufgaben; für alle 20 t-Tests gilt $|t(40)| > 4.1$, $p<.001$. Bei keiner Themengruppe unterscheiden sich die Speicherzahl von geläufigen und ungeläufigen Aufgaben. Die zehn t-Tests liegen mit $|t(40)| < 1.1$ durchweg unter dem Signifikanzniveau. Eine MANOVA mit Aufgabengruppe und Itemgeläufigkeit als Meßwiederholungs- sowie Valenz als unabhängiger Variable ergibt drei signifikante Haupteffekte: für Valenz $F(1,38)=6.72$, $p<.05$, Aufgabengruppe, $F(4,152)=15.94$, $p<.001$ und Geläufigkeit $F(2,76)=46.98$, $p<.001$; aber keinerlei signifikante Interaktion. Die Aufgaben der Typen 'Informationsspeicher' werden mit den meisten unterschiedlichen Speichern gelöst (3.1), gefolgt von dem Typ 'Termine' (2.68). Die Speicherzahl dieser beiden Typen liegen deutlich über den der Typen 'wiederkehrende Aufgaben', 'zu erledigen' und 'einkaufen' (2.45, 2.43 und 2.33), die sich nicht mehr voneinander unterscheiden. T-Tests bestätigen dieses Bild.

Zusammenfassend gilt für alle Aufgabengruppen, daß bei großer Geläufigkeit weniger unterschiedliche Speicher eingesetzt werden als bei mittlerer und geringer Geläufigkeit der Aufgaben. Zwischen letzteren beiden unterscheidet sich die Anzahl unterschiedlicher Speicher nicht. Erhöhung der Valenz der Aufgaben führt generell zu einer Erhöhung der Speicherzahlen - unabhängig von der Geläufigkeit oder Gruppe der Aufgaben.

Die Planungszeit in Abhängigkeit von der Speicherwahl

Die folgenden Analysen betrachten die Planungszeit zur Wahl eines Speichers in Abhängigkeit von dem gewählten Speicher. In der ersten Analyse werden alle 80 genannten Speicher nach der

Häufigkeit, mit der sie genannt wurden, in drei Gruppen unterteilt und die Planungszeit danach differenziert, zu welcher Gruppe der gewählte Speicher gehört. In der zweiten Analyse wird die gleiche Differenzierung mit den drei am häufigsten genannten Speichern durchgeführt. Beide Analysen beziehen sich wieder nur auf die 40 Probanden der realen Aufgabenstellung. Da die Speicherwahl frei erfolgte, sind Datenverluste unvermeidlich.

Nach der Häufigkeit ihres Einsatzes lassen sich die Speicher in drei Gruppen unterteilen: häufig (über 200), regelmäßig (20 bis 200) und selten eingesetzte Speicher (unter 20; insgesamt gab es 2400 Wahlen). Für jede der drei Speichergruppen wurde, getrennt nach der Geläufigkeit der Aufgaben, die Planungszeit errechnet, die die Probanden brauchten, um sich für einen Speicher der jeweiligen Gruppe zu entscheiden. Nur bei 16 Probanden war jede der neun Zellen gefüllt; auf sie bezieht sich die nachfolgende Analyse. Die Mittelwerte sind dargestellt in Abb. 17.

