

**DIE ENTWICKLUNG DES
PROZEDURALEN METAGEDÄCHTNISSES
ÜBER DIE LEBENSSPANNE**

**Inaugural-Dissertation zur Erlangung der Doktorwürde
der Fakultät für Humanwissenschaften
der Julius-Maximilians-Universität Würzburg**

**vorgelegt von
Elisabeth Therese Löffler
aus Würzburg**

Würzburg, März 2016



Erstgutachter: Prof. Dr. Wolfgang Schneider

Zweitgutachterin: Prof. Dr. Gerhild Nieding

Tag des Kolloquiums: 18.11.2016

VORWORT

Die vorliegende Arbeit entstand im Rahmen des von der DFG geförderten Projekts „Die Entwicklung prozeduralen metakognitiven Wissens in der Lebensspanne“ unter der Leitung von Prof. Dr. Wolfgang Schneider und Dr. Nicole von der Linden.

Mein Dank gilt Prof. Dr. Schneider für die Bereitstellung des Themas und die umfassende Unterstützung in allen Phasen des Projekts und Dr. Nicole von der Linden für die engagierte und äußerst unkomplizierte Betreuung.

Bedanken möchte ich mich auch bei den Kolleginnen und Kollegen des Lehrstuhls für Psychologie IV, besonders bei Catharina Tibken, Wienke Wannagat und Michaela Pirkner, für die angenehme Arbeitsatmosphäre und die konstruktiven Anregungen.

Vielen Dank auch an die zahlreichen studentischen Hilfskräfte, Diplomanden, Bachelor- und Masterstudenten, die ebenfalls maßgeblich zum Gelingen des Projekts und somit auch dieser Arbeit beigetragen haben. Außerdem danke ich allen Schülerinnen, Schülern und Erwachsenen, die an den Studien teilgenommen haben, sowie den beteiligten Eltern, Lehrkräften und Schulleitern.

Schließlich bin ich sehr dankbar für die großartige Unterstützung durch meinen Mann Stefan, meine Familie und Schwiegerfamilie.

INHALTSVERZEICHNIS

0. EINLEITUNG	1
1. METAGEDÄCHTNIS UND METAKOGNITION: BEGRIFFSKLÄRUNG UND KONZEPTENTWICKLUNG	4
1.1 Begriffsklärung.....	4
1.2 Historische Entwicklung des Konzepts Metagedächtnis.....	5
1.2.1 Konzeptualisierung von Flavell und Wellman.....	5
1.2.2 Konzeptualisierung von Brown.....	7
1.2.3 Konzeptualisierung von Paris und Kollegen.....	8
1.2.4 Konzeptualisierung von Kluwe.....	9
1.2.5 Modell von Nelson und Narens.....	10
1.2.6 Modell des guten Informationsverarbeiters.....	12
1.2.7 Abschließende Bewertung.....	13
2. DEKLARATIVES METAGEDÄCHTNIS	18
2.1 Erfassungsmethoden.....	18
2.2 Entwicklung.....	20
2.3 Zusammenhang mit der Gedächtnisleistung.....	26
2.4 Fazit.....	27
3. PROZEDURALES METAGEDÄCHTNIS	28
3.1 Metakognitive Überwachung.....	28
3.1.1 Erfassungsmethoden.....	28
3.1.2 Ease-of-Learning-Urteile.....	31
3.1.2.1 Allgemeinpsychologische Befunde.....	31
3.1.2.2 Entwicklungspsychologische Befunde.....	34
3.1.2.3 Fazit.....	38
3.1.3 Judgments of Learning.....	39
3.1.3.1 Allgemeinpsychologische Befunde.....	39
3.1.3.2 Entwicklungspsychologische Befunde.....	46
3.1.3.3 Fazit.....	51
3.1.4 Feeling-of-Knowing-Urteile.....	52
3.1.4.1 Allgemeinpsychologische Befunde.....	52
3.1.4.2 Entwicklungspsychologische Befunde.....	53
3.1.4.3 Fazit.....	55
3.1.5 Sicherheitsurteile.....	55
3.1.5.1 Allgemeinpsychologische Befunde.....	55
3.1.5.2 Entwicklungspsychologische Befunde.....	59
3.1.5.3 Fazit.....	63
3.2 Lernzeitallokation als Beispiel für metakognitive Kontrollprozesse.....	64
3.2.1 Allgemeinpsychologische Befunde.....	65
3.2.2 Entwicklungspsychologische Befunde.....	67
3.2.3 Fazit.....	70
3.3 Zusammenhang zwischen Überwachungs- und Kontrollprozessen.....	71
3.3.1 „Monitoring-affects-control“-Modell.....	71

3.3.2	„Control-affects-monitoring“-Modell	73
3.3.3	Wechselseitige Zusammenhänge zwischen Überwachungs- und Kontrollprozessen	75
3.3.4	Fazit	76
3.4	Zusammenhang zwischen prozeduralem Metagedächtnis und Gedächtnisleistung	77
3.4.1	Überwachungsleistung und Gedächtnisleistung	77
3.4.2	Kontrollprozesse und Gedächtnisleistung	78
3.4.3	Fazit	79
4.	EXKURS: EXPERTISE	81
4.1	Begriffsklärung und Operationalisierung	81
4.2	Besonderheiten von Expertenwissen	82
4.3	Entwicklung von Expertise	82
4.4	Zusammenhang von Expertise und Metagedächtnisleistung	86
4.5	Fazit	88
5.	ZUSAMMENFASSUNG	89
6.	EMPIRISCHER TEIL	91
6.1	Einführung	91
6.2	Studie 1: Einfluss der Komplexität der Lernmaterialien auf die Überwachungsleistung	94
6.2.1	Herleitung der Fragestellung	94
6.2.2	Methodik	96
6.2.2.1	Design	96
6.2.2.2	Stichprobe	97
6.2.2.3	Versuchsmaterial	98
6.2.2.4	Studienablauf	100
6.2.2.5	Überblick über die statistischen Auswertungen	101
6.2.3	Ergebnisse	102
6.2.3.1	Erinnerungsleistung	102
6.2.3.2	Judgments of Learning	103
6.2.3.3	Sicherheitsurteile	104
6.2.4	Diskussion	106
6.3	Studie 2: Einfluss des Vorwissens auf die Überwachungsleistung beim Lernen von Paar-Assoziationen	110
6.3.1	Herleitung der Fragestellung	110
6.3.2	Methodik	112
6.3.2.1	Design	112
6.3.2.2	Stichprobe	112
6.3.2.3	Versuchsmaterial	113
6.3.2.4	Studienablauf	114
6.3.2.5	Überblick über die statistischen Auswertungen	115
6.3.3	Ergebnisse	116
6.3.3.1	Erinnerungsleistung	116
6.3.3.2	Ease-of-Learning-Urteile	116
6.3.3.3	Judgments of Learning	118
6.3.3.4	Sicherheitsurteile	119

6.3.4 Diskussion.....	120
6.4 Studie 3: Einfluss des Vorwissens auf die Überwachungsleistung beim Lernen aus einem Text.....	123
6.4.1 Herleitung der Fragestellung.....	123
6.4.2 Methodik.....	124
6.4.2.1 Design.....	124
6.4.2.2 Stichprobe.....	125
6.4.2.3 Versuchsmaterial.....	126
6.4.2.4 Studienablauf.....	127
6.4.2.5 Überblick über die statistischen Auswertungen.....	128
6.4.3 Ergebnisse.....	128
6.4.3.1 Erinnerungsleistung.....	128
6.4.3.2 Globales Verständnisurteil.....	129
6.4.3.3 Judgments of Learning.....	130
6.4.3.4 Sicherheitsurteile.....	132
6.4.4 Diskussion.....	134
6.5 Studie 4: Einfluss eines Strategietrainings auf die Überwachungsleistung.....	139
6.5.1 Herleitung der Fragestellung.....	139
6.5.2 Methodik.....	141
6.5.2.1 Design.....	141
6.5.2.2 Stichprobe.....	142
6.5.2.3 Versuchsmaterial.....	143
6.5.2.4 Studienablauf.....	144
6.5.2.5 Überblick über die statistischen Auswertungen.....	145
6.5.3 Ergebnisse.....	145
6.5.3.1 Erinnerungsleistung.....	145
6.5.3.2 Judgments of Learning.....	147
6.5.3.3 Sicherheitsurteile.....	148
6.5.3.4 Urteile zu „Recollection“ oder Vertrautheit (RFN-Urteile).....	150
6.5.3.5 Spontaner versus instruierter Strategiegebrauch.....	151
6.5.4 Diskussion.....	152
6.6 Studie 5: Sequenzielle Verknüpfung von Überwachungs- und Kontroll- prozessen.....	155
6.6.1 Herleitung der Fragestellung.....	155
6.6.2 Methodik.....	157
6.6.2.1 Design.....	157
6.6.2.2 Stichprobe.....	157
6.6.2.3 Versuchsmaterial.....	158
6.6.2.4 Studienablauf.....	159
6.6.2.5 Überblick über die statistischen Auswertungen.....	160
6.6.3 Ergebnisse.....	161
6.6.3.1 Lerndurchgang 1.....	161
6.6.3.1.1 Erinnerungsleistung.....	161
6.6.3.1.2 Judgments of Learning.....	161
6.6.3.1.3 Sicherheitsurteile.....	162

6.6.3.1.4	Judgments of Learning als Hinweisreize für die Lernanstrengung	163
6.6.3.2	Lerndurchgang 2	164
6.6.3.2.1	Erinnerungsleistung	164
6.6.3.2.2	Judgments of Learning	165
6.6.3.2.3	Sicherheitsurteile	165
6.6.3.2.4	Lernzeit als Hinweisreiz für die Judgments of Learning	166
6.6.3.2.5	Validität der Lernzeit als Prädiktor für die Erinnerungsleistung	167
6.6.4	Diskussion	168
6.7	Studie 6: Simultane Verknüpfung von Überwachungs- und Kontrollprozessen	172
6.7.1	Herleitung der Fragestellung	172
6.7.2	Methodik	174
6.7.2.1	Design	174
6.7.2.2	Stichprobe	174
6.7.2.3	Versuchsmaterial	175
6.7.2.4	Studienablauf	176
6.7.2.5	Überblick über die statistischen Auswertungen	176
6.7.3	Ergebnisse	177
6.7.3.1	Lerndurchgang 1	177
6.7.3.1.1	Erinnerungsleistung	177
6.7.3.1.2	Judgments of Learning	178
6.7.3.1.3	Sicherheitsurteile	178
6.7.3.1.4	Lernzeit als Hinweisreiz für die Judgments of Learning	179
6.7.3.1.5	Validität der Lernzeit als Prädiktor für die Erinnerungsleistung	180
6.7.3.2	Lerndurchgang 2	181
6.7.3.2.1	Erinnerungsleistung	181
6.7.3.2.2	Judgments of Learning	182
6.7.3.2.3	Sicherheitsurteile	183
6.7.3.2.4	Auswirkung der Anreizwerte auf die Höhe der Judgments of Learning und die Lernzeit	184
6.7.3.2.5	Lernzeit als Hinweisreiz für die Judgments of Learning	185
6.7.3.2.6	Validität der Lernzeit als Prädiktor für die Erinnerungsleistung	186
6.7.3.2.7	Fragebogen zum Lernanreiz	187
6.7.4	Diskussion	188
6.8	Zusammenfassung	192
7.	GESAMTDISKUSSION	194
7.1	Überwachungsleistungen im Entwicklungsverlauf	194
7.1.1	Globales Verständnisurteil	195
7.1.2	Ease-of-Learning-Urteile	195
7.1.3	Judgments of Learning	197
7.1.4	Sicherheitsurteile	199
7.1.5	Fazit	201

7.2 Einfluss von Kontextfaktoren auf die Überwachungsleistung.....	203
7.2.1 Einfluss der Komplexität des Lernmaterials.....	203
7.2.2 Einfluss der Modalität der Erinnerungsabfrage.....	205
7.2.3 Einfluss des Vorwissens.....	207
7.2.4 Einfluss der verwendeten Strategie.....	208
7.2.5 Fazit.....	209
7.3 Ziel- und Datenorientierung im Entwicklungsverlauf.....	210
7.4 Ausblick.....	212
8. ZUSAMMENFASSUNG DER ARBEIT.....	216
LITERATURVERZEICHNIS.....	218
ANHANG.....	240

0. EINLEITUNG

Hochrechnungen des Statistischen Bundesamtes gehen davon aus, dass im Jahr 2060 ein Drittel aller Bürger in Deutschland über 65 Jahre alt sein wird. Schon im Jahr 2015 traf dies auf etwa 20% der Bevölkerung zu (Statistisches Bundesamt, 2016). Der kontinuierliche Anstieg dieser Bevölkerungsgruppe macht es notwendig, sie vermehrt in wissenschaftliche Untersuchungen einzubeziehen. Um Prozesse des Alterns im Kontext des gesamten Entwicklungsverlaufs nachvollziehen zu können, ist es zudem wichtig, die Forschung zu psychologischen Vorgängen auf die Lebensspanne auszuweiten. Denn erst Vergleiche über größere Altersspannen hinweg geben Hinweise auf Unterschiede und Gemeinsamkeiten zwischen Aufbau- und Abbauprozessen, auf den Erhalt von psychischen Funktionen und auf Möglichkeiten zu Förderung und Kompensation.

Im Bereich der Gedächtnisentwicklung ist diese Forderung bereits häufig umgesetzt worden (z.B. Cowan, Naveh-Benjamin, Kilb & Sauls, 2006; Kliegel, Mackinlay & Jäger, 2008), sodass ein reiches Wissen darüber vorhanden ist, wie sich Gedächtnisleistungen im Lauf des Lebens verändern. So ist bekannt, dass schon Neugeborene über die Fähigkeit zum Wiedererkennen verfügen (z.B. Slater, Morison & Rose, 1982). Über die ersten Lebensmonate und -jahre vollzieht sich dann eine rasante Zunahme der Gedächtnisfähigkeiten: Die entsprechenden Inhalte werden zunehmend explizit und bewusst; die Gedächtnisspanne steigt an. Ab dem Vorschulalter werden Einspeicher- und Abrufvorgänge zunehmend von strategischem Verhalten begleitet und das semantische Netzwerk, in welchem Assoziationen zwischen verschiedenen Objekten und Konzepten organisiert sind, wird durch die stetig wachsende Wissensbasis immer reichhaltiger (für einen Überblick s. Schneider, 2015). Im jüngeren Erwachsenenalter wird in der Regel in den meisten Gedächtniskomponenten der Leistungshöhepunkt erreicht. Für das höhere Erwachsenenalter herrscht die Annahme vor, dass hier die Erinnerungsfähigkeit grundsätzlich ab- und die Vergesslichkeit zunimmt. Die Ergebnisse psychologischer Studien bestätigen dieses Bild nur zum Teil: Das höhere Erwachsenenalter stellt sich hier sehr heterogen und differenziert dar. Zum einen existieren große interindividuelle Unterschiede, zum anderen bleiben Fähigkeiten, die auf Vorwissen oder Routinen beruhen, noch lange auf einem recht stabilen Niveau. Von Abbauprozessen sind eher Gedächtnisfunktionen betroffen, die auf Geschwindigkeitsaspekten oder geringem Vorwissen beruhen (für einen Überblick s. Lindenberger & Staudinger, 2012).

Während also die Gedächtnisentwicklung in lebensspannenübergreifenden Paradigmen als recht gut erforscht gelten kann, sind solche Ansätze für den Bereich des Metagedächtnisses,

also dem Wissen über und der Überwachung und Steuerung von Gedächtnisprozessen (Dunlosky & Metcalfe, 2009), noch extrem selten. Dies ist umso bedeutsamer, als Metagedächtnisfertigkeiten als ein wichtiger Motor der Gedächtnisentwicklung gelten (Schneider, 2015). Zwar sind isolierte Altersgruppen bereits in einzelnen Aspekten gut untersucht worden, es fehlt jedoch an Studien, die größere Altersspannen vergleichen. Auch sind – in Analogie zu Arbeiten zur Gedächtnisentwicklung – Befunde notwendig, die den Einfluss von Kontextfaktoren auf metakognitive Leistungen analysieren. Denn ähnlich wie bei Gedächtnisprozessen ist anzunehmen, dass auch Metagedächtnisleistungen vom Abfrageformat (z.B. Wiedererkennen vs. freie Erinnerung), von der Art des Lernmaterials (z.B. Paarassoziationen vs. alltagsnähere Aufgaben), vom Einsatz von Strategien oder vom bereichsspezifischen Vorwissen in verschiedenen Altersstufen auf unterschiedliche Weise beeinflusst werden. Eine weitere Forschungslücke stellt die Frage dar, inwiefern sich verschiedene Aspekte des Metagedächtnisses über die Lebensspanne hinweg wechselseitig beeinflussen. Mit dem Fokus auf die prozedurale Komponente des Metagedächtnisses (d.h. auf Überwachungs- und Steuerungsvorgänge) sollen in der vorliegenden Arbeit Entwicklungsverläufe über eine Altersspanne von sieben bis 80 Jahren dargestellt werden. Die einzelnen Experimente sollen dabei den Einfluss der genannten Kontextfaktoren bzw. uni- und bidirektionale Zusammenhänge zwischen Überwachungs- und Steuerungsprozessen beleuchten.