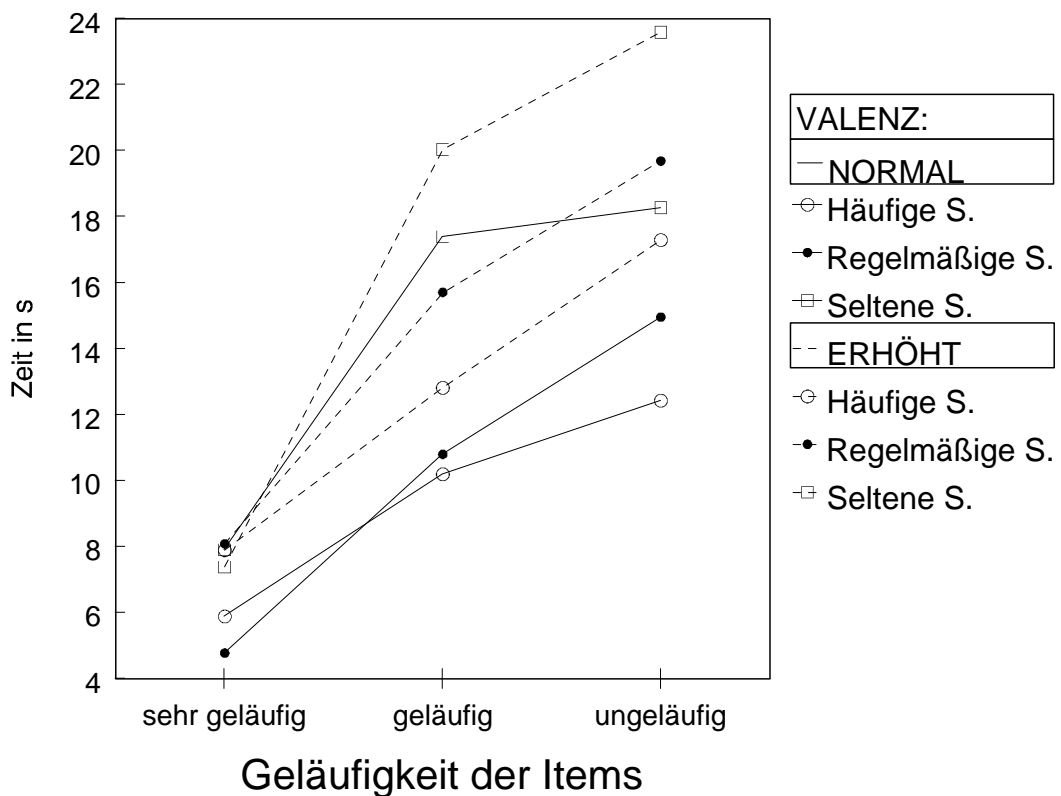


Abb. 17 Planungszeiten zur Speicherwahl, nach Häufigkeit des gewählten Speichers differenziert. Häufige Speicher=über 200, regelmäßige Speicher=20 bis 200, seltene Speicher=unter 20 von 2400 Wahlen. Die Mittelwerte beziehen sich auf die 16 Probanden, die in die Varianzanalyse aufgenommen wurden.

Die Varianzanalyse der Planungszeit zeigt zwei signifikante Haupteffekte: für die Speichergruppe, $F(2,28)=14.34$, $p<.001$ und die Geläufigkeit der Items, $F(2,28)=37.23$, $p<.001$. Die Valenz verfehlt nur knapp das Signifikanzniveau, $F(1,14)=4,36$; $p=.06$, signifikante Interaktionen treten nicht auf. Die Probanden brauchen bei allen drei Speichergruppen um so länger, einen Speicher zu wählen, je

ungeläufiger die Aufgabe. Auf der anderen Seite dauert die Wahl um so länger, je seltener der gewählte Speicher vom Probanden genutzt wird. Beide Effekte werden im wesentlichen von allen Paarvergleichen getragen; einzig der Unterschied zwischen Speichergruppe 1 und Speichergruppe 2 verpaßt mit $t(15)=-1.95$, $p=.08$, knapp das Signifikanzniveau (siehe Tab. 9).

Planungszeiten differenziert für häufig, regelmäßig und selten verwandte Speicher(Mittelwertsvergleiche über t-Tests)

Vergleich zwischen Speichergruppen und Geläufigkeitsstufen (t-Tests für Meßwiederholungsvariablen)						
	über alle Speichergruppen; Aufgabengruppe:			über alle Geläufigkeitsstufen; Speichergruppe:		
	sehr geläufig vs. geläufig	geläufig vs. ungeläufig	sehr geläufig vs. ungeläufig	häufig vs. regelmäßig	regelmäßig vs. selten	häufig vs. selten
	$t(15)=-6.55^{**}$	$t(15)=-2.12^*$	$t(15)=-11.13^{***}$	$t(15)=-1.95$	$t(15)=-3.11^{***}$	$t(15)=-5.41^{***}$

Tab. 9 t-Tests zum Vergleich der Mittelwerte der Anzahl der pro Aufgabenteilgruppe eingesetzter Speicher.

* für $p<.05$; ** für $p<.01$; *** für $p<.001$

Die drei am häufigsten eingesetzten Speicher sind das Gedächtnis, der Terminkalender und der Zettel an auffälligem Ort (769, 514 und 208 von insgesamt 2400 Nennungen). Die Planungszeiten zur Entscheidung für einen dieser drei Speicher sind in Abb. 18 dargestellt. Bei allen drei Speichern steigt die Planungszeit sowohl mit sinkender Geläufigkeit der Aufgaben als auch mit steigender Valenz. Da nur bei drei Probanden alle neun Zellen gefüllt sind, muß auf eine vollständige Analyse des Designs verzichtet werden. Stattdessen wird die ANOVA für jeden der drei Speicher getrennt gerechnet. So müssen für eine Analysen nur jeweils drei Zellen gefüllt sein und die Zahl auswertbarer Probanden steigt. Es ergibt sich eine hochsignifikante Wirkung der Aufgabengeläufigkeit: $F(2,62)=52.04$, $p<.001$ für Gedächtnis, $F(2,48)=17.68$, $p<.001$ für Terminkalender und $F(2,20)=19.7$, $p<.001$ für Zettel an auffälligem Ort. Weder die Valenz noch die Interaktion mit ihr erreicht das Signifikanzniveau. Die durchschnittliche Planungszeit über alle Geläufigkeitsstufen steigt von Gedächtnis (10s) über den Terminkalender (10.2s) bis zum Zettel am auffälligen Ort (12.3s). Kein t-Test zum Paarvergleich abhängiger Stichproben erreicht das Signifikanzniveau.

Auch eine Replikation dieser Analyse, bei statt die über alle Probanden am häufigsten eingesetzte Speicher für jede Versuchsperson individuell die drei von ihr am häufigsten eingesetzten Speicher Kriterium für die Auswertung der Planungszeit waren, verringert den Datenausfall kaum und verändert die Ergebnisse nur unwesentliche Veränderung. Auf ihre Darstellung wird deswegen an dieser Stelle verzichtet.

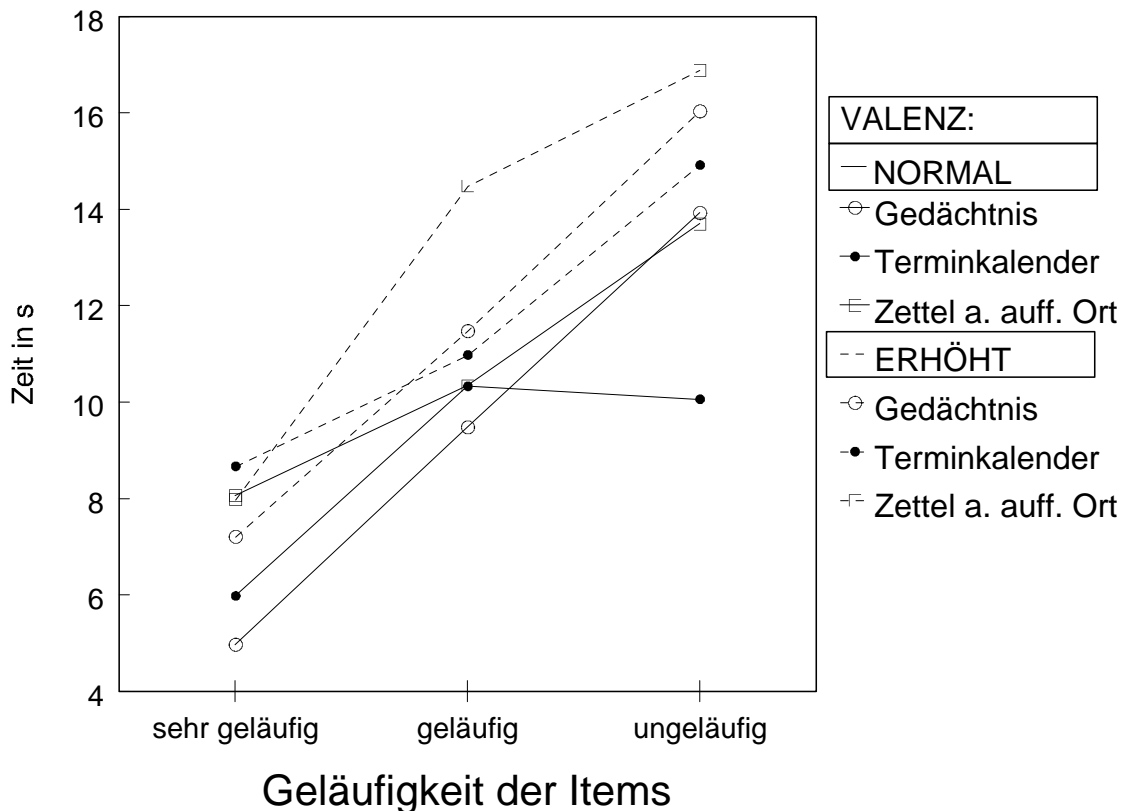


Abb. 18 Planungszeiten zur Speicherwahl, wenn einer der drei am häufigsten genannten Speicher gewählt wurde. Die Mittelwerte beziehen sich beim Gedächtnis auf 32, beim Terminkalender auf 25 und beim Zettel an auffälligem Ort auf 11 Probanden.

Zusammenfassend zeigen sich zwei Ergebnisse: Der Effekt der mit sinkender Geläufigkeit steigenden Planungszeit tritt bei allen Speichern bzw. Speichergruppen auf. Je seltener ein Speicher eingesetzt wird, um so länger ist die Planungszeit, die zu seiner Wahl benötigt wird. Dieser Effekt konnte nur bei den Speichergruppen varianzanalytisch überprüft werden. Die Tendenz zu längeren Planungszeiten bei erhöhter Valenz, besonders bei geringer Geläufigkeit der Items war zwar in den Mittelwerten vorhanden, konnte aber nicht statistisch abgesichert werden.

Diskussion

Die Ergebnisse werden zunächst in der Reihenfolge ihrer Präsentation diskutiert. Abschließend folgt eine Diskussion der unabhängigen Variablen über alle Ergebnisse hinweg.

Unterschiedliche Planungszeiten bei der Speicherwahl sind nach dem SWIEGS-Modell eine Funktion der Vorerfahrungen, die die Probanden mit Aufgaben ähnlichen Typs gemacht haben. Bezogen auf die Faktoren in Studie 4 sollten die Planungszeiten zum einen mit sinkender Aufgabengeläufigkeit steigen. Zum anderen schließt der virtuelle Kontext Vorerfahrungen mit den Aufgaben weitgehend aus.

Unterschiede in der Planungszeit zwischen verschiedenen Aufgabentypen müßten folglich bei dieser Aufgabenstellung verschwinden. Zum dritten sollte die Suche nach einem passenden Speicher bei Erhöhung der Valenz der Aufgaben in realem Kontext aufwendiger werden. Die Planungszeiten sollten sich bei erhöhter Valenz folglich verlängern (H7), und zwar um so mehr, je ungeläufiger die Aufgabe (H8).

Dieses durch das SWIEGS-Modell vorhergesagte komplexe Zusammenhangsmuster der Planungszeit mit den Faktorstufen wurde hochsignifikant bestätigt. Die Planungszeit steigt bei realer Aufgabenstellung mit sinkender Geläufigkeit hochsignifikant an. Das gilt sowohl für normale als auch für erhöhte Aufgabenvalenz. Bei erhöhter Aufgabenvalenz ist die Planungszeit generell signifikant höher als bei normaler Aufgabenvalenz. Dieser Unterschied ist für die ungeläufigen Aufgaben am größten. In der virtuellen Aufgabenstellung dagegen verschwinden die Unterschiede in der Planungszeit zwischen den verschiedenen Aufgaben weitgehend. Dabei bleibt noch ein kleiner, aber signifikanter Unterschied in der Planungszeit zwischen der sehr geläufigen und den beiden anderen Aufgabengruppen, der durch die unterschiedlichen Wortzahl und damit Verarbeitungszeit der einzelnen Aufgabengruppen plausibel wird (vgl. Studie 2).