Aufbau der Arbeit. Im theoretischen Teil der Arbeit wird zunächst der Begriff „Metagedächtnis“ definiert und die Entstehungsgeschichte dieses Konzepts sowie aktuelle Modellvorstellungen skizziert. Das zweite Kapitel beschäftigt sich mit Erfassungsmethoden und der Entwicklung des deklarativen Metagedächtnisses. Außerdem wird der Zusammenhang zu Gedächtnisleistungen aufgezeigt. Da diese Metagedächtnisfacette für die eigenen Studien von untergeordneter Bedeutung ist, beschränkt sich dieses Kapitel auf einen Überblick der wesentlichen Befunde. Im Zentrum der Arbeit steht das prozedurale Metagedächtnis. Dieses wird in Kapitel 3 ausführlich vorgestellt und diskutiert. Zunächst wird die Überwachungskomponente hinsichtlich der Erfassungsmethoden beleuchtet. Anschließend werden allgemein- und entwicklungspsychologische Befunde zu den einzelnen Überwachungsmaßen dargestellt. Hierbei wird besonders auf Einflussfaktoren auf die Überwachungsleistung eingegangen (z.B. Komplexität des Lernmaterials, Art der Erinnerungsabfrage) sowie mögliche theoretische Grundlagen für die verschiedenen Urteile betrachtet. Entwicklungsbedingte Veränderungen werden über die gesamte Lebensspanne hinweg beschrieben. Als Beispiel für Kontrollprozesse wird die Lernzeitallokation

herausgegriffen, da diese auch in den eigenen Studien erfasst wird, und ebenfalls bezüglich allgemein- und entwicklungspsychologischer Erkenntnisse vorgestellt. Anschließend werden beide Komponenten des prozeduralen Metagedächtnisses zusammengeführt und hinsichtlich wechselseitiger Beeinflussungen diskutiert. In einem Exkurs wird schließlich der Themenbereich Expertise und Vorwissen vorgestellt; hier werden schwerpunktmäßig Befunde zum Zusammenhang mit Überwachungsleistungen in verschiedenen Altersstufen berichtet. In einer Zusammenfassung werden Forschungslücken in den dargestellten Ergebnissen des Theorieteils aufgezeigt und Implikationen für die eigenen Studien abgeleitet.

Der empirische Teil der Arbeit ist im Wesentlichen nach den sechs Studien gegliedert. Nach einer Einführung, in der die Ziele der Untersuchungen sowie die Begründung für die einbezogenen Altersgruppen dargestellt werden, befassen sich die Kapitel 2 bis 7 mit der Durchführung und den Ergebnissen der eigenen Experimente. In Studie 1 wird der Einfluss der Komplexität der Aufgabe auf die Überwachungsleistungen untersucht. Die Studien 2 und 3 gehen der Frage nach dem Einfluss des bereichsspezifischen Vorwissens in der Domäne Fußball auf die Überwachungsfähigkeit nach. Studie 4 widmet sich dann dem Einfluss eines Strategietrainings auf die Überwachungsleistung und der Frage, inwiefern durch ein solches Training auch der Überwachung zugrunde liegende Prozesse verändert werden können. In den Studien 5 und 6 wird überprüft, inwiefern sich Überwachungs- und Kontrollvorgänge sowohl sequenziell als auch simultan gegenseitig beeinflussen. Alle Fragestellungen beziehen sich auf die Entwicklung der jeweiligen Aspekte über die Lebensspanne hinweg. In Kapitel 8 werden die Ergebnisse der eigenen Studien kurz zusammengefasst und in Kapitel 9 ausführlich diskutiert. Hierbei werden Entwicklungsverläufe in den einzelnen Überwachungsmaßen über alle Experimente hinweg betrachtet sowie der Einfluss von Kontextfaktoren auf die Überwachungsleistung diskutiert. Auch mögliche bidirektionale Zusammenhänge zwischen Überwachungs- und Kontrollprozessen werden noch einmal aufgegriffen. Zum Abschluss der Arbeit werden in einem Ausblick Anregungen für die weitere Forschung gegeben.

1. METAGEDÄCHTNIS UND METAKOGNITION:

BEGRIFFSKLÄRUNG UND KONZEPTENTWICKLUNG

Da sich das heutige Verständnis der Begriffe „Metagedächtnis“ und „Metakognition“ am besten aus der Geschichte der Konzepte ableiten lassen, beginnt der theoretische Teil dieser Arbeit nach einer kurzen Begriffsklärung mit einem Abriss über die chronologische Entwicklung der Termini.

1.1 Begriffsklärung

Ganz wörtlich sind „Metakognitionen“ zunächst „Kognitionen über Kognitionen“ (Flavell, Miller & Miller, 2002); sie stellen also eine Metaperspektive auf eigene oder auf die Denkvorgänge anderer Personen dar. Diese eher weit gefasste Definition des Begriffs schließt somit abstraktes Wissen über den Ablauf von Denkprozessen, über die Wirkungsweise von Strategien oder über andere Einflüsse auf die kognitive Leistung wie z.B. das Alter oder das Vorwissen der jeweiligen Person, aber auch Eigenschaften des Lernmaterials ein. Es sind jedoch auch Faktoren inbegriffen, die konkrete, aktuell ablaufende kognitive Prozesse betreffen, also unter anderem die Überwachung oder die Steuerung von Lern- und Erinnerungsvorgängen (vgl. Schneider, 2015). Metakognition als das Wissen über und das Eingreifen in Denkprozesse schließt demnach eine große Zahl einzelner Aspekte ein (Flavell, 1979). Als untergeordneter Terminus kann der zuerst von Flavell geprägte Begriff „Metagedächtnis“ (Flavell, 1971) gesehen werden; dieser umfasst im eigentlichen Wortsinn einen Teilbereich der Metakognition, nämlich das Wissen bzw. Kognitionen über Gedächtnisvorgänge. Die vorliegende Arbeit konzentriert sich auf die Forschung zum Metagedächtnis. Dennoch werden die Bezeichnungen Metakognition und Metagedächtnis im Folgenden, sofern dies nicht explizit erwähnt ist, synonym verwendet, da beide Begriffe inhaltlich sehr große Überschneidungen aufweisen – laut Flavell (1971) können Gedächtnisvorgänge als „angewandte Kognitionen“ gesehen werden – und dies in der einschlägigen Literatur ebenfalls üblich ist (z.B. Schneider & Lockl, 2006).

Detailliertere Untergliederungen und das aktuelle Verständnis der beiden Begriffe Metakognition und Metagedächtnis lassen sich anhand der historischen Entwicklung der Konzepte nachvollziehen.

1.2 Historische Entwicklung des Konzepts Metagedächtnis

Wie zuvor erwähnt, verwendete Flavell (1971) als erster den Begriff Metagedächtnis (*metamemory*), als er postulierte, dass man für das Verständnis der Gedächtnisentwicklung ganz grundsätzlich untersuchen müsse, welche Prozesse auf übergeordneter Ebene ablaufen, d.h. wie z.B. das intelligente Ablegen von Informationen im Gedächtnisspeicher, aber auch das intelligente Überwachen dieser Speichervorgänge erlernt wird. Diese erste Namensgebung führte sehr bald zu einer intensiven theoretischen (z.B. Brown, 1978; Flavell, 1979; Flavell & Wellmann, 1977) und empirischen (z.B. Flavell, Friedrichs & Hoyt, 1970; Kreuzer, Leonard & Flavell, 1975) Auseinandersetzung mit metakognitiven Prozessen. Das zunächst recht allgemein gefasste Konzept wurde in Folge weiter differenziert und unterlief verschiedenen Modellentwicklungen.

1.2.1 Konzeptualisierung von Flavell und Wellman

Flavell und Wellman (1977) ordneten das Metagedächtnis als eine von vier Facetten des Gedächtnisses ein. Sie postulierten als erste Kategorie grundlegende Gedächtnisvorgänge wie Wiedererkennen und Repräsentieren (*basic operations*). Im Weiteren orientierten sich Flavell und Wellman (1977) an den Begrifflichkeiten von Brown (1975) und sahen Wissen (*knowing*), d.h. semantische und konzeptuelle Gedächtnisinhalte und deren Verbindungen zueinander, Strategien (*knowing how to know*) und das Metagedächtnis (*knowing about knowing*) als weitere wesentliche Bestandteile des Gedächtnisses an; letzteres stellte in ihren Augen einen maßgeblichen Motor der Gedächtnisentwicklung dar.

Den Autoren zufolge lässt sich das Metagedächtnis wiederum in zwei Aspekte untergliedern. Zum einen ist die *Sensitivitätskategorie* zu nennen; diese beinhaltet das Wissen einer Person darüber, welche Situationen den willkürlichen und planvollen Einsatz von Gedächtnisprozessen verlangen und welche nicht. An dieser Stelle unterschieden Flavell und Wellman (1977) zwischen Gedächtnisaktivitäten, die von außen initiiert werden, wie z.B. der Aufforderung, sich etwas einzuprägen, und spontanen Anstrengungen, sich an bestimmte Inhalte zu erinnern. Die Autoren gingen davon aus, dass sich Letztere deutlich später entwickeln und es ein generalisiertes Produktionsdefizit bei jüngeren Kindern geben muss, welches ganz allgemein Defizite im bewussten Vorbereiten auf zukünftige Abrufprozesse beinhaltet und nicht nur den fehlenden spontanen Gebrauch von Strategien einschließt. Flavell und Wellman (1977) betonten, dass ein Großteil des Lernens gerade bei Kindern unwillkürlich stattfindet; sie sahen

jedoch in der sich entwickelnden Sensitivität dafür, wann intentionales Lernen notwendig ist, einen maßgeblichen Faktor für die Verbesserung von Lern- und Erinnerungsprozessen. Neben der Sensitivitätskategorie postulierten Flavell und Wellman (1977) die *Variablenkategorie*, die sich als deutlich einflussreicher für die weitere Konzeptbildung in der Metagedächtnisforschung herausstellte. Damit ist das potenziell verbalisierbare Wissen über Faktoren gemeint, die die Gedächtnisleistung einer Person in einer bestimmten Situation beeinflussen. Die *Personenvariable* umfasst dabei vorübergehende und dauerhafte Merkmale von Personen, die sich auf Gedächtnisleistungen auswirken. Hierunter fassten die Autoren zunächst ein sog. *mnemonisches Selbstkonzept*, also Annahmen über das eigene Gedächtnis wie z.B. das Wissen „Ich kann mir schlecht Namen merken“. Im Artikel von 1977 wurden unter der Personenvariable allerdings auch State-Merkmale aufgeführt, d.h. personenbezogene Gedächtniserfahrungen, die im Hier und Jetzt stattfinden wie z.B. das Gefühl „Es liegt mir auf der Zunge“ (*tip of the tongue-feeling*) und ganz allgemein das Überwachen von Lern- und Erinnerungsvorgängen. Später wurde die Personenvariable von Flavell genauer spezifiziert: Er unterschied zwischen einer intrapersonellen Differenzierung bezüglich des Wissens über personenbezogene Eigenschaften (z.B. „Sarah kann sich besser Gedichte merken als mathematische Formeln“), einer interpersonellen Differenzierung (z.B. „Sarah hat ein besseres Gedächtnis als Franz“) und *universals of cognition* (Flavell, 1979, S. 907). Letztere entsprechen in etwa den State-Merkmalen von Flavell und Wellman (1977), denn sie umfassen Einsichten, wie gut Inhalte gelernt und verstanden wurden bzw. zu einem späteren Zeitpunkt abrufbar sind. Einen weiteren Faktor innerhalb der Variablenkategorie stellt die *Aufgabenvariable* dar. Diese beinhaltet das Wissen darüber, welche Merkmale einer Aufgabe das Lernen und Erinnern erleichtern oder erschweren. Dabei kann es sich um Wissen handeln, das die Struktur (z.B. „Zusammenhängende Informationen sind leichter zu merken als unzusammenhängende“) oder die Anforderungen (z.B. „Das wörtliche Erinnern eines Textes ist schwieriger als das Wiedergeben des Inhalts“) einer Aufgabe betrifft. Es ist ebenfalls das Wissen darüber eingeschlossen, welche Aufgabenmerkmale irrelevant für das Erinnern sind (z.B. „Grün und blau gedruckte Wörter sind gleich schwer zu lernen“). Schließlich enthielt das Konzept von Flavell und Wellman (1977) die *Strategievariable*, d.h. das Wissen darüber, welche Strategien den Erinnerungsvorgang erleichtern können. Dies können sowohl interne (z.B. Wiederholung) als auch externe (z.B. einen wichtigen Termin im Kalender notieren) Strategien sein. Die Autoren hoben hervor, dass eine gute Metagedächtnisleistung immer auch mit dem Wissen um die Interaktion der genannten Variablenkategorien einhergeht. So kann sich eine Person mit großem Vorwissen im Bereich Fußball wahrscheinlich die Ergebnisse des letzten Bundesliga-

Spieltages besser merken als ein Fußball-Laie (Interaktion von Person und Aufgabe). Personen mit einem differenzierten metakognitiven Wissen können auch berücksichtigen, dass bestimmte Personen bestimmte Strategien nur bei bestimmten Aufgaben gewinnbringend einsetzen können (Interaktion von Person, Strategie und Aufgabe).

Später veröffentlichte John Flavell eine weitere Ausdifferenzierung seines Metagedächtniskonzepts: Das *Model of Cognitive Monitoring* umfasste den Bereich *metacognitive knowledge* (Flavell, 1979), worunter die von Flavell und Wellman (1977) genannten Kategorien Personen-, Aufgaben- und Strategievariable fallen. Den zweiten Bereich stellten *metacognitive experiences*, also bewusste kognitive oder affektive Erfahrungen, die intellektuelle Prozesse begleiten, dar. Diese können als eine Art „Standortbestimmung“ des Individuums gesehen werden und dazu führen, dass Ziele, Strategien oder auch der Inhalt des metakognitiven Wissens angepasst werden. *Goals (or tasks)* und *actions (or strategies)* bildeten laut Flavell (1979) zwei weitere Bereiche des Modells. Die metakognitive Überwachung (*monitoring*) besteht diesem Modell zufolge im Zusammenspiel und der wechselseitigen Beeinflussung dieser vier Bereiche.

Es lässt sich also erkennen, dass die Kategorie „metakognitives Wissen“ bzw. die „Variablenkategorie“ von Flavell und seinem Kollegen Wellman schon früh recht eindeutig definiert wurde und verschiedene Modellbildungen überdauerte (Flavell, 1979; Flavell & Wellman, 1977). Andere Konzepte waren weniger genau abgegrenzt; so teilen beispielsweise die „Sensitivitätskategorie“ und der Terminus „metakognitive Erfahrungen“ einige inhaltliche Aspekte, denn beide beziehen sich auf die Fähigkeit des menschlichen Intellekts, bei der Bearbeitung von Lern- oder Denkaufgaben sensibel für deren kognitive und metakognitive Anforderungen zu sein und darauf zu reagieren. Gleichzeitig blieben Flavell und Wellman (1977) eine spezifische Definition dieser Metagedächtniskonponenten schuldig. Es bleibt ferner unklar, wie genau die verschiedenen Aspekte der Modelle zusammenwirken, damit ein effektives Überwachen von Gedächtnisprozessen möglich wird.

1.2.2 Konzeptualisierung von Brown

In der chronologischen Abfolge sind die Konzepte von Flavell und Wellman (1977) auf der einen und von Ann Brown (1978) auf der anderen Seite schwer zu trennen. Dennoch können die Arbeiten von Brown inhaltlich als eine Weiterentwicklung des oben beschriebenen Ansatzes gelten.

Brown (1978) kritisierte an den Forschungsarbeiten der 70er-Jahre, dass häufig nur einzelne Komponenten isoliert betrachten würden und dadurch z.B. die Interaktion von Person-, Aufgaben- und Strategievariable außer Acht gelassen werde. Studien, die solchermaßen alleinstehende Metagedächtnisaspekte untersuchten, unterschätzten die Komplexität von metakognitiven Prozessen. Laut Brown kann die Introspektion eigener Lern- und Erinnerungsvorgänge unterschiedlich anspruchsvoll ausfallen und bedarf exekutiver Funktionen, um gewinnbringend abzulaufen. Sie griff deshalb auf Flavell und Wellmans (1977) Ansatz die im Hier und Jetzt stattfindenden, personenbezogenen *Gedächtniserfahrungen* auf, also die Beurteilung und Überwachung des eigenen, aktuellen Lernstandes, und betonte die Anforderungen, die diese an die Exekutive stellt. Zu diesen gehören das Wissen um Kapazitätsgrenzen, das Wissen um Routinen und deren sinnvolle Anwendung, die Fähigkeit, aktuelle Probleme zu erfassen, geeignete Problemlösestrategien anzuwenden und deren Effektivität zu überwachen sowie schließlich die Bewertung von Erfolg oder Misserfolg kognitiver Prozesse, um auch die Beendigung von strategischem Vorgehen zu planen. Im Verständnis von Brown (1978) ist die Bedingung für diese Auseinandersetzung mit geistigen Prozessen die Überwindung der *secondary ignorance*, also die Unterscheidung zwischen Inhalten, die ein Lerner versteht und im Gedächtnis gespeichert hat, und solchen, auf die dies nicht zutrifft. Auf der Grundlage dieses Bewusstseins kann ein Individuum dann beurteilen, wie schwierig Aufgaben sind oder wie sicher es sich über die Richtigkeit einer gegebenen Antwort ist; es kann Vorhersagen über künftige Erinnerungsleistungen treffen, die potenzielle Wirkung des Einsatzes von Strategien beurteilen, seine eigenen kognitiven Ressourcen sinnvoll einteilen oder auch Inkonsistenzen in Aufgaben aufdecken, um nur einige Beispiele zu nennen (Brown, 1978). Brown (1978) fokussierte sich also weniger auf Aspekte, die das Wissen um Gedächtnisprozesse betreffen, sondern auf solche, die die Überwachung und Steuerung beinhalten. Beide Facetten wurden von ihr auch in der Nomenklatur deutlich getrennt in eine *deklarative* und eine *prozedurale* Komponente (Brown, 1980; Brown, Bransford, Ferrara & Campione, 1983).