Wie schon in der Diskussion zu Studie 2 vermutet, zeigt sich, daß es zu inhaltlich schwer interpretierbaren Ergebnissen führen würde, wenn man die Lesezeit der einzelnen Aufgaben von der Bearbeitungszeit subtrahierte, um die Planungszeit zu erhalten. Die Bearbeitungszeiten im virtuellen Kontext von Studie 4 lagen durchweg unter der für die einzelnen Aufgaben erhobenen Lesezeit (vgl. Abb. 10 in Studie 2). Eine Subtraktion würde folglich zu negativen Planungszeiten führen - einer psychologisch nicht sinnvoll interpretierbaren Größe. Probanden in der virtuellen Bedingung berichteten, die Texte hauptsächlich in Bezug auf die Wortzahl als Indikator der Schwierigkeit des Behaltens im Kopf und in Bezug auf Zahlen im Text als Indikator für ein einfaches Wiederfinden der Aufgabe im Karteikartenstapel beurteilt zu haben. Ein Verständnis des Inhalts der Aufgabe war folglich für die Speicherwahl nicht erforderlich (vgl. Diskussion von Studie 2). Bei realer Aufgabenstellung wirkte sich die Erhöhung der Valenz für die verschiedenen Geläufigkeitsstufen unterschiedlich auf die Planungszeit aus. Bei virtueller Aufgabenstellung dagegen führte erhöhte Valenz zu einer für alle Aufgaben gleich großen Verringerung der Planungszeit. Die Gleichmäßigkeit der Verringerung über alle drei Aufgabengruppen unterstützt die Hypothese, daß bei virtuellem Kontext die Speicherwahlen durch die Gleichschaltung der Vorerfahrungen für alle Aufgaben über den gleichen Prozeß liefen.

Daß die Planungszeit bei virtuellem Kontext durch Valenzerhöhung kürzer statt länger wird, widerspricht der Hypothese, daß Planungszeiten für ungeläufige Aufgaben durch Valenzerhöhung steigen (H8). In den postexperimentellen Interviews berichteten die Probanden in der virtuellen

Bedingung generell von Schwierigkeit, sich die Aufgabe zusammen mit der komplexen Speichernutzung vorzustellen. Die Operationalisierung der Valenzerhöhung durch einen fiktiven Wettkampf erhöht die Komplexität dessen, was vorzustellen ist, noch einmal. Eine Erklärung wäre folglich, daß die Probanden durch die Valenzerhöhung überfordert waren und die Bereitschaft, sich mit der Aufgabe intensiv auseinanderzusetzen, dadurch von vornherein gering war. Dies stimmt mit den Angaben einzelner Probanden im postexperimentellen Interview überein.

Ein hypothesenkritischer Teilaspekt der Ergebnisse zur Planungszeit ist der signifikante Unterschied zwischen normaler und erhöhter Valenz bei sehr geläufigen Aufgaben in realem Kontext. Wenn sehr geläufige Aufgaben über Scriptanwendung gelöst werden, sollte Valenzerhöhung hier nämlich keinerlei Einfluß haben. Diese Unregelmäßigkeit bei der Valenzerhöhung wird am Schluß der Diskussion im Kontext der anderen Ergebnisse noch einmal beleuchtet.

Die differenzierte Erhebung der metakognitiven Parameter Aufwand und Erfolgswahrscheinlichkeit für je zwei Speicheralternativen pro Aufgabe ermöglichte die Vorhersage, welche dieser Alternativen gewählt werden müßte, wenn der Proband der Logik seiner Parameter folgt. Überraschend war das Ergebnis zur Häufigkeit der verschiedenen Urteils kombinationen (vgl. Abb. 19). Über die Hälfte der Urteile war uneindeutig, d.h. daß der gewählte Speicher den Probanden sowohl weniger aufwendig als auch erfolgreicher schien. Die hohe Parallelität der Einschätzung von Aufwand und Erfolg fand sich auch in der Bewertung der eigenen Lösung wieder. Je ungeläufiger die Aufgabe, desto mehr sinkt die Einschätzung des Erfolges der Speicherung. Spiegelbildlich dazu steigt die Einschätzung des Aufwandes der Speicherung. Trotzdem ist es notwendig, beide Parameter zu erheben. Bei den eindeutigen Urteilen weist nämlich nur einer der beiden Parameter auf einen Speicher, während der andere unentschieden bleibt. Diese Kombination trat in 24 % der Fälle auf. Aufwand und Erfolg scheinen also in der Entscheidung für eine Speicherform in der Regel keine Antagonisten zu sein. Sie werden aber als Ergänzung eingesetzt: kann der Proband beispielsweise keinen Unterschied in der Erfolgswahrscheinlichkeit von verschiedenen Speichermöglichkeiten feststellen, so wird der weniger aufwendige Speicher gewählt und umgekehrt. Ein echter Konflikt zwischen Aufwand und Erfolgswahrscheinlichkeit - die zwar-aber Urteile - lag nur bei 13% der Aufgaben vor. In der virtuellen Bedingung ist der Anteil an zwar-aber Urteilen deutlich höher als in der realen Bedingung (18% vs. 8%). Akzeptiert man die Interpretation des engen Zusammenhangs zwischen Aufwand und Erfolg als ein generelles Muster, läßt sich dieser Unterschied als weiteres Indiz der diskutierten Überforderung der Probanden in der virtuellen Bedingung verstehen: die geringere Reliabilität der Einschätzung führt eher zu widersprüchlichen Aussagen.

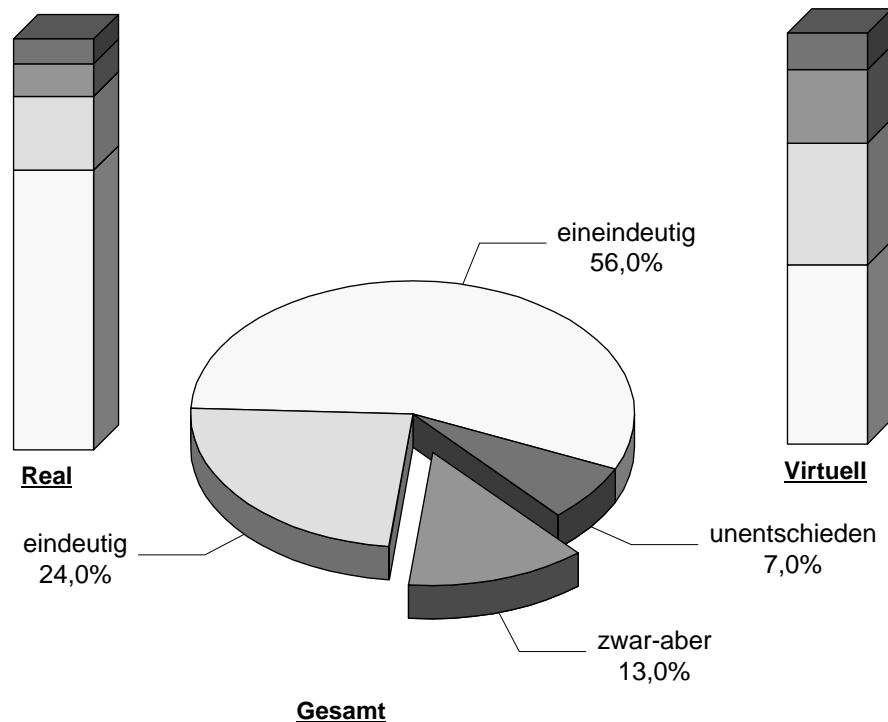


Abb. 19 Die Verteilung der Urteilkombinationen - getrennt nach den beiden Aufgabenstellungen 'virtuell' und 'real' sowie für beide Aufgabenstellungen gemeinsam. Insgesamt wurden 2400 Urteile abgegeben.

Die Geläufigkeit der Aufgabe sollte bestimmen, in welchem Maße die Vorhersage der Speicherwahl aus den Parametern möglich ist. Bei sehr geläufigen Aufgaben wird der Speicher nach dem Modell durch ein Script gewählt, also ohne Einfluß der metakognitiven Parameter. Diese leiten aber die Speicherwahl bei weniger geläufigen Aufgaben in den Prozessen der Evaluation bzw. Konstruktion (H5). Sind die Aufgaben durch den realen Kontext unterschiedlich geläufig, sollte die Vorhersage mit sinkender Geläufigkeit der Aufgabengruppe steigen. Im virtuellen Kontext, wo der Unterschied der Geläufigkeit zwischen den Aufgabengruppen verschwindet, sollte dagegen die Vorhersage der Speicherwahl für alle Aufgaben gleich gut möglich sein. Da durch den Kontext alle Aufgaben ungeläufig werden, sollte die Vorhersagbarkeit darüber hinaus hoch sein.

Diese Zusammenhänge wurden hochsignifikant bestätigt. Bei realer Aufgabenstellung stimmte der gewählte Speicher bei sehr geläufigen Aufgaben kaum über die Zufallswahrscheinlichkeit von 50% hinaus mit dem tatsächlich gewählten überein. Bei geläufigen Aufgaben dagegen stieg die Übereinstimmung auf über 70 % und bei ungeläufigen Aufgaben auf über 80%. Unerwartet war dabei, daß die Erhöhung der Valenz zu einem deutlichen Anstieg der Vorhersagbarkeit bei sehr geläufigen Aufgaben führte. Zwar erreichte weder der Haupteffekt der Valenz noch eine Interaktion mit ihr das

Signifikanzniveau. Trotzdem soll am Schluß der Diskussion noch einmal auf dieses Ergebnis eingegangen werden.

Die Vorhersagbarkeit ist bei der virtuellen Aufgabenstellung für alle Aufgabengruppen gleich groß - Beleg der Annahme, daß hier die Speicher alle über den gleichen Prozeß unter Einbeziehung der metakognitiven Parameter ablief. Trotzdem bleibt sie deutlich geringer als die der ungeläufigen Aufgaben bei realer Aufgabenstellung (68% vs. 83%). Nach dem SWIEGS-Modell müßten die Speicher in beiden Fällen durch den Prozeß der Konstruktion gewählt werden, da für beide Aufgabengruppen Vorerfahrungen in hohem Maße fehlen. Die Vorhersagbarkeit der Entscheidung für die virtuelle Aufgabenstellung müßte also der der ungeläufigen Aufgaben bei realer Aufgabenstellung entsprechen. Eine mögliche Erklärung der tatsächlichen Differenz zwischen beiden ist die schon diskutierte komplexe Aufgabenstellung in der virtuellen Bedingung. Wenn die Probanden Schwierigkeiten hatten, sich die virtuelle Aufgabenstellung und vor allem den komplexen Gebrauch des externen Speichers vorzustellen, sind die Parameter insgesamt wenig reliabel und die Vorhersageleistung sinkt.