1.2.3 Konzeptualisierung von Paris und Kollegen

Die Arbeiten von Paris und Kollegen (z.B. Paris, Lipson & Wixson, 1983) zur Taxonomie metakognitiver Prozesse sind insofern erwähnenswert, als sie einen Beitrag dazu leisten, die Ansätze von Flavell und Brown konzeptuell zusammenzuführen. Die Autoren griffen Browns Trennung in eine deklarative und eine prozedurale Komponente auf (s.o.), spezifizierten beide

jedoch genauer und erweiterten sie um eine *konditionale Komponente*. Unter dem *deklarativen (Metagedächtnis-)Wissen* werden Inhalte zusammengefasst, die die Aufgabenstruktur oder die Lernziele betreffen. Es geht dabei auch um Überzeugungen bezüglich eigener kognitiver Fähigkeiten. Paris et al. (1983) nahmen hier also Bezug auf jene Aspekte, die Flavell und Wellman (1977) mit der Variablenkategorie bezeichneten. Das *prozedurale (Metagedächtnis-)Wissen* entspricht dem, was Brown (1978) mit der Überwachung und Steuerung von Lernprozessen im Hier und Jetzt meinte. Paris und Kollegen (1983) bezogen sich in ihrer Taxonomie auf einschlägige Gedächtnismodelle, die ebenfalls eine prozedurale und einen deklarative Komponente beinhalten (z.B. Resnick, 1983). Während sich prozedurales Wissen in Gedächtnismodellen auf das Ausführen von Handlungen bezieht (z.B. Französisch sprechen können), geht es im Metagedächtnismodell von Paris et al. (1983) um Handlungen zweiter Ordnung, d.h. um das Verstehen der Handlung. Wenn also beispielsweise ein Schüler als wirkungsvolle Strategie gelernt hat, beim Lesen von Texten wichtige Wörter zu unterstreichen (= prozedurales Wissen), hat er unter Umständen auch ein Verständnis dafür, wie das Unterstreichen wirkt und warum es sinnvoll ist (= prozedurales Metagedächtniswissen). Schließlich erweiterten Paris et al. (1983) die vorherrschenden Modelle um das *konditionale (Metagedächtnis-)Wissen*. Dieses beinhaltet das Wissen darüber, wann und warum welches Vorgehen bei Lern- und Gedächtnisaufgaben effektiv ist. Hier werden also deklarative und prozedurale Aspekte miteinander in Verbindung gesetzt, denn ein gut ausgeprägtes konditionales Metagedächtniswissen unterstützt dabei, beide Facetten, also z.B. das Wissen darüber, dass man noch nicht genug über das Thema eines Textes weiß (deklaratives Metagedächtniswissen), und das Wissen darüber, wie Unterstreichen einem Leser hilft, sich Informationen aus Texten einzuprägen (prozedurales Metagedächtniswissen), in der konkreten Situation zielführend anzuwenden.

1.2.4 Konzeptualisierung von Kluwe

Kluwe (1982; Kluwe & Schiebler, 1984) bezog sich in seinen Veröffentlichungen zwar nicht direkt auf die Arbeiten von Paris und Kollegen, nutzte aber auch die beiden Begriffe „deklarativ“ und „prozedural“, um das metakognitive Wissen nach Flavell (1979) und exekutive Prozesse nach Brown (1978) voneinander abzugrenzen. Das deklarative Metagedächtnis wurde von ihm recht vergleichbar mit anderen Arbeiten (Flavell & Wellman, 1977; Paris et al., 1983) als Wissen über eigene Gedächtnisinhalte und die anderer Personen definiert (Kluwe, 1982). Der Fokus lag bei Kluwe (1982; Kluwe & Schiebler, 1984) jedoch auf

dem prozeduralen Metagedächtnis; hier stellten seine Arbeiten eine wesentliche Weiterentwicklung und Spezifizierung des Konzepts dar. Denn Kluwe (1982) differenzierte die exekutiven Prozesse, also das prozedurale Metagedächtnis, in *Kontrollprozesse (monitoring processes)*¹, d.h. solche Aktivitäten, bei denen ein Individuum Informationen über den aktuellen Stand der Informationsverarbeitung ableitet, und in *Regulationsprozesse*, d.h. dem Eingreifen in ablaufende Lernvorgänge durch die Steuerung von Kapazität, Intensität oder Geschwindigkeit des Lernens. Neben den exekutiven Prozessen postulierte Kluwe (1982) sog. *Lösungsprozesse*, bei denen die Informationen und Ergebnisse aus Kontroll- und Steuerungsprozessen angewandt werden, um ein Problem zielgerichtet zu bearbeiten.

1.2.5 Modell von Nelson und Narens

Es ist den bisher vorgestellten Konzeptualisierungen zu entnehmen, dass das Verständnis des Metagedächtnisses seit den 70er-Jahren sehr raschen Entwicklungen unterlegen war. Gemeinsam ist den bisher vorgestellten Konzepten die zunehmende Trennung zwischen deklarativen und prozeduralen Aspekten. Während sich die Einteilung des deklarativen Metagedächtnisses in Personen-, Aufgaben- und Strategiewissen seit Flavell und Wellman (1977) bzw. Flavell (1979) im Wesentlichen durchgesetzt hatte, waren die prozeduralen Aspekte deutlich weniger einheitlich gefasst.

Hier war es das Verdienst von Nelson und Narens (1990, 1994), bisher weitgehend unabhängig voneinander betrachtete Aspekte des prozeduralen Metagedächtnisses zusammenzuführen und eine fundierte theoretische Begründung zu liefern. Ihre Modellentwicklung basiert auf drei Prinzipien (Nelson & Narens, 1990): Die Autoren postulierten erstens, dass kognitive Prozesse in mindestens zwei, sich wechselseitig beeinflussende Ebenen aufgeteilt werden können; dies sind die *Metaebene* und die *Objektebene*. Zweitens beinhaltet die Metaebene laut Nelson und Narens (1990) ein dynamisches Modell, also eine innere Repräsentation der Vorgänge auf Objektebene. Drittens gingen die Autoren von zwei vorherrschenden Wegen aus, wie die Information zwischen den beiden Ebenen fließt: Zum einen wird Information aus der Metaebene an die Objektebene weitergegeben; hierbei handelt es sich also um Kontroll- oder Steuerungsprozesse (*control*). Das heißt, die kognitive Handlung wird von der Metaebene aus

¹ In der eigenen Arbeit werden *monitoring processes* mit *Überwachungsprozesse* übersetzt; *Regulationsprozesse* werden *Kontroll- bzw. Steuerungsprozesse* genannt (vgl. Nelson & Narens, 1990, 1994).

beeinflusst. Zum anderen fließen Informationen von der Objektebene zur Metaebene; es handelt sich hier folglich um Informationen, die der Überwachung dienen (*monitoring*; vgl. Abb. 1).

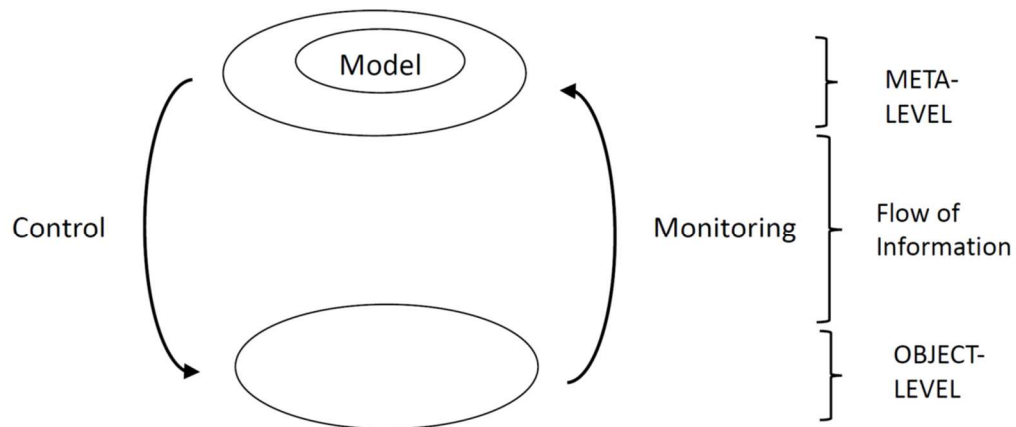


Abbildung 1: Meta- und Objektebene nach Nelson und Narens (1990)

Beide Prozesse wurden von Nelson und Narens (1990, 1994) weiter spezifiziert; ihnen wurden Maße zugeordnet, die auch vorher schon Verwendung fanden, aber im hier beschriebenen Modell genauer definiert und an bestimmten Zeitpunkten im Enkodier-, Speicher- und Abrufprozess lokalisiert wurden (s. Abb. 2). Vor dem Lernen legt ein Individuum laut Nelson und Narens (1990) die eigenen Lernziele (*norm of study*) fest und macht sich einen Plan, wie diese zu erreichen sind. An dieser Stelle kommen die verschiedenen Formen der Überwachung ins Spiel. Hier können grundsätzlich prospektive (d.h. vor dem Abruf stattfindende) und retrospektive (d.h. nach dem Abruf stattfindende) Urteile unterschieden werden. Vor dem eigentlichen Lernen werden also bereits *Ease-of-Learning-Urteile* (EOLs) abgegeben, also Urteile darüber, wie leicht oder schwierig die künftig zu lernenden Inhalte sein werden. Auf Grundlage dieser Information wird dann entschieden, welches Vorgehen am ehesten zum Ziel führt bzw. welche Strategie die erfolgreichste sein wird (*selection of kind of processing*) und wieviel Zeit wohl den einzelnen Lerneinheiten zugeteilt werden sollte (*initial plan for allocation of study time*). Bei *Judgments of Learning* (JOLs) handelt es sich um Einschätzungen, die während oder nach dem Lernvorgang abgegeben werden und Vorhersagen über die künftige Erinnerungsleistung darstellen. Wenn die Information aus den JOLs besagt, dass Inhalte beherrscht werden, endet der Vorgang des Einprägens an dieser Stelle (*termination*

of study), anderenfalls wird den entsprechenden Items noch einmal Lernzeit zugeteilt (*allocation of study time*). *Feeling-of-Knowing-Urteile* (FOKs) beziehen sich auf Vorhersagen, inwiefern aktuell nicht erinnerte Items zu einem späteren Zeitpunkt erinnert werden können. FOKs nehmen beim Abrufprozess darauf Einfluss, ob die Suche nach Inhalten fortgesetzt wird oder nicht (*selection of search strategy, termination of search*). Zum Schluss wird ein Sicherheitsurteil (SU; *confidence judgment*) abgegeben, d.h. ein retrospektives Urteil, ob die gegebene Antwort korrekt ist oder nicht.

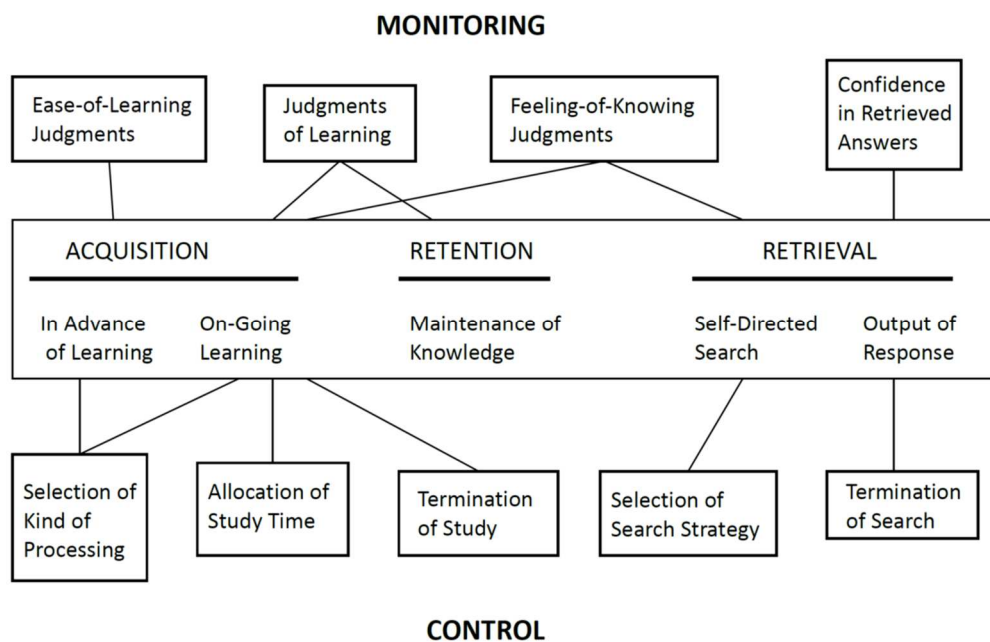


Abbildung 2: Modell des prozeduralen Metagedächtnisses nach Nelson und Narens (1990,1994)

Insgesamt betrachtet förderte das Modell von Nelson und Narens (1990, 1994) aufgrund seiner Ausdifferenziertheit und der zeitlichen Verortung der Überwachungs- und Steuerungsprozesse im Lernablauf eine Vereinheitlichung der gängigen Nomenklatur zumindest für das prozedurale Metagedächtnis.

1.2.6 Modell des guten Informationsverarbeiters

Obwohl chronologisch etwas vor dem Modell von Nelson und Narens (1990, 1994) angesiedelt, soll das *Modell des Guten Informationsverarbeiters* (Pressley, Borkowski & Schneider, 1989)

am Ende dieses Überblicks über die konzeptuelle Entwicklung des Metagedächtnisses stehen. Denn dieses Modell wie auch seine Vorgänger schafften erstmalig wieder eine Verbindung zwischen den deklarativen und den prozeduralen Aspekten des Metagedächtnisses. Im Jahr 1985 entstand mit dem Modell des *Metamemory about Strategies* (MAS; Pressley, Borkowski & O'Sullivan, 1985) ein integratives und anwendungsbezogenes Konzept, welches darauf abzielte, zu erklären, wie Strategiewissen, die Auswahl und Anwendung von Strategien sowie die stetige Weiterentwicklung dieses Wissens durch Feedback aus der Überwachung des Erfolgs oder Misserfolgs eines Lernprozesses zusammenwirken und eine effiziente Informationsverarbeitung fördern. Im Modell des *Good Strategy User* (Pressley, Borkowski & Schneider, 1987) wurde das MAS wiederum in einen größeren Zusammenhang eingeordnet, indem zusätzlich zum Metagedächtnis über Strategien auch nicht strategisches Weltwissen und die – bei ausreichender Übung – automatisierte Anwendung und Koordination strategischen Handelns als wesentlich für eine erfolgreiche kognitive Verarbeitung aufgeführt wurden. Hervorzuheben ist, dass der Gute Strategienutzer auch über allgemeines Wissen über das Zustandekommen kognitiver Leistungen verfügt, also z.B. Wissen darüber, dass und wie Anstrengung, Motivierung und eine geeignete Aufmerksamkeitslenkung den Erfolg von Informationsverarbeitungsprozessen beeinflussen. Noch umfassender war das kurz darauf erschienene Modell des *Good Information Processor* (Pressley et al., 1989), das Einflussfaktoren auf kognitive Performanz um günstige neurologische Voraussetzungen sowie förderliche, d.h. intellektuell stimulierende Umweltbedingungen erweiterte.

1.2.7 Abschließende Bewertung

Dieser Überblick über die Konzeptentwicklung kann nicht erschöpfend sein, zumal es sich hierbei um einen dynamischen Prozess handelt. Ganz allgemein bleibt festzuhalten, dass es sich bei den Begriffen „Metakognition“ oder „Metagedächtnis“ nicht um ein vollständig neues Konstrukt handelt. Die Fähigkeit zur Selbstreflexion ist eine Eigenschaft, die dem denkenden Menschen genuin zu eigen ist. Mindestens seit der Antike finden sich Zeugnisse von „Kognitionen über Kognitionen“ und vom Nachdenken über eigene Gedächtnisprozesse, so z.B. Sokrates' Ausspruch „Ich weiß, dass ich nichts weiß“, die Inschrift im Orakel von Delphi „Erkenne dich selbst!“ oder die Legende von Simonides, der die sog. Loci-Methode als Gedächtnisstrategie anwendete (vgl. Dunlosky & Metcalfe, 2009). Zu Beginn der wissenschaftlichen Psychologie im 19. Jahrhundert war die Introspektion eine der wesentlichen Methoden, um Erkenntnisse über kognitive Vorgänge zu gewinnen. So wurden z.B. von

Wissenschaftlern der „Würzburger Schule“ um Oswald Külpe nach Denk- oder Assoziationsaufgaben systematische Selbstberichte von Probanden erhoben, um davon ausgehend Rückschlüsse auf zugrunde liegende Prozesse ziehen zu können (Schneider, 1999). Mit Aufkommen des Behaviorismus (z.B. Watson, 1913) verlor die Methode der Introspektion an Bedeutung; erst im Zuge der „Kognitiven Wende“ (Dember, 1974; Neisser, 1967) rückte das Interesse an mentalen Prozessen wieder in den Vordergrund.