Die Bewertung der eigenen Speicherwahl sollte hauptsächlich die Valenz der Aufgaben widerspiegeln. Je bedeutsamer eine Aufgabe für den Probanden, um so strengere Prüfkriterien für die Evaluation und Konstruktion eines Speichers sagt das SWIEGS-Modell vorher. Höhere Valenz einer Aufgabe sollte ja zu höheren Prüfparametern und damit zu höherer Einschätzung von Erfolg und Aufwand einer Speicherung führen (H10). Speicher, die mit diesen strengeren Kriterien gewählt wurden, sollten folglich auch erfolgreicher und aufwendiger eingeschätzt werden als die Speicher, die bei normaler Valenz gewählt werden. In den vorliegenden Daten veränderte die erhöhte Valenz der Aufgabe aber keineswegs die Einschätzung der gewählten Lösung.

Eine genauere Analyse der überraschenden Verteilungen der Urteilkombinationen zeigt allerdings, daß dieser Zusammenhang auch theoretisch nicht zu erwarten war. Tatsächlich bestand in 80% der Fälle für die Probanden kein Widerspruch zwischen Aufwand und Erfolg (vgl. Abb. 19). Der erfolgreichere Speicher war also entweder genauso oder gar weniger aufwendig als der weniger erfolgreiche. In diesen Fällen liegt aber kein Konflikt vor; der gewählte Speicher ist optimal in Bezug auf alle Kriterien. Bei einer optimalen Wahl aber können erhöhte Prüfparameter auch nicht mehr zu einer Verbesserung führen.

Nur bei den zwar-aber Entscheidungen besteht tatsächlich ein Konflikt zwischen Aufwand und Erfolg. Bei den 13% der zwar-aber Urteile könnte sich der Proband für den weniger aufwendigen oder aber für den erfolgreicheren Speicher entscheiden. Hier müßten eine erhöhte erforderlicher Erfolgswahrscheinlichkeit (EES) und ein erhöhter akzeptierter Aufwand (AAS) tatsächlich zu einer

anderen Wahl führen. Eine Analyse dieser Teilgruppe ergibt Mittelwertsunterschiede in der erwarteten Richtung: bei realer Aufgabenstellung führt erhöhte Valenz der Aufgaben sowohl zu einer höheren Einschätzung der Erfolgswahrscheinlichkeit als auch des Aufwandes der Speicherung. Bei virtueller Aufgabenstellung bleibt die Erfolgswahrscheinlichkeit unverändert, der Aufwand aber steigt (siehe Abb. 20).

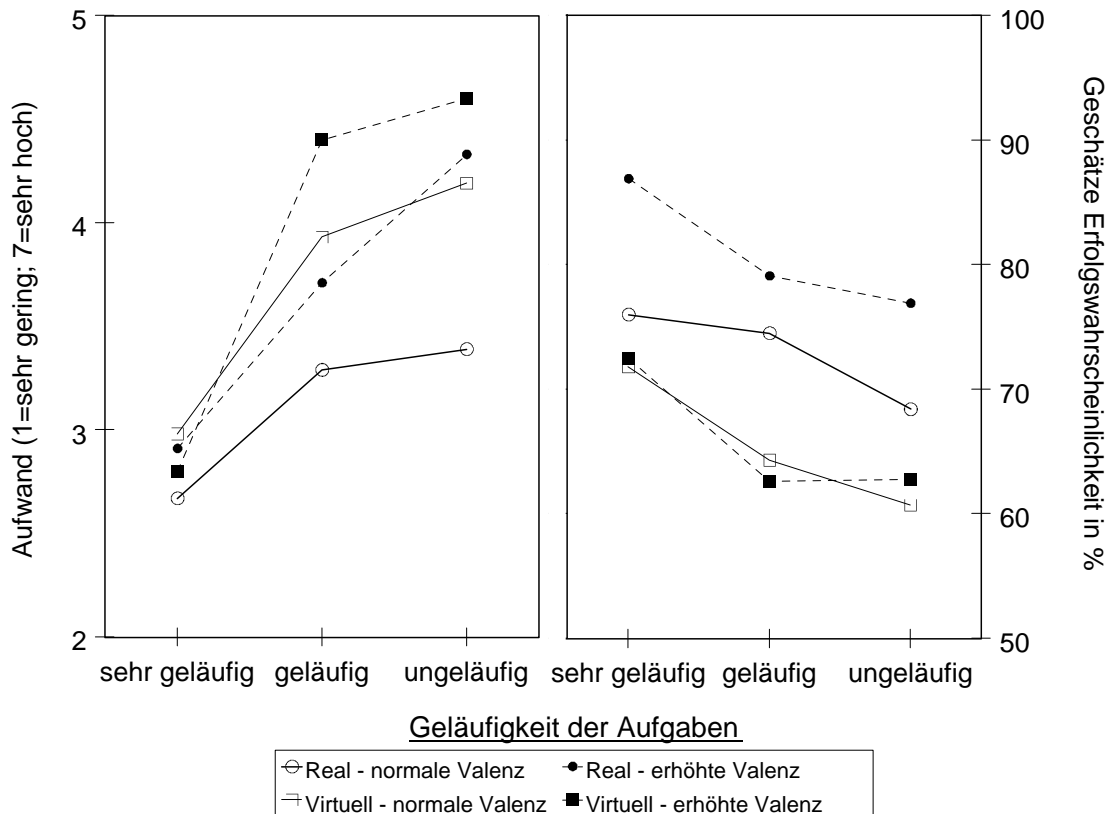


Abb. 20 Die Einschätzungen der Erfolgswahrscheinlichkeit und des Aufwandes desjenigen Speichers, der von den Probanden für die jeweilige Aufgabe gewählt wurde. Ausgewertet sind aber nur die Wahlen, bei denen ein Konflikt zwischen Aufwand und Erfolg bestand.