Flavell (1971) steht damit als Entwicklungspsychologe in einer Tradition von kognitiven Psychologen. Seine Pionierleistung kann zum einen in der Bündelung bislang nebeneinander existierender Forschungslinien zur Gedächtnisentwicklung gesehen werden. Verschiedenste „Kognitionen zweiter Ordnung“ (Kuhn, 1999a), die besonders im Lauf der ersten zehn Lebensjahre dazu beitragen, Lern- und Erinnerungsprozesse effektiver, sparsamer und planvoller zu gestalten, wurden von Flavell (1971) erstmals in einem einzigen Konzept zusammengefasst. Zum anderen schuf er mit der Begründung des Terminus „Metagedächtnis“ einen Begriff, der den Austausch zwischen Wissenschaftlern mit verschiedenen Ansätzen ermöglichte. Den in rascher Folge entstandenen theoretischen und praktischen Artikeln ist der Enthusiasmus über die neue Forschungsrichtung zu entnehmen, den schon Flavell (1971) formulierte: „Let’s all go out and study it!“ (S. 277). Eine Vielzahl an selbstreflexiven kognitiven Prozessen wurde mit der Vorsilbe „meta“ versehen (vgl. Brown, 1978; Cavanaugh & Perlmutter, 1982; Dunlosky & Metcalfe, 2009) und empirisch untersucht (z. B. Flavell et al., 1970; Markman, 1977).

Kritisch ist zu sehen, dass Flavells (1971) bzw. Flavell und Wellmans (1977) erste Definitionen und Modelle eher unscharf blieben und keine klare Trennung zwischen exekutiven und Wissensaspekten vornahmen (vgl. Cavanaugh & Perlmutter, 1982). Weiterhin wurde bemängelt, dass die einzelnen Subkategorien, wie z.B. die Sensitivitäts- und die Aufgabenkategorie zum Teil schwer voneinander abzugrenzen sind (Groeben, 1982). Insgesamt handelte es sich also um ein eher offenes Konzept (Yussen, 1985). Dies erschwerte einerseits die Operationalisierung und somit auch die Vergleichbarkeit verschiedener experimenteller Arbeiten zum Metagedächtnis. Andererseits hat die Unschärfe des Begriffs sicher auch die intensive Auseinandersetzung und die Weiterentwicklung des Verständnisses von Metakognitionen angeregt.

Ann Brown (1975, 1978) präziserte und erweiterte den Ansatz von Flavell und Wellman (1977). Ihre Konzeptualisierung trug der Beobachtung Rechnung, dass erfolgreiches Lernen in hohem Maße einer Selbstüberwachung und -steuerung bedarf. Ihre Arbeiten lenken also den

Fokus von reinen Wissensaspekten hin zu Prozessen, die während des Lernens (online) ablaufen. D.h. Brown (1978) leistete wichtige Vorarbeiten dazu, dass prozedurales und deklaratives Metagedächtnis später genauer abgegrenzt wurden. Ihr Artikel von 1978 stellt eine Zusammenschau verschiedenster prozeduraler Metagedächtnisvorgänge dar, die später zum Teil auch in das Modell von Nelson und Narens (1990, 1994) einfließen. Insgesamt kann Browns (1975, 1978) Verständnis des Metagedächtnisses ebenfalls noch als unscharf bezeichnet werden kann – es bleibt z.B. unklar, wie die einzelnen Vorgänge genau interagieren oder wie sie operationalisiert werden können. Sie stimulierte jedoch die Weiterentwicklung des Konzeptes und stellte den Anwendungsbezug her (z.B. zum Lesenlernen; Brown, 1980).

In der weiteren Entwicklung des Konzepts „Metagedächtnis“ etablierte sich die auch heute noch gültige Unterscheidung zwischen deklarativen und prozeduralen Aspekten (Kluwe, 1982; Paris et al., 1983). Die Arbeitsgruppe um Paris (z.B. Paris et al., 1983; Paris & Lindauer, 1982; Paris & Oka, 1986) aus der Entwicklungs- und Pädagogischen Psychologie sowie Kluwe (1982) aus der Allgemeinen Psychologie trieben beide die Ausdifferenzierung und Vereinheitlichung der Begriffe trotz unterschiedlicher Schwerpunktsetzung voran. In Bezug auf das prozedurale Metagedächtnis waren schließlich Nelson und Narens (1990, 1994) diejenigen, welche eine große Anzahl der bisher untersuchten Überwachungs- und Kontrollprozesse in einem Modell zusammenführten. Doch auch dieses Modell kann nicht als vollständig betrachtet werden, sondern es bleibt offen für weitere Überwachungs- und Steuerungsmaße im zeitlichen Verlauf des Lernprozesses, wie z.B. für Maße der Quellenüberwachung (*source monitoring*). Auch wird der postulierte uni-direktionale Zusammenhang zwischen beiden Facetten (Überwachung und Steuerung) in neuerer Zeit zunehmend kritisiert (z.B. Koriat, Ma'ayan & Nussinson, 2006).

Die theoretische Weiterentwicklung und empirische Arbeiten standen und stehen weiterhin in enger Verflechtung miteinander und beeinflussen sich gegenseitig. Die Ausführungen haben verdeutlicht, dass – nachdem die Metagedächtnisforschung zunächst vorrangig entwicklungspsychologisch motiviert war (z.B. Flavell, 1971; Brown, 1975) – recht bald auch Wissenschaftler aus dem Bereich der Informationsverarbeitung (z.B. Kluwe, 1982; Nelson & Narens, 1990, 1994) Beiträge zur Konzeptualisierung leisteten. Seit einigen Jahren gibt es zunehmend Arbeiten aus dem Feld der Neuropsychologie, die sich mit neuronalen Korrelaten metakognitiver Aktivitäten auseinandersetzen (z.B. Fleming & Dolan, 2012; McCurdy et al., 2013; Pannu & Kaszniak, 2005). Weiterhin ist hervorzuheben, dass sich nicht nur die Gedächtnisforschung, sondern auch andere Gebiete der Kognitionspsychologie mit Metaprozessen beschäftigen, wie beispielsweise die „Theory of Mind“-Forschung (ToM). Diese widmet sich vor allem im Kleinkind- und Vorschulbereich der Entwicklung des Wissens

um mentale Prozesse anderer. D.h., es geht um die Frage, inwiefern junge Kinder bereits Wissen darüber erworben haben, dass anderen Personen eine mentale Welt zu eigen ist, die von der eigenen abweichen kann (vgl. Lockl & Schneider, 2006). Kuhn (1999b, 2000) trug wesentlich zur Abgrenzung der genannten Termini voneinander bei, indem sie das Wissen über mentale Zustände (hier: *metacognitive knowing*), wie es die ToM-Forschung untersucht, und das Metagedächtnis sensu Flavell (1971; hier: *metastrategic knowing*) als zwei große Teilbereiche des Oberbegriffs Metakognition einordnete. Kuhn (1999a) verwendete in ihrem *Developmental Model of Critical Thinking* zwar zum Teil von etablierten Konzepten abweichende Begriffe; sie bietet jedoch ein gutes Ordnungssystem der verschiedenen Facetten von „Kognitionen zweiter Ordnung“, das von Schneider & Lockl (2006) weiter ausgearbeitet wurde (s. Abb. 3).

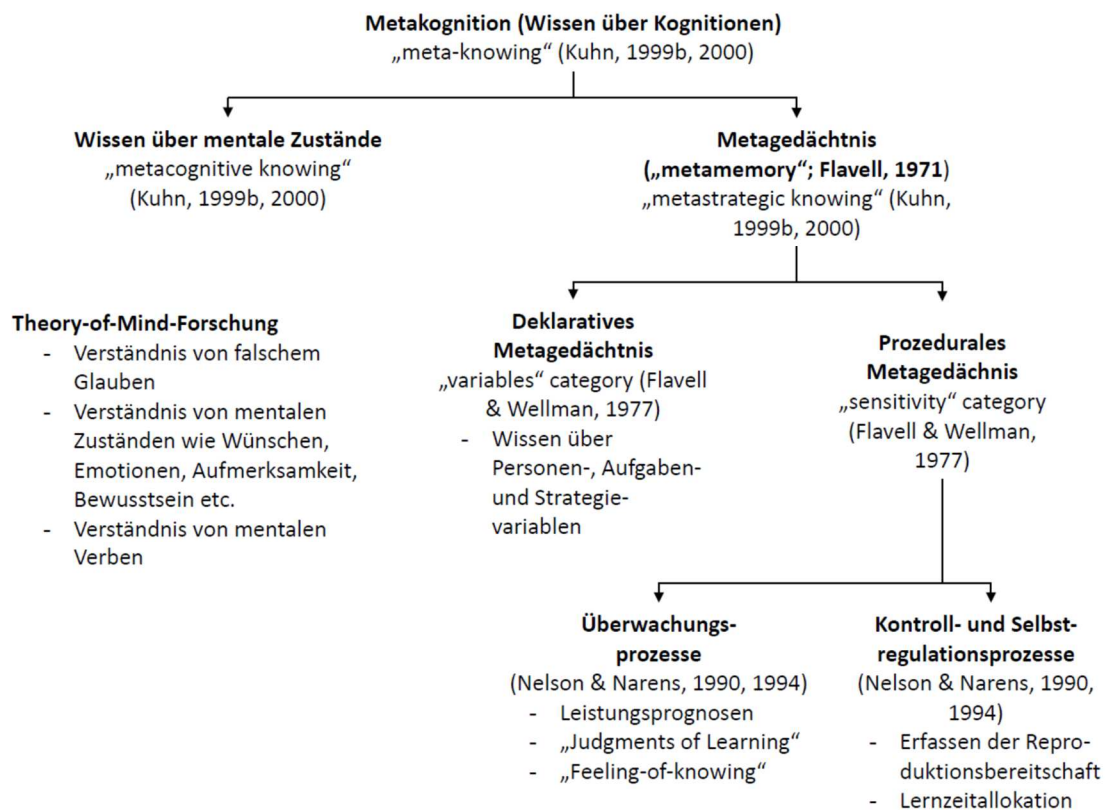


Abbildung 3: Klassifikation verschiedener Komponenten der Metakognition (nach Schneider & Lockl, 2006)

Es werden auch in Zukunft Konzepte nötig sein, die – ähnlich wie das Modell des Guten Informationsverarbeiters (Pressley et al., 1989) oder die Arbeiten von Kuhn (1999a, 1999b; 2000) – verschiedene Forschungslinien bündeln. Denn mit zunehmendem Volumen der

existierenden Studien steigt auch die Zahl an zugrunde liegenden Maßen und Annahmen. Für die eigene Arbeit wurde die Trennung nach prozeduralen und deklarativen Aspekten übernommen, da diese sich in der Literatur etabliert hat und vor Kurzem auch faktorenanalytisch untermauert werden konnte (Scott & Levy, 2013). Innerhalb des prozeduralen Metagedächtnisses kommen die Maße des Modells von Nelson und Narens (1990, 1994) zur Anwendung, da diese bislang am eindeutigsten operationalisierbar sind.

2. DEKLARATIVES METAGEDÄCHTNIS

Auch wenn das deklarative Metagedächtnis nicht im Fokus der eigenen Arbeit steht, soll im Folgenden auf die Erfassungsmethoden, die Entwicklung über die Lebensspanne sowie den Zusammenhang zu Gedächtnisleistungen eingegangen werden. Denn trotz der prinzipiellen Unabhängigkeit von prozeduralen und deklarativen Aspekten (Scott & Levy, 2013) können wechselseitige Beeinflussungen gerade im Entwicklungsverlauf nicht ausgeschlossen werden (Flavell et al., 2002; Schraw, 1994).

2.1 Erfassungsmethoden

Als wissensbasierte Komponente ist das deklarative Metagedächtnis laut Flavell (1979) strukturell nicht von anderen Inhalten des Langzeitgedächtnisses zu unterscheiden. Das heißt, Informationen daraus können willentlich abgerufen und verbalisiert werden, können aber auch automatisch durch Hinweisreize in einer kognitiven Aufgabe aktiviert werden und den aktuellen Vorgang beeinflussen (wie z.B. die unbewusste Nutzung einer Merkstrategie). Das Wissen über Gedächtnisvorgänge kann – als Teil des Weltwissens einer Person – sowohl zutreffend als auch fehlerhaft sein, d.h. es kann bei einer Aktivierung auch keine oder nachteilige Auswirkungen haben (Flavell, 1979).

Grundsätzlich wird das deklarative metakognitive Wissen nicht während gleichzeitig ablaufender Gedächtnisaktivitäten, sondern in sog. unabhängigen Messungen (Cavanaugh & Perlmutter, 1982) erfasst. Die befragte Person soll also allgemeine Aussagen über ihr Gedächtniswissen treffen.

Hierfür werden zum einen Interviews eingesetzt. In der Regel werden den Probanden Szenarien vorgelegt, bei denen z.B. Strategien zur Anwendung kommen (z.B. Kreutzer et al., 1975). Die Teilnehmer sollen dann in offenem oder geschlossenem Antwortformat darüber Auskunft geben, welche Strategien ihnen einfallen, um sich die Aufgabe zu erleichtern, oder ob sich ältere oder jüngere Kinder Dinge besser merken können, um nur zwei Beispiele zu nennen. Besonders bei Kindern im Kindergartenalter ist die Interpretation der Ergebnisse dieser Studien allerdings unter Vorbehalt zu sehen, da diese Verfahren ausgeprägte verbale Fertigkeiten verlangen, die in diesem Alter nicht grundsätzlich gegeben sind. Um dieses Problem abzumildern, setzten Wellman (1977a) und Justice (1989) Bilder bzw. Videos ein, auf denen Kinder beispielhaft verschiedene Strategien anwendeten, und ließen deren Effektivität von den Teilnehmern beurteilen.

Trotz dieser Verbesserungen ist der Einsatz von Interviews zur Erfassung des deklarativen Metagedächtniswissens kritisch zu sehen, da sie zum einen sehr aufwändig sind, zum anderen Validität, Reliabilität und Objektivität der Messung eingeschränkt bzw. nicht eindeutig zu beurteilen sind. 1988 wurde von Belmont und Borkowski ein erster Gruppentest zur Erfassung des deklarativen Metagedächtnisses bei Sechs- bis Zwölfjährigen entwickelt, der den Erhebungsaufwand deutlich verringerte und die Qualität der Ergebnisse erhöhte. In den Aufgaben werden jedoch prozedurale und deklarative Inhalte vermischt. Außerdem verfügt das Verfahren nicht über eine Normierung oder den Vergleich mit einem Standard, was eine genaue Einordnung der Ergebnisse eines einzelnen Probanden erschwert. Die Würzburger Testbatterie zum deklarativen Metagedächtnis (Schlagmüller, Visé & Schneider, 2001) zur Anwendung bei Dritt- und Viertklässlern stellt insofern eine Verbesserung dar, als sie – neben einer hohen Objektivität – verbesserte Reliabilitätskennwerte liefert, die Validität empirisch überprüft sowie eine große Vergleichsstichprobe heranzieht. Die Antworten der Teilnehmer in den erfassten Wissenskomponenten werden mit den Antworten von Experten verglichen, ein Vorgehen, das auch schon Schneider (1986) sowie Justice (1989) einsetzten. Gute psychometrische Eigenschaften weist auch ein neueres Verfahren für jüngere Kinder von Haberkorn, Lockl, Pohl, Ebert und Weinert (2014) auf. Die Autorinnen beziehen sich in ihrem Gruppentest für Erst- und Zweitklässler u.a. auf Items von Kreutzer et al. (1975) sowie Schlagmüller et al. (2001) und setzen ebenfalls Paarvergleiche zur Bewertung der gegebenen Antworten ein.

In einzelnen Domänen wie Lesen oder Mathematik sind in den letzten Jahren normierte Testverfahren entstanden, die deklaratives Metagedächtniswissen bereichsspezifisch erfassen (z.B. Würzburger Lesestrategie-Wissenstest für die Klassen 7 – 12, WLST 7-12; Schlagmüller & Schneider, 2007; Mathematisches Strategiewissen für 5. und 6. Klassen, MAESTRA 5-6+; Lingel, Götz, Artelt & Schneider, 2014).

Für das Erwachsenenalter gibt es bislang keine vergleichbaren Maße. Einen Ansatz zur Erfassung von deklarativen Metagedächtnisprozessen stellen Fragebögen zur „metakognitiven Bewusstheit“ (*metacognitive awareness*) dar. Schraw und Dennison (1994) zielen z.B. mit dem von ihnen entwickelten MAI (Metacognitive Awareness Inventory) darauf ab, sowohl Wissen als auch die Regulation von Kognitionen zu erfassen. Die Autoren finden empirische Belege für die Zwei-Faktoren-Struktur ihres Fragebogens (Wissen und Regulation). Im Gegensatz zu den oben vorgestellten Fragebögen zum Gedächtniswissen erfasst der MAI jedoch Selbsteinschätzungen, deren Bezugsnorm ungeklärt bleibt. Somit ist der Nutzen eines „Wissensitems“ wie beispielsweise „Ich habe die Kontrolle darüber, wie gut ich lerne“ fraglich,

da sich die Aussage nicht auf eine konkrete Aufgabe bezieht und auch ein bedeutsamer Einfluss der sozialen Erwünschtheit nicht ausgeschlossen werden kann. Eine flächendeckende Anwendung des MAI für erwachsene Probanden ist deshalb kritisch zu sehen, weil die Entwicklung ausschließlich unter College-Studenten erfolgte. Auch die Güte des JuniorMAI (Sperling, Howard, Miller & Murphy, 2002), der für Dritt- bis Neuntklässler entwickelt wurde, ist kritisch zu sehen: So fanden die Autoren keine konsistenten Korrelationen ihres Fragebogens mit anderen Maßen zur Erfassung des deklarativen Metagedächtnisses.

Besonders für das höhere Erwachsenenalter liegt der Fokus bei Instrumenten zur Erfassung des Gedächtniswissens vor allem auf der Frage nach spezifischen Annahmen über das eigene Gedächtnis, weniger auf dem Wissen über Personen-, Aufgaben- und Strategievariablen. So deckt der von Dixon und Hultsch (1983) entwickelte Fragebogen *Metamemory in Adulthood* (MIA) beispielsweise Selbsteinschätzungen auf den Dimensionen Gebrauch von Gedächtnis-Strategien, Gedächtniskapazität und Veränderungen derselben oder auch wahrgenommene Kontrolle über (altersbedingte) Verluste in der Gedächtnisleistung ab. Auf ähnliche Weise ist auch der *Memory Functioning Questionnaire* (MFQ; Gilewski, Zelinski & Schaie, 1990) konzipiert, der Selbstbeurteilungen auf verschiedenen Faktoren (z.B. Häufigkeit von Vergessen oder Häufigkeit der Nutzung von Gedächtnisstrategien) abfragt und hohe Zusammenhänge mit dem MIA aufweist (Hertzog, Dixon & Hultsch, 1990).

Zusammenfassend kann also festgehalten werden, dass es bis heute kein Standardverfahren zur Erfassung des deklarativen Metagedächtniswissens gibt – besonders wenn verschiedene Altersgruppen über die Lebensspanne hinweg verglichen werden sollen. Testaufgaben, bei denen die Einschätzung von Probanden zur Effektivität bestimmter Strategien oder zur Auswirkung von Personen- oder Aufgabenmerkmalen mit denen von Experten verglichen werden, stellen nach heutigem Stand die beste Methode dar, das deklarative Metagedächtnis zu erfassen. Bislang sind hier die domänenspezifischen Verfahren die elaboriertesten Maße; diese umfassen jedoch nur einen eingeschränkten Altersbereich.

2.2 Entwicklung

Es soll nun ein Überblick über die Entwicklung von deklarativem Metagedächtniswissen über die Lebensspanne gegeben werden. Beginnend mit wichtigen Vorläuferfertigkeiten wird im Kindes- und Jugendalter vor allem auf die Entwicklung von Wissen über Personen-, Aufgaben- und Strategiemerkmale sowie deren Interaktion eingegangen (vgl. Flavell & Wellman, 1977).

Im höheren Erwachsenenalter beziehen sich Erhebungen zum deklarativen Metagedächtnis wie oben dargestellt eher auf allgemeine Selbsteinschätzungen der eigenen kognitiven Leistungsfähigkeit.

Vorläuferfertigkeiten: Das Verständnis mentaler Verben. Damit Kinder über ihr eigenes Gedächtnis Auskunft geben können, muss ihnen ein geeignetes Vokabular zur Beschreibung innerer kognitiver Vorgänge, d.h. sogenannte *mentale Verben*, zur Verfügung stehen. Bereits in den Anfängen der Forschungen zum Metagedächtnis wurden Kindergartenkinder anhand konkreter Situationen zu ihrem Verständnis von mentalen Verben untersucht. Es scheint demzufolge, dass es Dreijährigen im Gegensatz zu Vierjährigen noch schwer fällt, zwischen „wissen“ und „raten“ (Miscione, Marvin, O'Brien & Greenberg, 1978) oder zwischen „vergessen“ und „erinnern“ (Wellman & Johnson, 1979) zu unterscheiden. Auch Vierjährige hatten jedoch in diesen Studien Schwierigkeiten, den Gebrauch dieser Verben flexibel zu handhaben. Bei Miscione et al. (1978) fokussierten die Vierjährigen noch stark auf das Handlungsergebnis, um zwischen „wissen“ und „raten“ zu differenzieren, bei Wellman und Johnson (1979) gelang es erst den Fünfjährigen, neben der Erinnerungsleistung auch das Vorwissen als wichtige definitorische Komponente von „vergessen“ und „raten“ zu berücksichtigen.

Mit dem Beginn der Grundschule ist die Entwicklung des Verständnisses von mentalen Verben noch nicht abgeschlossen, sondern erfährt weitere Verfeinerungen. So unterschieden sich in einer Studie von Schwanenflugel, Fabricius und Alexander (1994) Acht- und Zehnjährige in dem Maß, in dem sie mentale Aktivitäten (z.B. Wiedererkennen, Planen, Verstehen) nach dem Einfluss von Gedächtnis- und Aufmerksamkeitsvorgängen und dem von internalen und externalen Hinweisreizen differenzierten. Auch die Bedeutung der Sicherheit des eigenen Urteils bei der Einordnung von mentalen Verben nimmt im Lauf der Grundschulzeit zu (Schwanenflugel, Fabricius & Noyes, 1996). Astington und Olson (1990) baton Probanden der 6. bis 12. Klasse sowie Erwachsene darum, in Texten „sagen“ und „denken“ durch geeignetere Wörter zu ersetzen und fanden auch in diesem Altersbereich noch bedeutsame Fortschritte im Verständnis und der Differenziertheit dieser Verben.

Basale Fertigkeiten, zwischen einfachen mentalen Verben zu unterscheiden, sind also ab dem Alter von etwa vier Jahren vorhanden. Welche Faktoren aber in welchem Maß verschiedene Lern- und Erinnerungsprozesse beeinflussen sowie die entsprechenden Begriffe dazu, lernen Kinder erst im Vorschul- bzw. Grundschulalter. Oben zitierte Arbeiten aus dem Bereich der

Theory-of-Mind-Forschung (Schwanenflugel et al., 1994, 1996; Astington & Olson, 1990) zeigen, dass es bis ins Erwachsenenalter hinein noch Fortschritte in der Differenzierungsfähigkeit zwischen verschiedenen mentalen Verben gibt. Insgesamt gesehen lassen diese Ergebnisse darauf schließen, dass die Beurteilung eigener Gedächtnisprozesse besonders bei sehr jungen Kindern wohl nur sehr eingeschränkt möglich ist, da grundlegende verbale Fertigkeiten hierfür noch fehlen. Doch auch im Kindes- und Jugendalter können Begriffe, denen komplexere Konzepte zugrunde liegen und für deren genaues Verständnis die Interaktion verschiedener Aspekte (Vorwissen, Aufmerksamkeit, Sicherheit etc.) berücksichtigt werden muss, nicht grundsätzlich als gegeben vorausgesetzt werden.

Entwicklung im Kindesalter. Eine der ersten und einflussreichsten Arbeiten zur Entwicklung des deklarativen Metagedächtnisses im Kindesalter stellt die Interviewstudie von Kreuzer et al. (1975) dar. Die Autoren befragten 80 Kinder zwischen dem Kindergartenalter und der fünften Klasse zu verschiedenen Gedächtnisbereichen. Sie fanden, dass schon die meisten der jüngsten Teilnehmer basales Wissen über Gedächtnisvorgänge hatten, so z.B., dass sie bisweilen auch Inhalte vergessen oder dass es einfacher ist, Dinge wieder zu lernen, die man bereits einmal wusste. Je älter die teilnehmenden Kinder waren, desto mehr Wissen hatten sie jedoch darüber erworben, dass die Gedächtnisleistung in verschiedenen Situationen und zwischen verschiedenen Personen variiert, dass es leichter ist, sich gegensätzliche Begriffspaare zu merken als vollkommen unzusammenhängende, oder dass die Kategorisierung von Objekten eine effektive Strategie sein kann, um nur einige Beispiele zu nennen. Rudimentäres Wissen über bestimmte Personen- und Aufgabenmerkmale konnte also in dieser Studie schon für die Kindergartenkinder gefunden werden. Das Wissen über Strategien zeigte sich erst bei den Grundschulern. Insgesamt konnten bei nahezu allen Interviewfragen Verbesserungen mit zunehmendem Alter nachgewiesen werden.

Die angemessene Beurteilung der Interaktion mehrerer gedächtnisrelevanter Variablen (z.B. Aufgaben- und Strategiemerkmale) wurde etwas später von Wellman (1978) bei Fünf- und Zehnjährigen untersucht. Obwohl die Kindergartenkinder in der Lage waren, die Bedeutung isolierter Merkmale korrekt zu identifizieren, gelang es ihnen in der Regel nicht, mehrere Dimensionen zu berücksichtigen, um die Schwierigkeit einer Gedächtnisaufgabe zu beurteilen. Hier waren die älteren den jüngeren Teilnehmern deutlich überlegen. Nachfolgende Studien zum deklarativen Metagedächtnis bezogen sich in der Regel auf die klassische Arbeit von

Kreutzer et al. (1975) und konnten die gefundenen Entwicklungstrends im Wesentlichen bestätigen (z.B. O'Sullivan, 1993; Schneider & Sodian, 1988).

Intuitiv ist anzunehmen, dass der Zuwachs im deklarativen Metagedächtniswissen nach dem Eintritt in die Schule stärker ausgeprägt ist als während der Kindergartenzeit, da hier die Auseinandersetzung mit formalisierten Lern- und Gedächtnisaufgaben deutlich zunimmt. Im Gegensatz zu frühen querschnittlich angelegten Studien (z.B. Kreutzer et al., 1975; Wellman, 1977a) liegen inzwischen auch Längsschnittbefunde vor, die deutlich machen, dass die Anstiegsraten – auch über den Schuleintritt hinweg – linear verlaufen. So erfassten z.B. Annevirta und Vauras (2001) Wissen über kognitive Prozesse in den Dimensionen Gedächtnis, Verständnis und Lernen von der Vorschule bis zur dritten Klasse und fanden insgesamt gleichmäßige Anstiege über die Zeit. Hervorgehoben werden soll hierbei der Befund, dass auch Erstklässler noch nicht den Verstand („mind“) als wichtigstes Werkzeug für kognitive Aktivität ansahen, sondern ihrer Rolle als aktiver Lerner die größte Relevanz beimaßen. Ferner zeigte sich, dass im Gegensatz zu den Vorschulkindern, die bereits ein Konzept des Begriffes „Gedächtnis“ aufwiesen, sich erst die Teilnehmer im Alter von acht bis neun Jahren auf den Begriff „Verständnis“ bezogen, wenn es um die Beschreibung mentaler Vorgänge ging. Es zeigt sich also eine zunehmende Ausdifferenzierung kognitiver Prozesse.

Belege für einen graduellen Anstieg des deklarativen Metagedächtniswissens finden auch Lockl und Schneider (2006) für den Übergang vom Kindergarten zur Grundschule sowie Grammer, Purtell, Coffman und Ornstein (2011) für die ersten beiden Schuljahre. Mit Blick auf beeinflussende Faktoren konnten letztere nachweisen, dass das deklarative Metagedächtnis zu Beginn der ersten Klasse vor allem durch die Arbeitsgedächtnisleistung und Merkmale der häuslichen Umwelt (z.B. Häufigkeit des gemeinsamen Lesens o.ä.) vorhergesagt werden konnte. Der Anstiegsgradient im Verlauf der beiden ersten Schuljahre wurde dagegen überwiegend vom Bildungsniveau der Mutter bestimmt – ein Befund, der wahrscheinlich auf Unterschiede im Umgang mit den Kindern (z.B. Gesprächsstil, Häufigkeit von kognitiven Aktivitäten) begründet ist.

Generell scheinen sich verschiedene Dimensionen des deklarativen Metagedächtnisses im Kindesalter eher unabhängig voneinander, aber dennoch gleichmäßig zu entwickeln (z.B. Schlagmüller et al., 2001; Haberkorn et al., 2014). Es konnte kein Nachweis dafür erbracht werden, dass Kinder, die zu Beginn der Grundschulzeit eher Defizite im deklarativen Metagedächtnis aufweisen, diese später aufholen können (Grammer et al., 2011).

Entwicklung im Jugendalter. Selbst wenn also beim Übertritt in die Sekundarstufe das Wissen um Gedächtnisvorgänge im Wesentlichen schon vorhanden ist, kann die Entwicklung zu diesem Zeitpunkt nicht als abgeschlossen betrachtet werden.

So fanden z.B. Wellman, Collins und Glibermann (1981) bei einem Vergleich von 5-, 8-, 10- und 19-Jährigen auch noch Vorteile zugunsten der ältesten Gruppen bei der Aufgabe, verschiedene Einflussfaktoren auf die Gedächtnisleistung (u.a. Anstrengung, Umfang der Aufgabe) gegeneinander abzuwägen. Im Unterschied zu dieser Studie konzentrieren sich die meisten Arbeiten über das Jugendalter allerdings auf bereichsspezifische Anteile von deklarativem Metagedächtniswissen bzw. auf die Frage, inwiefern Belege für spezifische und generalisierte Anteile gefunden werden können. Für die Fächer Englisch und Deutsch konnten im Verlauf der fünften und sechsten Klasse bedeutsame Verbesserungen im domänenspezifischen Metagedächtnis gezeigt werden; die mittleren Lösungsraten dieser Studie machen aber auch deutlich, dass die Entwicklung in diesem Alter noch nicht abgeschlossen sein kann (Artelt, Neuenhaus, Lingel & Schneider, 2012). Der von Grammer et al. (2011) gezeigte Befund, dass kein Aufholen schwächerer Schülerinnen und Schüler vorliegt, wird von Artelt et al. (2012) bestätigt. Für den Vergleich der Bereiche Geschichte und Mathematik wiesen van der Stel und Veenman (2014) in einer Längsschnittstudie nach, dass der Anstieg im Metagedächtniswissen im Alter zwischen 12 und 15 Jahren nicht durchgehend linear verläuft wie für die Vor- und Grundschule gezeigt, sondern dass mit 15 Jahren ein vorläufiges Leistungsplateau erreicht wird. Eine Arbeit zum metakognitiven Strategiewissen im Kontext „Deutschaufsatz“ mit Schülern der 10. und 11. Klasse fand im Verlauf eines Schuljahres keine weiteren Anstiege. Hier wurde zusätzlich gezeigt, dass ein höherer sozioökonomischer Status, weibliches Geschlecht sowie Lernmotivation und Selbstwirksamkeit mit höheren Metagedächtniswerten einhergehen (Karlen, Maag Merki & Ramseier, 2014).

Obwohl insgesamt die domänenspezifische Erfassung von deklarativem Metagedächtniswissen in der Sekundarstufe überwiegt, gibt es Nachweise für einen starken Zusammenhang zwischen fächerübergreifenden und generalisierten Anteilen (Neuenhaus, Artelt, Lingel & Schneider, 2011; van der Stel & Veenman, 2014). Es besteht also die Notwendigkeit, auch entsprechende domänenübergreifende Messverfahren zu konstruieren, die in diesem Altersbereich noch ausreichend differenzieren, da die Entwicklung des deklarativen Metagedächtnisses auch bis ins Erwachsenenalter noch nicht als abgeschlossen betrachtet werden kann (vgl. Schneider, 2015).

Entwicklung im höheren Erwachsenenalter. Die Erfassung von deklarativem Metagedächtniswissen im Kindes- und Jugendalter erfolgt also mehrheitlich über den Einsatz von Tests, in denen der Einfluss verschiedener Personen-, Aufgaben- und Strategiemerkmale beurteilt werden muss. Wie weiter oben beschrieben geht der Ansatz bei älteren Erwachsenen in eine andere Richtung: Hier werden mehrheitlich Selbsteinschätzungen über das eigene Gedächtnis in Bezug auf die intraindividuelle Entwicklung erfasst. Obwohl also die Vorgehensweisen zur Erfassung des deklarativen Metagedächtnisses in den verschiedenen Lebensabschnitten nicht direkt vergleichbar sind, sollen an dieser Stelle dennoch ausgewählte Ergebnisse aus entsprechenden Studien mit älteren Erwachsenen vorgestellt werden.

Recht durchgängig belegen empirische Studien die alltägliche Beobachtung, dass ältere Erwachsene häufig berichten, mehr Schwierigkeiten als früher zu haben, neue Informationen zu erlernen, und dass ihre Gedächtnisleistung insgesamt abnimmt (Hultsch, Hertzog & Dixon, 1987). Weiterhin gehen ältere Erwachsene im Vergleich zu jüngeren weniger davon aus, dass Defizite in Gedächtnisleistungen potenziell kontrollierbar sind (M. E. Lachman, 2006). Hier stellt sich jedoch die Frage, ob diese Selbsteinschätzungen mit tatsächlichen Verschlechterungen in Gedächtnisleistungen einhergehen oder ob vor allem Faktoren wie der Einfluss von Altersstereotypen (*stereotype threat*; z.B. Heckhausen, Dixon & Baltes, 1989) wirksam werden. Mehrfach zeigten sich zwar in Längsschnittstudien korrelative Zusammenhänge mit (negativen) Selbstbeurteilungen und (schlechteren) Gedächtnisleistungen. Diese fielen in der Regel aber recht gering aus (z.B. McDonald-Miszczak, Hertzog & Hultsch, 1995; Valentijn et al., 2006), d.h., dass der Einfluss impliziter Theorien über gedächtnisbezogene Altersprozesse insgesamt mehr Gewicht zu haben scheint als tatsächliche Veränderungen (vgl. auch Lineweaver & Hertzog, 1998). Einschränkend muss auch berücksichtigt werden, dass Selbstbeurteilungsfragebögen generell mehr Ungenauigkeiten unterliegen als (standardisierte) Wissenstests (Rabbitt, Maylor, McInnes, Bent & Moore, 1995) und dass sich weitere, zumeist unberücksichtigte Einflussfaktoren, wie z.B. ein aktueller depressiver Affekt (z.B. Reid & MacLulich, 2006), bedeutsam auf Antworttendenzen auswirken.