Bei realer Aufgabenstellung wirkt die erhöhte Valenz auch in der erwarteten differenzierten Form. Die Bewertung der gewählten Speicher ändert sich nur bei geläufigen und ungeläufigen Aufgaben; bei sehr geläufigen Aufgaben bleibt sie gleich - da der Speicher über das Script gewählt wird, haben veränderte Prüfparameter keinen Einfluß auf die Wahl bzw. die Bewertung dieser Wahl. Die ungleiche Besetzung der Zellen - 13 Versuchspersonen fallen aus der Analyse heraus, da sie keine 'zwar - aber' Kombination haben - und vor allem unterschiedliche Häufigkeit, mit der diese Kombination bei einer einzelnen Versuchsperson (zwischen 2 und 35 mal) und in beiden Aufgabenstellungen vorkommt (virtuell N=429, real N=181) verzerren die Ergebnisse, wenn man sie, wie für die varianzanalytische Auswertung notwendig, über die Probanden zusammenfaßt. Eine tatsächliche Überprüfung des vorhergesagten Zusammenhangs kann mit diesen Daten also nicht durchgeführt werden; es müßte dazu

ein Versuch konzipiert werden, der die 'zwar - aber' Kombination der metakognitiven Parameter erzwingt.

Bei einer Gruppe von Aufgaben ein und desselben Typs bestimmt die Geläufigkeit einer Teilgruppe dieser Aufgaben auch, mit wie vielen unterschiedlichen Speichern die Aufgaben repräsentiert werden. Betrachtet man beispielsweise die Teilgruppe der Termine, so sollten die sehr geläufigen Termine allesamt mit einem Terminalsript und dadurch mit dem gleichen Speicher repräsentiert werden. Werden für ungeläufigere Termine dagegen Speicher evaluiert oder konstruiert, so könnte für jede einzelne Aufgabe ein anderer Speicher zum Einsatz kommen. Ungeläufige Aufgaben eines bestimmten Typs sollten also mit mehr unterschiedlichen Speichern gelöst werden als sehr geläufige Aufgaben (H6). Diese Hypothese kann nur über die Teilgruppe mit realer Aufgabenstellung getestet werden, weil nur hier die Speicherwahl frei war. Der geforderte Zusammenhang war hochsignifikant vorhanden. Eine Interaktion mit dem Aufgabentyp trat nicht auf; tatsächlich ließ sich dieser Zusammenhang auch für jeden der fünf Aufgabentypen für sich gesehen hochsignifikant nachweisen. Kritische Größe für die Zahl unterschiedlicher Speicher war also alleine die Geläufigkeit, nicht der Typ der Aufgaben.

Im Zusammenhang mit der Valenz sollte sich dieses Muster weiter differenzieren. Erhöhte Valenz sollte bei evaluierten und konstruierten Speichern zu differenzierteren Lösungen führen, nicht aber bei Speicherwahlen über die Anwendung eines Scripts. Die Anzahl unterschiedlicher Speicher sollte also bei erhöhter Valenz größer sein als bei normaler Valenz (H9). Dieser Effekt sollte um so geringer sein, je geläufiger die Aufgabe (H10). Die Ergebnisse bestätigen Hypothese 9, stehen aber im Widerspruch zu Hypothese 10. Die Anzahl der eingesetzten Speicher steigt durch die erhöhte Valenz tatsächlich hochsignifikant an, allerdings ohne Unterschied bei den verschiedenen Geläufigkeitsstufen der Aufgaben. Eine Betrachtung nach den verschiedenen Aufgabentypen getrennt zeigt, daß dies nicht für alle Typen gilt. Bei wiederkehrenden Aufgaben führt erhöhte Valenz, wie von Hypothese 10 gefordert, nur bei geläufigen und ungeläufigen Aufgaben, nicht aber bei sehr geläufigen Aufgaben zu einer Steigerung der Speicherzahl. Die entsprechende Interaktion zwischen Aufgabentyp, Aufgabengeläufigkeit und Valenz erreicht das Signifikanzniveau aber nicht. Die entsprechende Hypothese konnte also mit der vorliegenden Untersuchung nicht belegt werden.

Eine letzte Analyseebene ist die Auswertung der Planungszeit in Abhängigkeit von dem gewählten Speicher. Drei Hypothesen sollten damit geprüft werden. Der Effekt steigender Planungszeit mit sinkender Geläufigkeit der Aufgaben sollte unabhängig vom gewählten Speicher sein (H2). Zwei Analysen wurden zur Klärung dieser Hypothesen gerechnet. Zum einen wurden alle eingesetzten Speicher nach der Häufigkeit ihrer Nutzung in drei Gruppen unterteilt, zum anderen die drei meistgenutzten Speicher betrachtet. In beiden Fällen wurde nun die Planungszeit für alle Kombinationen

von Aufgabengeläufigkeit und Valenz getrennt für den Einsatz jedes Speichers bzw. jeder Speichergruppe berechnet. Da immer nur bei einem kleinen Teil der Probanden alle Zellen gefüllt waren, waren beide Auswertungen mit dem Problem geringer Probandenzahlen behaftet.