Eine der wenigen Arbeiten, die im höheren Erwachsenenalter das deklarative Metagedächtnis als Wissen und nicht als Selbsteinschätzung misst, stellt die Studie von Schneider und Uhl (1990) dar, die Drittklässler, jüngere und ältere Erwachsene kontrastiert. Angelehnt an den Test von Belmont und Borkowski (1988) wurde das Wissen über die Effektivität verschiedener Strategien erfasst. Insgesamt gesehen zeigten sich wenig Unterschiede im Vergleich von

jüngeren und älteren Erwachsenen; lediglich beim Wissen über semantische Kategorisierungsstrategien schnitten die älteren Erwachsenen schlechter ab als die jüngeren.

Die berichteten Befunde weisen also darauf hin, dass die wahrgenommene Verschlechterung der Gedächtnisleistung bei älteren Menschen nicht immer mit der tatsächlichen Veränderung korrespondiert, sondern dass hier implizite Theorien einen großen Beitrag leisten. Außerdem wird die Notwendigkeit deutlich, Messverfahren zu entwickeln, mit denen über die Lebensspanne hinweg das deklarative Metagedächtnis als Wissen über Personen-, Strategie- und Aufgabenmerkmale erfasst werden kann. Denn da auch bei jüngeren Erwachsenen häufig keine Deckeneffekte im deklarativen Metagedächtnis nachgewiesen werden können (vgl. Schneider, 2015) und sich die Entwicklung auch immer im Wechselspiel mit anspruchsvollen Lern- und Gedächtnisaufgaben vollzieht (Schneider, 2010), ist zu vermuten, dass feinere Messinstrumente auch differenzierte Veränderungen und u.U. auch Defizite im höheren Erwachsenenalter finden können.

2.3 Zusammenhang mit der Gedächtnisleistung

Da der Ausgangspunkt für die Auseinandersetzung mit dem Metagedächtnis dessen angenommener Einfluss auf die Gedächtnisleistung war (Flavell, 1971), existieren auch seit dem Beginn der Metagedächtnisforschung Studien, die sich diesem Zusammenhang widmen. Womöglich auch wegen der heterogenen Definitionen von Metagedächtnis sprachen die Befunde zunächst nicht für eine Korrelation zwischen Metagedächtniswissen und Leistung (Cavanaugh & Perlmutter, 1982). In späteren Metaanalysen auf einer breiteren Datenbasis und mit exakteren Definitionen wurden jedoch substantielle, wenn auch eher moderate Zusammenhänge zwischen Metagedächtnis und Erinnerungsleistung gefunden (z.B. Schneider, 1985, 1989). In Bezug auf das deklarative Metagedächtnis fielen die Zusammenhänge etwas geringer aus als für das prozedurale Metagedächtnis. Weiterhin waren die Korrelationen zwischen Metagedächtniswissen und Leistung im Vorschulalter und zu Beginn der Grundschule niedriger als bei älteren Kindern und Jugendlichen (Schneider, 1989). Belege für die Wirkrichtung des Zusammenhangs liefern Studien, denen multivariate Analysen zugrunde liegen. So erbrachten z.B. Schneider, Schlagmüller und Visé (1998) bei Dritt- und Viertklässlern mithilfe eines Strukturgleichungsmodells sowohl den Nachweis für einen kleinen, direkten Effekt vom deklarativen Metagedächtnis auf die Erinnerungsleistung, aber auch für einen deutlich größeren indirekten Effekt, der über die Strategienutzung vermittelt wird. Latente Wachstumskurvenmodelle aus der oben bereits erwähnten Studie von Grammer

et al. (2011) zeigen im Längsschnitt für die Klassenstufen 1 und 2, dass der Erwerb von Metagedächtniswissen dem von Kategorisierungsstrategien zeitlich vorausgeht und dieses zu verschiedenen Zeitpunkten vorhersagt (vgl. auch Schlagmüller & Schneider, 2002). Trotz dieser Belege dafür, dass der Zusammenhang von (deklarativem) Metagedächtnis und Gedächtnisleistung im Wesentlichen durch den Einfluss des Metagedächtniswissens auf einen effektiven Einsatz von Strategien und eine nachfolgend bessere Leistung erklärt werden kann, sind wechselseitige Beeinflussungen nicht grundsätzlich auszuschließen. Denn auch Feedback aus Lern- und Erinnerungsprozessen führt dem deklarativen Metagedächtnis im Umkehrschluss neue Informationen zu (vgl. Schneider, 2015).

2.4 Fazit

Zusammenfassend lässt sich sagen, dass das Wissen über mentale Verben als Vorläuferfertigkeit und grundlegendes metakognitives Wissen schon bei Kindergartenkindern vorhanden sind. Dieses nimmt dann im Verlauf der Grundschul- und frühen Sekundarschulzeit weitgehend linear zu (vgl. Schneider & Lockl, 2006). Die vorwiegend domänenspezifische Erfassung des deklarativen Metagedächtnisses in der Sekundarschulzeit zeigt nach einem ersten weiteren Anstieg ein Abflachen der Wachstumskurve (Karlen et al., 2014; van der Stel & Veenman, 2014). Dennoch ist auch im jüngeren Erwachsenenalter noch von einem Verbesserungspotenzial auszugehen (Schneider, 2015). Für das höhere Erwachsenenalter lässt die aktuelle Studienlage keine weiter reichenden Schlüsse über die Entwicklung im Metagedächtniswissen zu. Befunde bezüglich der Selbsteinschätzung zeigen, dass das negative gedächtnisbezogene Selbstbild älterer Probanden häufig nicht mit den tatsächlichen Leistungen korrespondiert (vgl. Hertzog & Hultsch, 2000).

Insgesamt betrachtet kann die Bedeutung des deklarativen Metagedächtnisses als ein wichtiger Faktor für die Erklärung von kognitiven Leistungen als erwiesen angesehen werden. Nicht nur bei Gedächtnisaufgaben im Labor, sondern auch im Schulkontext und für verschiedene Altersgruppen zeigte sich, dass das Abschneiden in Tests zum metakognitiven Wissen die Schulleistungen z.T. intelligenzunabhängig vorhersagt (Artelt et al., 2012; van der Stel & Veenman, 2014).

3. PROZEDURALES METAGEDÄCHTNIS

Im Zentrum der eigenen Arbeit steht die Entwicklung verschiedener Aspekte des prozeduralen Metagedächtnisses über die Lebensspanne. Dabei wird auf das Modell von Nelson und Narens (1990, 1994) zurückgegriffen, welches Überwachungs- von Steuerungsprozessen trennt (vgl. Abschnitt 1.2.5). Da Steuerungsprozesse lediglich für die Studien 5 und 6 eine Rolle spielen, soll der Überwachung auch im Theorieteil eine größere Bedeutung zukommen. Im Folgenden wird die Messung von Überwachungsprozessen sowie die Entwicklung der einzelnen Facetten vom Vorschul- bis zum höheren Erwachsenenalter ausführlich beschrieben, bevor als Beispiel für Kontrollprozesse die Lernzeitallokation sowohl aus allgemein- als auch aus entwicklungspsychologischer Sicht vorgestellt wird.

3.1 Metakognitive Überwachung

Bevor Ergebnisse zu Leistungen in der metakognitiven Überwachung in den einzelnen Altersgruppen dargestellt werden, sollen zunächst grundsätzliche Erfassungs- und Analysemethoden kontrastiert werden.

3.1.1 Erfassungsmethoden

Im Gegensatz zum deklarativen Metagedächtniswissen wird die metakognitive Überwachung immer im Zusammenhang mit konkreten, aktuell zu bearbeitenden Lern- und Erinnerungsaufgaben, also „online“ bzw. abhängig (Cavanaugh & Perlmutter, 1982) erfasst. Es haben sich in den gut 40 Jahren der Metagedächtnisforschung verschiedene Methoden dazu etabliert: Hier sind u.a. Protokolle des Lauten Denkens zu nennen (z.B. Bannert & Mengelkamp, 2008) oder die Erfassung von Fehlerentdeckungsraten (z.B. L. Baker, 1984). Bei beiden handelt es sich um Messverfahren, die vorwiegend in anwendungsbezogenen Feldern, wie z.B. der Forschung zum Leseverständnis, eingesetzt werden und besonders im Fall von Denkprotokollen nicht eindeutig zwischen Überwachungs- und Steuerungsprozessen unterscheiden.

Neben diesen hat sich – obwohl es nicht als erschöpfend bezeichnet werden kann – das oben beschriebene Modell zu den Maßen des prozeduralen Metagedächtnisses von Nelson und Narens (1990, 1994; vgl. Abschnitt 1.2.5) durchgesetzt, auf dem auch die eigenen Studien beruhen. In Bezug auf die Überwachungsmaße werden hier EOLs, JOLs, FOKs und SUs

benannt und zeitlich im Informationsverarbeitungsprozess verortet. Prinzipiell können auch nicht explizit von Nelson und Narens (1990, 1994) einbezogene Maße in deren Modell integriert werden. So benennt beispielsweise die Forschung zum Überwachen des Textverstehens (*Metacomprehension*) auch Maße, die erfragen, wie sicher eine Person ist, dass sie Informationen aus einem zusammenhängenden Text richtig verstanden hat, die analog zu den erwähnten Urteilen ebenfalls zeitlich vor, während und nach dem Lesevorgang eingeordnet werden können (z.B. Dunlosky, Rawson & Hacker, 2002).

Die Güte des Überwachungsprozesses wird durch den Vergleich der subjektiven mit der tatsächlichen Leistung bestimmt. Sowohl die Abfrage der Urteile als auch die Qualität der Überwachung kann entweder für jedes zu lernende Item einzeln oder aggregiert für den gesamten Lernvorgang erfolgen (z.B. van Overschelde & Nelson, 2006). Daraus leiten sich zwei grundsätzlich verschiedene Berechnungsmethoden ab. Zum einen kann die *relative Genauigkeit* der Überwachungsurteile bestimmt werden (*Resolution*). Diese beinhaltet eine Aussage darüber, ob die vorhergesagte Leistung bei einem Item im Vergleich zur vorhergesagten Leistung bei einem anderen Item mit der Rangfolge dieser Items bei der beobachteten Leistung übereinstimmt (Nelson, 1984, 1996). „Relativ“ bezieht sich hier also darauf, dass die Vorhersagegüte bei einem Item in Relation mit anderen Items beurteilt wird. Zur Berechnung der relativen Genauigkeit von Überwachungsmaßen kommen Goodman-Kruskal-Gammakorrelationen zum Einsatz. Dieses Zusammenhangsmaß trifft eine Aussage über die Assoziation zwischen zwei mindestens ordinalskalierten Variablen; es setzt sich zusammen aus dem Vergleich zwischen Konkordanz, d.h. Fällen, in denen bei der Kontrastierung von zwei Items dasjenige ein höheres Überwachungsurteil erhält, das auch besser erinnert wird, und Diskordanz, d.h. Fällen, in denen das Item ein höheres Überwachungsurteil erhält, das später schlechter erinnert wird (vgl. Nelson, 1984). Gammakorrelationen können Werte zwischen -1 (sehr hohe relative Ungenauigkeit) und +1 (sehr hohe relative Genauigkeit) annehmen. Der Vorteil von Gammakorrelationen gegenüber anderen Maßen kann u.a. in der Unabhängigkeit von der Gesamtleistung, der Anwendbarkeit auf verschiedene Leistungsmaße (Wiedererkennen, freies Erinnern) und bei Rangbindungen gesehen werden (Gonzales & Nelson, 1996). Außerdem kann die relative Genauigkeit auch über die Differenzierungsfähigkeit zwischen erinnerten und nicht erinnerten Items berechnet werden. Hierzu werden die mittleren Überwachungsurteile vor bzw. nach richtigen Antworten mit denen vor bzw. nach falschen oder nicht erinnerten Antworten verglichen. Je höher die Differenz, desto besser ist die Qualität der Überwachung. Auch hier wird also die metakognitive Leistung einer Person im Vergleich einzelner Items miteinander betrachtet. Die Interpretation

erlaubt – im Gegensatz zu Gammakorrelationen – Aussagen über die absolute Höhe der Urteile und kann deshalb als ergänzend zu diesen gesehen werden (vgl. z.B. Krebs & Roebers, 2012; von der Linden, Schneider & Roebers, 2011).

Eine zweite Methode zur Bestimmung der Güte von Überwachungsvorgängen stellen die sog. *Kalibrierungsmaße* dar. Mit diesen können Aussagen über die *absolute* Genauigkeit metakognitiver Urteile getroffen werden. Hierbei wird die Höhe der eigenen Leistungseinschätzung mit der Höhe der tatsächlichen Leistung kontrastiert. Dies kann durch die Differenz eines globalen Urteils mit der Gesamtleistung geschehen; so entspricht dann z.B. eine Vorhersage, 70% der Aufgaben richtig lösen zu können, und darauf folgend 70% korrekter Antworten einer perfekten Kalibrierung (z.B. Maki, Foley, Kajer, Thompson & Willert, 1990). Alternativ können auch Vorhersagen für einzelne Lerneinheiten mit der darauffolgenden Leistung verglichen werden (z.B. Stone, 2000) oder aber auch Kalibrierungskurven auf Itemebene analysiert werden (z.B. Connor, Dunlosky & Hertzog, 1997). Der Unterschied zwischen Resolutions- und Kalibrierungsmaßen besteht also vor allem darin, dass bei ersteren einzelne Items verglichen werden, bei letzteren absolut gesehen die Vorhersage mit der Gesamtleistung in Bezug gesetzt wird.

Der Vergleich zwischen beiden Maßen zeigt, dass diese nicht immer deckungsgleich sind. Maki, Shields, Wheeler und Zacchilli (2005) zeigten z.B. für Studenten, dass die absolute Genauigkeit deutlich stärker in Abhängigkeit der Aufgabenschwierigkeit variierte. Die Teilnehmer schienen also ihre (globalen) Urteile nicht ausreichend an den konkreten Kontext anzupassen und ihre eigene Leistung eher konstant einzuschätzen. Da die absolute Höhe der Einschätzungen keinen Einfluss auf die Gammakorrelationen hat, fiel die relative Genauigkeit insgesamt besser aus. Für den Vergleich von älteren und jüngeren Erwachsenen ergab die oben erwähnte Studie von Connor et al. (1997), dass sich die Gammakorrelationen beider Gruppen nicht unterschieden. Die absolute Genauigkeit wurde aber von der Höhe der tatsächlichen Erinnerungsleistung beeinflusst und fiel deshalb bei den älteren Probanden etwas geringer aus.

Es muss an dieser Stelle auch darauf hingewiesen werden, dass dieser Absatz die möglichen Berechnungsmethoden nicht umfassend auflisten kann. Schon 1984 stellte Nelson acht verschiedene Maße für die Auswertung von FOK-Urteilen vor. Schraw (2009) unterschied zwischen fünf Möglichkeiten die Überwachungsfähigkeit zu analysieren (absolute und relative Genauigkeit, Urteilstendenzen und Streuungen sowie die Diskriminierungsfähigkeit). Zunehmend wird auch der aus der Entscheidungstheorie stammende Signalentdeckungsansatz

zur Bestimmung der Güte metakognitiver Überwachung in die wissenschaftliche Diskussion eingebracht (z.B. Fleming & Dolan, 2012).

Für die vorliegenden Studien sollen Maße zur Bestimmung der relativen Genauigkeit zum Einsatz kommen. Hierbei sollen sowohl Gammakorrelationen als auch der Vergleich der mittleren Überwachungsurteile vor bzw. nach richtigen vs. falschen Antworten berechnet werden. Erstere erlauben eine recht präzise und von der absoluten Leistung unbeeinflusste Aussage über die Überwachungsleistung; weiterhin sind sie wegen ihrer verbreiteten Anwendung gut mit anderen Studien vergleichbar. Die Berechnung der Differenzierungsfähigkeit gibt zusätzlich Informationen über die Höhe der Urteile, die in den Gammakorrelationen nicht mehr enthalten ist. Ein Kalibrierungsmaß zur globalen Einschätzung des eigenen Textverständnisses soll ergänzend in Studie 3 zum Einsatz kommen. Davon abgesehen wird auf Kalibrierungsmaße verzichtet, da unklar ist, ob auch Grundschulkinder diese eher abstrakte Wahrscheinlichkeitsschätzung vornehmen können, und da – wie oben dargelegt – diese auch von Leistungs- oder Altersfaktoren stärker beeinflusst sind. Dieser Aspekt ist gerade für die eigenen Studien relevant, da hier eine große Altersspanne und, in den Studien zum Einfluss von Expertise, verschiedene Leistungsgruppen verglichen werden. In der Darstellung der Ergebnisse wird dann von „Genauigkeit“ gesprochen, wenn auf Gammakorrelationen Bezug genommen wird, und auf „Differenzierungsfähigkeit“, wenn die Höhe der Überwachungsurteile verglichen wird.

3.1.2 Ease-of-Learning-Urteile

Beginnend bei EOLs als zeitlich erstem Urteil sollen nun jeweils zu den einzelnen Überwachungsmaßen aus dem Modell zum prozeduralen Metagedächtnis von Nelson und Narens (1990,1994) allgemeinspsychologische Befunde vorgestellt werden, um auf allgemeine Effekte und Grundlagen der Urteile einzugehen, aber auch um den Stand bei jüngeren Erwachsenen darzustellen, der dann als Vergleichsmaßstab dient. Anschließend werden die Entwicklungsverläufe im Kindes-, Jugend- und höheren Erwachsenenalter beschrieben.

3.1.2.1 Allgemeinspsychologische Befunde

Noch bevor der Terminus „Metagedächtnis“ von Flavell (1971) geprägt und verschiedene Maße der prozeduralen Komponente von Nelson und Narens (1990, 1994) in einem einzigen Modell zusammengefasst wurden, gab es erste Experimente zu Schwierigkeitseinschätzungen von

Items. Diese wurden zwar noch nicht unter den Begriff „Ease-of-Learning“-Urteile (EOLs) gefasst, sind aber aus heutiger Sicht als solche zu verstehen. So ließen Richardson und Erlebacher (1958) Studenten Wortpaare und einzelne Silben hinsichtlich ihrer semantischen Ähnlichkeit und ihrer Schwierigkeit beurteilen und fanden hohe Übereinstimmungen zwischen diesen Maßen. Der prädiktive Wert dieser Einschätzungen wurde von den Autoren jedoch nicht erfasst. Erst Underwood (1966) zeigte, dass Annahmen über die Schwierigkeit von Items (z.B. Assoziativität, Lernschwierigkeit, Ausspracheschwierigkeit) hoch mit späteren Lern- und Erinnerungsleistungen korreliert sind. Demnach wurden beispielsweise als leicht bzw. stark zusammenhängend eingeschätzte Items nach einer festgesetzten Lernzeit auch besser erinnert. Für Paradigmen mit frei eingeteilter Lernzeit postulierten Nelson und Leonesio (1988), dass dieser Zusammenhang hier gegen Null gehen sollte, da die Probanden schwierige Items länger lernen sollten und sich dadurch die Gedächtnisleistung angleichen sollte. Diese *Monitoring-Neutralization-Hypothese* konnte jedoch experimentell nicht belegt werden. Die gefundenen Korrelationen lagen zwar mit Werten im mittleren Bereich deutlich niedriger als bei Underwood (1966), sind aber dennoch als bedeutsam anzusehen. Die Autoren interpretierten dieses Ergebnis dahingehend, dass das metakognitive Urteil zum Zeitpunkt seiner Abgabe zutreffend ist, aber durch die Speicherung im Kurzzeit- und nicht im Langzeitgedächtnis für die spätere Lernphase nicht hinreichend valide ist (Nelson & Leonesio, 1988). Kelemen, Frost und Weaver (2000) zeigten für die Einschätzung von Englisch-Suaheli-Wortpaaren, die später mit einer festgelegten Lernzeit dargeboten wurden, auch nur mittlere bis niedrige Genauigkeiten der EOLs. Insgesamt gesehen konnten also für das Lernen von Buchstabenkombinationen und Paarassoziationen die zunächst bei Underwood (1966) gefundenen sehr großen Zusammenhänge nicht belegt werden (für ein ähnliches Ergebnis s. auch Leonesio & Nelson, 1990).

Studien, die EOLs beim Lernen von komplexeren Materialien, also z.B. Informationen aus Texten, erfassen, sind noch spärlicher. Hier fällt auf, dass nicht in jedem Bericht erfasst wurde, wie die EOLs mit der späteren Lernleistung zusammenhängen (z.B. Thiede, Anderson & Theriault, 2003). Außerdem divergiert die exakte Erfassung der EOLs in den wenigen vorhandenen Studien voneinander. Burkett und Azevedo (2012) präsentierten ihren Probanden beispielsweise offene Fragen, die später anhand der Lektüre eines Textes und der Interpretation von Grafiken beantwortet werden sollten, und ließen sie beurteilen, wie leicht oder schwer es ihnen fallen würde, die benötigte Information zur Beantwortung dieser Frage zu lernen. Es ergab sich kein Zusammenhang dieser Einschätzung mit der anschließenden Leistung. Bei Thiede et al. (2003) beurteilten die Teilnehmer anhand des Titels eines Textes dessen

Lernschwierigkeit. Shiu und Chen (2013) griffen dieses Design auf, gaben den Probanden aber zusätzlich die Gelegenheit, die dargebotenen Texte jeweils kurz zu überfliegen, bevor sie zu jedem ein EOL-Urteil abgaben. Vergleichbar mit Burkett und Azevedo (2012) fanden auch Shui und Chen (2013) sehr niedrige Gammakorrelationen (zwischen $-.06$ und $.19$).

Die hier vorgestellten Ergebnisse lassen, da insgesamt nur wenige Studien vorliegen, keine weitreichenden Schlussfolgerungen bezüglich des prädiktiven Wertes von EOLs zu. Grundsätzlich scheinen jedoch die eher hohen Korrelationen, die in der ersten Studie gefunden wurden (Underwood, 1966), nach unten korrigiert werden zu müssen. Bei Studien zum Textlernen fallen die Zusammenhänge der EOLs mit der späteren Lern- und Erinnerungsleistung noch niedriger aus als bei Studien zum Paar-Assoziationslernen. Insgesamt liegt die Vermutung nahe, dass die Komplexität der Items die Höhe der Korrelationen beeinflusst. So bildeten bei Underwood (1966) Trigramme das Stimulusmaterial; spätere Studien, die niedrigere EOL-Leistungszusammenhänge fanden, nutzten in der Regel konkrete Substantive (z.B. Leonesio & Nelson, 1990). Bei Studien zum Textlernen ist das Lernmaterial noch einmal deutlich komplexer (z.B. Burkett & Azevedo, 2012).

Die Frage nach der Grundlage von EOLs, also den kognitiven Prozessen beim Fällen eines solchen Urteils, muss als weitgehend ungeklärt gelten. Im Gegensatz zur JOLs liegt hierzu wesentlich weniger Literatur vor. Schon Leonesio und Nelson (1990) argumentierten, dass metakognitive Überwachungsurteile auf multidimensionalen Strukturen, also auf verschiedenen Faktoren, beruhen. Ähnlich wie für JOLs wird auch für EOLs ein *Cue-utilization-Ansatz* (Koriat, 1997) diskutiert, demzufolge metakognitive Urteile inferentiell auf der Grundlage verschiedener Hinweisreize (*cues*) abgegeben werden (Jönsson & Lindström, 2010; für eine ausführlichere Darstellung des Cue-utilization-Ansatzes s. Abschnitt 3.1.3.1). Da EOLs vor dem eigentlichen Lernvorgang erhoben werden, sollten sie vorrangig auf intrinsischen Hinweisreizen beruhen. Denn dies sind Hinweisreize, die a priori für die zu lernenden Items charakteristisch sind. Mögliche Hinweisreize für EOLs können also – wie oben beschrieben – die Komplexität der Items, aber auch andere, wie die Vertrautheit der Probanden mit den Items, die Konkretheit oder die Assoziativität der Stimuli (vgl. Underwood, 1966) sein. Jönsson und Lindström (2010) untersuchten experimentell die Rolle intrinsischer Hinweisreize für EOLs; sie erbrachten Nachweise dafür, dass die Wortlänge, die Häufigkeit dieser Wörter im Schwedischen und die Gegenständlichkeit (*concreteness*) als Anhaltspunkte für die Höhe der Urteile genutzt werden. Vergleichbar dazu fanden auch Mazzoni, Cornoldi, Tomat und Vecchi (1997), dass die EOLs von Probanden beim Memorieren einer Einkaufsliste davon beeinflusst wurden, wie typisch ein Item ist und wie häufig es gekauft wird, obwohl sich

diese Merkmale nicht bzw. nur geringfügig auf die tatsächliche Erinnerungsleistung der Probanden auswirkten. Extrinsische Faktoren wie Wiederholung oder Dauer der Darbietung spielen zu diesem Zeitpunkt noch keine Rolle, auch mnemonische Faktoren wie die Vertrautheit mit den Items oder die Leichtigkeit der Verarbeitung können nur zum Teil in EOLs einfließen (Koriat, 1997).

Für Fragestellungen zum Textlernen fehlen bislang Studien, die den Einfluss verschiedener Hinweisreize auf die Höhe und Genauigkeit der EOLs untersuchen. Neben den erwähnten intrinsischen Faktoren, wie die erwartete Komplexität des Textes, können hier auch das Vorwissen oder Annahmen über den eigenen Expertenstatus in Bezug auf das jeweilige Thema eine Rolle spielen. Diese Frage soll in Studie 3 aufgegriffen werden.

Insgesamt gesehen weist die Forschung zu EOLs noch etliche Lücken auf. Es kann als weitgehend gesichert gelten, dass Probanden in allgemeinspsychologischen Studien grundsätzlich in der Lage sind, die Schwierigkeit von Items für einen späteren Lernprozess zu beurteilen. Die Genauigkeit dieser Einschätzung lässt jedoch noch viel Raum zur Verbesserung. Weiterhin fehlen Studien, die den Einfluss verschiedener Hinweisreize auf EOLs genauer aufschlüsseln. Erste Ansätze aus Paar-Assoziationsdesigns zeigen, dass intrinsische Reize – wie auch theoretisch postuliert – in die Höhe, aber nicht zwangsläufig auch in die Genauigkeit der EOLs einfließen (Jönsson & Lindström, 2010); doch für komplexere Stimuli, aber auch für die gesamte Bandbreite der möglichen Hinweisreize ist die Datenlage bislang unzureichend.

3.1.2.2 Entwicklungspsychologische Befunde

Entwicklung im Vor- und Grundschulalter. Im Gegensatz zur Datenlage bei Erwachsenen sind EOLs bei Kindern recht gut erforscht. Das gängige Untersuchungsparadigma weicht hier etwas von dem bei Erwachsenen ab. In der Regel werden Kinder aufgefordert, ihre Gedächtnisspanne vorherzusagen, indem sie immer längere Itemlisten vorgelegt bekommen und angeben sollen, ob sie eine Liste dieser Länge noch auswendig lernen könnten. Anschließend wird die Vorhersage mit der tatsächlichen Erinnerungsleistung verglichen. Das heißt, auch hier findet eine Leistungsvorhersage ohne aktuelle Lernerfahrung statt, allerdings nicht auf Itemebene, sondern bezogen auf die Gesamtheit der Items. Eine der ersten Studien hierzu stammt von Flavell et al. (1970), die die Einschätzungen von Kindergarten- und Vorschulkindern sowie Zweit- und Viertklässlern miteinander verglichen. Insgesamt überschätzten die Teilnehmer aller Altersgruppen ihr Erinnerungsvermögen. Es zeigte sich

jedoch, dass die vorhergesagte Gedächtnisspanne bei den beiden älteren Gruppen niedriger ausfiel als bei den beiden jüngeren; die objektive Leistung war jedoch bei den jüngeren Teilnehmern deutlich geringer. Nachfolgende Arbeiten replizierten diese Befunde im Wesentlichen (z.B. Worden & Sladewski-Awig, 1982; Yussen & Levy, 1975), sodass die Annahme nahe liegt, dass die deutliche Überschätzung mit etwa Beginn der Grundschule abnimmt.

Verschiedene Gründe sind für die beobachtete überoptimistische Einschätzung der eigenen Gedächtnisleistung bei Kindergarten- und Vorschulkindern denkbar. Schneider (1998) sowie Visé und Schneider (2000) fanden bei Kindergarten- und Grundschulkindern genauere Vorhersagen der eigenen Leistung bei bekannten und auch einfacheren Aufgaben (z.B. Weitsprung) als bei neuen Aufgaben (z.B. Wortspanne). Für eine ähnliche Altersgruppe wiesen Cunningham und Weaver (1989) Vorteile hinsichtlich der Selbsteinschätzung nach, wenn diese nicht explizit verbalisiert werden musste, sondern nonverbal angezeigt werden konnte. Wenn also die Aufgabe einfach genug konzipiert wird, ist die Tendenz zur Überschätzung bei jüngeren Kindern etwas geringer ausgeprägt. Zudem scheint im Kindergarten- und frühen Grundschulalter das *gewünschte* Ergebnis stark in die Höhe der Vorhersage einzufließen. In der bereits erwähnten Studie von Visé und Schneider (2000) waren auch die jüngeren Teilnehmer in der Lage, ihre Leistung nach der Ausführung der Aufgabe recht akkurat einzuschätzen; allerdings unterschieden die Vier- und Sechsjährigen im Gegensatz zu den Neunjährigen bei ihren Leistungsprognosen nicht zwischen gewünschtem und erwartetem Ergebnis, d.h. ihr Wunschdenken hatte als motivationaler Faktor einen stärkeren Einfluss auf die Vorhersage als das Resultat der metakognitiven Überwachung vorheriger Aufgaben (vgl. auch Stipek, 1984). Shin, Bjorklund und Beck (2007) überprüften den potenziellen adaptiven Nutzen der Überschätzung jüngerer Kinder experimentell und beobachteten, dass Kinder, die ihre Leistung stärker überschätzten, sich in nachfolgenden Aufgaben auch stärker verbesserten, was die Autoren mit einer höheren Motivation und Ausdauer bei überoptimistischen Kindern begründen. Es kann also nicht davon ausgegangen werden, dass in dieser Altersgruppe ein allgemeines metakognitives Defizit vorliegt, wie zunächst postuliert wurde (Flavell et al., 1970), sondern dass das Maß der Überschätzung durch Faktoren wie der Vertrautheit mit der Aufgabe, der Reduktion von kognitiver Beanspruchung und motivationalen Einflüssen maßgeblich mitbestimmt wird.

Studien zu EOLs mit älteren Kindern sind deutlich seltener. Hier ist lediglich die oben berichtete Arbeit von Visé und Schneider (2000) zu nennen, die Kinder bis zum Alter von neun Jahren integrierte.

Entwicklung im Jugendalter. Studien mit Jugendlichen sind in diesem Zusammenhang ebenfalls rar. Labuhn, Zimmerman und Hasselhorn (2010) zeigten für eine Stichprobe von Fünftklässlern, dass Schüler, die sich in ihren Vorhersagen stark überschätzten, von Leistungsrückmeldungen profitierten, und dass dies sogar leichte Effekte auf die Leistung in der vorliegenden Aufgabe hatte. Weitere Untersuchungen zur Qualität von EOLs bei Jugendlichen existieren nur noch in Bezug auf den Vergleich von Probanden mit und ohne Lernbehinderung. Die Arbeiten von Klassen (Job & Klassen, 2012; Klassen, 2007) zeigen, dass sich Jugendliche mit einer Lernbehinderung stärker überschätzten als eine Vergleichsgruppe. Dies war sowohl für schulnahe Aufgaben wie Rechtschreiben und Lesen der Fall, als auch für nicht akademische Anforderungen, wie z.B. einen Ball werfen (Job & Klassen, 2012). Für beide Studien muss einschränkend bemerkt werden, dass das Ausgangsniveau der Leistungen jeweils nicht zwischen den beiden untersuchten Gruppen angepasst wurde, sodass Überschätzungen der Schüler mit Lernbehinderung auch auf eine Tendenz zur Mitte zurückzuführen sein können. Interessanterweise reagierten jedoch die Schüler mit einer Lernbehinderung bei der praktischen Aufgabe stärker auf den Anstieg der Schwierigkeit; die Überschätzung wurde also bei dieser intelligenzferneren Aufgabe noch größer, sodass darauf geschlossen werden kann, dass die überoptimistische Selbsteinschätzung bei lernbehinderten Schülern z.T. aufgabenunabhängig war. Über einen möglichen adaptiven Wert dieser Überschätzung existieren für diese Gruppe bislang keine empirischen Untersuchungen.

Entwicklung im höheren Erwachsenenalter. Für die älteren Erwachsenen ist die Datenlage zu EOLs ebenfalls nicht so umfangreich wie für jüngere Kinder, jedoch deutlich besser als für Jugendliche. Wie in Kapitel 2 zum deklarativen Metagedächtnis beschrieben, ist die aufgabenunabhängige Selbsteinschätzung älterer Probanden von negativen Altersstereotypen beeinflusst. Ob sich dies auch in den Urteilen zu konkreten Aufgaben zeigt, wurde mehrfach untersucht. Hertzog et al. (1990) sowie Hertzog, Saylor, Fleece und Dixon (1994) verglichen Fragebogendaten zur *Memory self-efficacy* (MSE; d.h. deklarative Annahmen über die eigenen kognitiven Fähigkeiten) mit globalen Prädiktionen zu konkreten Lern- und Erinnerungsaufgaben verschiedener Formate. Sowohl für jüngere als auch für ältere Erwachsene fanden sie substantielle Korrelationen zwischen den MSE-Werten und der Höhe der Leistungsvorhersage, sodass davon ausgegangen werden kann, dass die Vorhersagen zumindest z. T. auf dem allgemeinen gedächtnisbezogenen Selbstkonzept beruhten (Hertzog et al., 1994). Generell fanden die Autoren in mehreren Telexperimenten in beiden Altersgruppen eher eine Unterschätzung, die bei den älteren Probanden sogar noch geringer ausfiel als bei den

jüngeren. In einigen Teilstudien setzten die Autoren vor der Selbsteinschätzung einen Anker im mittleren Bereich („Im Schnitt erinnern sich Probanden an x Wörter“). Die höhere Genauigkeit der Leistungsvorhersage bei älteren Erwachsenen lässt sich in diesen Experimenten auch damit erklären, dass ihre objektive Leistung näher an diesem Anker lag; doch auch ohne einen Referenzwert orientierten sich beide Altersgruppen in ihren Vorhersagen an einem mittleren Wert. Neben diesen Befunden, die eher für eine höhere Genauigkeit bei älteren Erwachsenen in der Leistungsvorhersage im Sinne einer geringeren Überschätzung sprechen, finden sich auch Studien, die eher eine Überschätzung der eigenen Gedächtnisleistung bei älteren Probanden finden (z.B. Rabbit & Abson, 1991), solche, die inkonsistente Ergebnisse hervorbringen (z.B. Devolder, Brigham & Pressley, 1990; Lachman & Jelalian, 1984), und schließlich solche, die keine Unterschiede in der Vorhersagegenauigkeit zwischen jüngeren und älteren Erwachsenen nachweisen können (z.B. Woo, Schmitter-Edgecombe & Fanher, 2008). Insgesamt gesehen scheint es notwendig, neben den aufgabenunabhängigen, meist eher negativen Annahmen über eigene kognitive Fähigkeiten bei älteren Probanden weitere Einflussfaktoren auf die Leistungsvorhersagen wie die Aufgabenschwierigkeit, die Vertrautheit mit der Aufgabe, den Einfluss von Vorwissen oder Attributionsmuster zu berücksichtigen. Bei Lachman und Jelalian (1984) machten jüngere und ältere Teilnehmer jeweils bei der Aufgabenart genauere Vorhersagen, bei der sie schließlich auch besser abschnitten, überschätzten sich hier also weniger stark. Von der Vertrautheit mit der Aufgabe profitierten bei Hertzog et al. (1990) nur die jüngeren Erwachsenen hinsichtlich ihrer Vorhersagen. Bezüglich der Attributionsmuster wurden keine Unterschiede zwischen jüngeren und älteren Erwachsenen gefunden (Lachman & Jelalian, 1984). Studien zum Einfluss des Vorwissens auf die Genauigkeit der Leistungsvorhersage bei älteren Probanden liegen bislang noch nicht vor, obwohl diese Frage durchaus von Relevanz ist, da davon auszugehen ist, dass hier negativen Effekte von Altersstereotypen abgemildert werden können.