Für die drei Speichergruppen wie auch für die drei einzelnen Speicher steigt die Planungszeit mit sinkender Geläufigkeit. Dieser Effekt war bei beiden Analysen trotz geringer Probandenzahlen hochsignifikant. Interaktionen zwischen Speicher bzw. Speichergruppe und Geläufigkeit traten dabei nicht auf; im Gegenteil zeigten Partialanalysen, daß dieser Effekt auch von jedem Speicher bzw. jeder Speichergruppe für sich gesehen getragen wurde. Wenn bei sinkender Aufgabengeläufigkeit die Planungszeiten steigen, so ist das kein Effekt davon, daß seltenere Speicher benötigt werden und es einfach länger dauert, bis diese den Probanden einfallen. Die Wahl des Gedächtnisses dauert so je nach Aufgabentyp zwischen 5s und 16s. Es muß also eindeutig der Prozeß, der zur Wahl führt, für die unterschiedliche Planungszeit verantwortlich sein.

Gleichzeitig sollte aber über alle Geläufigkeiten hinweg die Planungszeit für den Einsatz eines Speichers um so länger sein, je seltener er insgesamt eingesetzt wird (H3). Dieser Effekt sollte um so stärker sein, je ungeläufiger die Aufgabe (H4). Die Analyse der Planungszeiten über die Speichergruppen hinweg bestätigte Hypothese 3. Je häufiger ein Speicher eingesetzt wird, um so schneller wird er - bei gleicher Aufgabengeläufigkeit - gewählt. Dieser Effekt wurde zwar in den Mittelwerten mit sinkender Aufgabengeläufigkeit stärker; die erforderliche Interaktion erreichte das Signifikanzniveau aber nicht. Gleiches gilt für die Aufgabervalenz. Höhere Planungszeiten bei erhöhter Valenz (H3) und zwar um so mehr, je ungeläufiger die Aufgabe, wurde über das gesamte Design hochsignifikant bestätigt (s.o.). Bei der Differenzierung der Planungszeiten nach den Speicherwahlen waren beide Effekte in den Mittelwerten deutlich, erreichten aber nicht immer das Signifikanzniveau. Es überrascht nicht, daß die Analyse um so seltener zu einem signifikanten Ergebnis kommt, je mehr Faktoren in sie aufgenommen werden. Hinzu kommt, daß sich der Ausfall von Probanden für die Analyse über die Speicherwahl prinzipiell nicht vermeiden läßt, da die Speicherwahl frei erfolgen muß. Daß der erwartete Effekt in den Mittelwerten vorhanden war, ermutigt immerhin zu weiteren Studien, bei denen die Probandenzahl stark erhöht werden müßte, um die Ergebnisse auch inferenzstatistisch absichern zu können.

Über alle Analyseebenen hinweg verursachten die unabhängigen Variablen fast vollständig die geforderten Effekte. Sinkende Aufgabengeläufigkeit führte zu steigender Planungszeit, besserer Vorhersage der Speicherwahl aus den metakognitiven Parametern, höherer Speicherzahl und schlechterer Bewertung der gewählten Speicher. Die virtuelle Aufgabenstellung egalisierte für alle Aufgaben die Planungszeit und die Vorhersagbarkeit. Erhöhte Valenz schließlich führte zu längeren

Planungszeiten, besserer Bewertung der gewählten Speicher (zumindest bei den zwar-aber Urteilkombinationen, wo das überhaupt nur möglich war) und höherer Speicherzahl. Offen blieben zum Teil die nach Speicherzahlen differenzierten Analysen, wo aber die geringe Zahl analysierbarer Versuchspersonen eine plausible Erklärung für fehlende Signifikanzen lieferte.

Eine differenziertere Betrachtung zeigt allerdings einen Schwachpunkt. Die Valenzerhöhung wirkte sich fast durchgängig auch auf die sehr geläufigen Aufgaben aus, wo das nicht erwartet worden war. Werden die sehr geläufigen Aufgaben über Scriptanwendung gelöst, hat die Valenz der Aufgabe nach dem SWIEGS-Modell keinen Einfluß. Tatsächlich führte sie aber zu erhöhter Planungszeit, zur Erhöhung der Vorhersagbarkeit der Speicherwahl und bei vier der fünf Aufgabentypen zu einem Anstieg der Speicherzahl. Da die Wirkung der Valenz auf sehr geläufige Aufgaben aber der einzige Teilaspekt ist, wo Modellvorhersagen systematisch scheitern, liegt die Vermutung nahe, daß dieser Effekt mit der Operationalisierung der Valenz zusammenhängt. Es wäre denkbar, daß die hypothetische Wettkampfsituation die sehr geläufigen Aufgaben so veränderte, daß das entsprechende Repräsentationsscript nicht mehr paßte. Beispielsweise würde die Aufgabe, ein Buch in die Bibliothek zurückzubringen, zu einem anderen Aufgabentyp, wenn die erfolgreiche Erinnerung an diese Aufgabe über den Sieg in einem Wettkampf mitentscheidet. Um diese Überlegungen prüfen zu können, müßten weitere Untersuchungen mit andersartig operationalisierter Valenzerhöhung durchgeführt werden.

Am Rande sei angemerkt, daß die Probanden in der Gruppe mit erhöhter Valenz größtenteils berichteten, daß diese hypothetische Wettkampfsituation keinen Einfluß auf ihre Aufgabenbearbeitung gehabt habe. Dies bestätigt die im theoretischen Teil diskutierte mangelhafte Validität metakognitiver Urteile.