Insgesamt gesehen weist die Datenlage für Leistungsvorhersagen bei älteren Erwachsenen also noch einige Lücken auf. Es kann bislang lediglich festgehalten werden, dass die teils widersprüchlichen Befunde nicht für eine allgemeine Verschlechterung dieser Überwachungsfähigkeit sprechen (vgl. Hertzog et al., 1994). Erste längsschnittliche Ergebnisse über fünf Jahre sprechen auch nicht dafür, dass die Genauigkeit über diesen Zeitraum hinweg abnimmt (Woo et al., 2008), sodass generell ab dem Erwachsenenalter wohl von einer recht stabilen Fähigkeit gesprochen werden kann.

3.1.2.3 Fazit

Zusammenfassend lässt sich sagen, dass – obwohl erste Studien sehr hohe Zusammenhänge zwischen EOLs und späterer objektiver Leistung fanden – auch bei jüngeren Erwachsenen noch Verbesserungspotenzial hinsichtlich der Genauigkeit der Einschätzungen vorhanden ist, da die Korrelationen zumeist im mittleren Bereich liegen.

Studien zum Entwicklungsverlauf von EOL-Einschätzungen zeigen, dass besonders jüngere Kinder eine sehr optimistische Einstellung zu ihren eigenen Fähigkeiten haben. Die ungenauen Leistungsvorhersagen sind jedoch nicht durch metakognitive Defizite schlüssig zu erklären; es scheinen vielmehr motivationale Faktoren eine Rolle zu spielen, wie z.B. der Einfluss von Wunschenken. Die wenigen Studien bei Jugendlichen zeigen ebenfalls Tendenzen zur Überschätzung in Bereichen, in denen die Teilnehmer leistungsschwächer sind, was die Annahme nahe legt, dass hier ebenfalls motivationale Aspekte wirksam werden. Für den weiteren Entwicklungsverlauf über die Lebensspanne scheinen die metakognitiven Fähigkeiten bezüglich der Leistungsvorhersage recht stabil zu bleiben; Unterschiede in den zitierten Befunden sind auf wechselnde Aufgabenschwierigkeiten, Unterschiede in den Ausgangsleistungen der untersuchten Gruppen oder die Vertrautheit mit den Aufgaben zurückzuführen.

An dieser Stelle muss einschränkend angemerkt werden, dass die Studien zu Entwicklungsverläufen in EOLs nicht direkt mit allgemeinspsychologischen Studien vergleichbar sind, da hier in der Regel globale Leistungsvorhersagen anstelle von Einschätzungen auf Itemebene erfasst werden. Forschungsbedarf besteht weiterhin bezüglich der Frage nach den Grundlagen von EOLs, und das auch für verschiedene Altersstufen. Es herrscht zwar weitgehend Einigkeit darüber, dass intrinsische Hinweisreize hierfür wohl die größte Bedeutung haben (Vertrautheit mit der Aufgabe, wahrgenommene Aufgabenkomplexität). Für andere Faktoren (z.B. wahrgenommene Expertise im einschlägigen Gebiet, motivationale Faktoren) können zum jetzigen Zeitpunkt noch keine allgemeingültigen Aussagen getroffen werden. Weiterhin fehlen im gesamten Entwicklungsverlauf Studien zu EOLs bei komplexeren Materialien, die nachweisen könnten, inwieweit die Ergebnisse zu jüngeren Erwachsenen auf die anderen Altersgruppen übertragbar sind.

3.1.3 Judgments of Learning

Im zeitlichen Verlauf nach den EOLs werden die JOLs erfasst; sie werden während oder nach dem Lernvorgang erhoben und stellen eine Prognose über zukünftige Erinnerungsleistungen dar. Auch in diesem Kapitel werden JOLs zunächst allgemeinspsychologisch und bezüglich ihrer zugrunde liegenden Mechanismen betrachtet, bevor die Entwicklung über die Lebensspanne dargestellt wird.

3.1.3.1 Allgemeinspsychologische Befunde

Zu JOLs existieren ebenfalls frühe Arbeiten, die aus heutiger Sicht als Pionierarbeiten gelten können, selbst wenn hier noch nicht explizit von JOLs die Rede ist. So präsentierten z.B. Arbuckle und Cuddy (1969) Studenten kurze Listen von zusammenhängenden und nicht zusammenhängenden Wortpaaren und ließen sie einschätzen, ob sie diese zu einem späteren Zeitpunkt erinnern würden. Sie fanden zum einen, dass die Genauigkeit dieser Einschätzungen über dem Zufallsniveau lag, zum anderen, dass Unterschiede in der Assoziativität der Items als wesentliche Grundlage für die Vorhersagen herangezogen wurden. Auch für Wiedererkennensparadigmen wies Groninger (1976) einen Zusammenhang zwischen vorhergesagter und tatsächlicher Rekognitionsleistung nach. Vergleichende Aussagen über das Ausmaß der JOL-Genauigkeit können erst seit der Einführung von Goodman-Kruskal-Gammakorrelationen (Nelson, 1984) gemacht werden. Insgesamt gesehen liegen diese in den früheren Arbeiten eher im mittleren bis niedrigen Bereich; bei Leonesio und Nelson (1990) beispielsweise erreichten die Probanden in einem Paar-Assoziationsparadigma eine mittlere Gammakorrelation von etwa .30, d.h. die Teilnehmer waren durchaus in der Lage, ihre spätere Erinnerungsleistung vorherzusagen, die Qualität dieser Vorhersage lässt allerdings viel Raum für Verbesserungen (für Korrelationen ähnlicher Höhe vgl. auch Begg, Duft, Lalonde, Melnick & Sanvito, 1989). Als Konsequenz entstanden verschiedene Ansätze mit dem Ziel, die Güte der JOLs zu steigern, die gleichzeitig auch dazu beitrugen, die zugrunde liegenden Mechanismen dieser Urteile besser zu verstehen.

So zeigten Shaughnessy und Zechmeister (1992) z.B. eine deutliche Steigerung der JOL-Differenzierung zwischen später korrekten und später falschen Antworten durch zwischengeschaltete Probetests, selbst wenn die Erinnerungsleistung konstant gehalten wurde. Diese Übungsabfragen tragen zur Verbesserung der JOL-Genauigkeit bei, indem die Probanden

vorab eine interne Rückmeldung über ihre Leistung erhalten (vgl. auch Dunlosky, Rawson & McDonald, 2002).

Weiterhin existieren seit Kurzem einige wenige Studien, die den Einfluss einer effektiven Lernstrategie auf JOLs untersuchen. Im Bereich des Paar-Assoziationslernens wirkt sich besonders die Generierung interaktiver Bilder (*Interactive imagery* oder *Visual imagery*) positiv auf die Erinnerungsleistung aus (z.B. Hertzog & Dunlosky, 2006; Paivio, 1971). Der Transfer dieses Befundes von der Gedächtnisleistung auf die Metagedächtnisleistung wurde bisher nur selten erbracht. Hertzog, Sinclair und Dunlosky (2010) vermittelten den Teilnehmern ihrer Studie vor Beginn des Lernvorgangs Informationen zu der Elaborationsstrategie („Interactive imagery“) und erfragten retrospektiv, welche Strategie bei welchem Wortpaar genutzt wurde. Die Autoren konnten hier auch ohne die explizite Instruktion, die Strategie einzusetzen, zeigen, dass deren Nutzung die JOL-Genauigkeit vor allem bei nicht zusammenhängenden Itempaaren verbesserte.

Die größte Verbesserung der JOL-Genauigkeit konnte jedoch bislang durch die Verzögerung des JOL-Abrufs erzielt werden (*Delayed-JOL-effect*; Nelson & Dunlosky, 1991). Die Autoren zeigten, dass die mittlere Gammakorrelation für JOLs, die unmittelbar nach der Präsentation des jeweiligen Items erhoben wurden, bei .38 lag, für JOLs, die zeitlich ein wenig verzögert abgefragt wurden, dagegen bei .90, was also einer nahezu perfekten Vorhersageleistung entspricht. Dieser Effekt wurde erfolgreich bei verschiedenen Personengruppen, mit verschiedenen Stimuli und verschiedenen Verzögerungsintervallen repliziert (z.B. Weaver & Kelemen, 1997; Maki, 1998a; van Loon, de Bruin, van Gog & van Merriënboer, 2013), sodass seine Gültigkeit inzwischen auch metaanalytisch belegt werden konnte (Rhodes & Tauber, 2011). Trotz der weitreichenden Gültigkeit des Delayed-JOL-Effekts muss an dieser Stelle darauf hingewiesen werden, dass die nahezu perfekten Korrelationen lediglich für Aufgaben zum freien Erinnern bzw. cued-recall gefunden wurden; bei Wiedererkennens-Paradigmen liegen die mittleren Gammakorrelationen auch für verzögert erfasste JOLs in der Regel etwas niedriger, jedoch immer noch im mittleren bis hohen Bereich, da hier die Genauigkeit durch korrektes Raten bzw. korrektes Zurückweisen von falschen Disktraktoren verringert wird (Dunlosky & Nelson, 1997; Thiede & Dunlosky, 1994). Als Erklärung für den Delayed-JOL-Effekt postulierten Nelson und Dunlosky bereits in der Arbeit von 1991 die sog. *Monitoring-Dual-Memories Hypothese*, die annimmt, dass bei der Abgabe von JOLs sowohl auf Informationen aus dem Kurzzeit-, als auch auf Informationen aus dem Langzeitgedächtnis zurückgegriffen wird. Bei unmittelbaren JOLs befindet sich das entsprechende Lern-Item noch im Kurzzeitspeicher, sodass die Aussage, ob dieses Item in den Langzeitspeicher übergehen

wird, erschwert ist. Bereits eine Verzögerung von etwa 30 Sekunden bewirkt, dass sich eine Information dann im Langzeitgedächtnis befindet, was die Aussage darüber, ob dieses Item auch zum Zeitpunkt der Erinnerung im Langzeitgedächtnis verfügbar ist, deutlich erleichtert. Dieser Erklärungsansatz gilt auch heute nicht als widerlegt, wenngleich er rasch kritisiert wurde. So argumentierten Spellman und Bjork (1992), dass es vor der Abgabe von verzögerten JOLs zu Probeabrufen des gesuchten Items kommt – eine Erklärung, die meines Erachtens als Erweiterung der Monitoring-Dual-Memories Hypothese gesehen werden kann, da ein Probeabruf bei verzögerter JOL-Abfrage ja auf Informationen aus dem Langzeitgedächtnis zurückgreift (vgl. auch Dunlosky & Metcalfe, 2009). Auch der im Folgenden beschriebene Cue-utilization-Ansatz von Koriat (1997) bietet einen Erklärungsansatz für den Effekt: Hier wird ebenfalls angenommen, dass verzögerte JOLs auf anderen Grundlagen beruhen als unmittelbar abgegebene. Er unterscheidet sich von den anderen vorgestellten Ansätzen in erster Linie dadurch, dass er die diagnostische Validität dieser Grundlagen (Hinweisreize) mit berücksichtigt. Da der Delayed-JOL-Effekt als sehr robust gilt, werden in den Experimenten der vorliegenden Arbeit JOLs standardmäßig erst nach dem vollständigen Durchlaufen der Lernphase erfasst, sodass potenzielle Verbesserungen in den JOL-Genauigkeiten durch Interventionen oder Kontextfaktoren immer als zusätzlich zu diesem Effekt interpretiert werden können.

Cue-utilization-Ansatz. Wie soeben erwähnt fasste Koriat (1997) die Vielfalt der verschiedenen Grundlagen, auf denen JOLs beruhen können, im Cue-utilization-Ansatz zusammen. Sein Ansatz wendet sich gegen die ursprüngliche Ansicht, dass Personen in der Lage sind, die Stärke einer bestimmten Gedächtnisspur zu beurteilen und dementsprechend hohe oder niedrige JOLs abzugeben (*Direct-access-Hypothese*; King, Zechmeister & Shaughnessy, 1980). Im Cue-utilization-Ansatz wird postuliert, dass JOLs inferentieller Natur sind und auf impliziten Heuristiken beruhen. Wie im letzten Abschnitt erwähnt, ist die Genauigkeit des metakognitiven Urteils davon abhängig, wie stark die genutzten Hinweisreize mit den Faktoren, die die spätere Leistung bestimmen, zusammenhängen. Koriat (1997) unterscheidet hierbei drei Kategorien von Hinweisreizen. Zunächst sind intrinsische Hinweisreize zu nennen, also solche, die mit Eigenschaften der zu lernenden Items zusammenhängen wie z.B. die Itemschwierigkeit, der Grad der Assoziativität bei Itempaaren, die Bildhaftigkeit (imagery), Gegenständlichkeit oder Auftretenshäufigkeit. Die zweite Kategorie, die extrinsischen Hinweisreize, beinhaltet Informationen, die sich aus der Lernsituation ergeben, wie z.B. die Zahl der Lerndurchgänge, die Präsentationsdauer, aber auch

Besonderheiten des Lernvorgangs wie z.B. die Verwendung von Elaborationsstrategien. Als dritte Kategorie schlägt Koriat (1997) die mnemonischen Hinweisreize vor. Diese geben das Ausmaß an, zu dem ein Item gelernt wurde und zukünftig abgerufen werden kann. Mnemonische Hinweisreize sind also erfahrungs- und nicht theoriebasiert und umfassen Mechanismen wie die Leichtigkeit von Verarbeitung und Abruf, aber auch die Vertrautheit mit dem Stimulus. Sie sind laut Koriat (1997) nicht als vollständig unabhängig von intrinsischen und extrinsischen Hinweisen zu sehen, sondern werden von diesen beeinflusst.

Der Beitrag der einzelnen Hinweisreize auf die JOL-Abgabe ist verschieden gut erforscht. Bezüglich der intrinsischen Reize liegen bislang die meisten empirischen Befunde vor. Die erste wichtige Arbeit ist die oben zitierte Studie von Arbuckle und Cuddy (1969) zur Bedeutung der semantischen Assoziativität von Items. Auch wenn die Berücksichtigung intrinsischer Hinweisreize häufig zu einer erhöhten Genauigkeit von JOLs führt, wurden einige Bedingungen identifiziert, unter denen diese die Vorhersageleistung verschlechtern. So wurden z.B. für die Erinnerungsleistung subtilere Unterschiede zwischen Itempaaren unterschiedlicher Assoziationsgrade gezeigt als für die entsprechenden JOLs (Koriat & Bjork, 2005), identische Items wurden mit höheren JOLs bewertet als hoch assoziative, was sich jedoch gegenläufig in der Erinnerungsleistung zeigte (Castel, McCabe & Roediger, 2007). Weiterhin wurden seltene Wörter als schwerer zu erinnern eingeschätzt als häufige Wörter, obwohl bezüglich der tatsächlichen Erinnerung das Gegenteil der Fall war (Begg et al., 1989). Annahmen über die Leichtigkeit eines Lernvorgangs und sich daraus ergebende Schlussfolgerungen über die spätere Erinnerungsleistung können also auch irreführend sein; intrinsische Hinweisreize können demnach nicht in jedem Fall als „diagnostisch“ (vgl. Koriat, 1997) gewertet werden.

Zum Einfluss extrinsischer Hinweisreize auf die Höhe und Güte der JOLs liegen etwas weniger Befunde vor. Dunlosky und Nelson (1994) kontrastierten in diesem Zusammenhang die Wirkung von verteiltem vs. gehäuftem Lernen als Faktor der Lernsituation sowie den Einfluss einer Instruktion in „Interactive imagery“ im Gegensatz zu reinem Auswendiglernen als Besonderheit des Lernvorgangs. Sowohl das verteilte Lernen als auch die Verwendung der Elaborationsstrategie hatte erwartungsgemäß höhere JOLs zur Folge; es zeigte sich jedoch, dass der Effekt der jeweiligen Lernbedingung für verzögerte JOLs deutlich größer ausfiel als bei einer JOL-Abgabe unmittelbar nach dem Lernvorgang. Als weiterer extrinsischer Hinweisreiz kann die Lernzeit gelten. Neuere Studien zum Einfluss der selbstgewählten Lernzeit (z.B. Koriat, Ackerman, Adiv, Lockl & Schneider, 2014) zeigen, dass Personen etwa ab dem Jugendalter Informationen, die implizit aus der Lernzeit abgeleitet werden, als Basis für spätere JOLs heranzogen, und zwar in dem Sinn, dass kurz gelernte Items in der Regel höhere JOLs

erhielten als länger gelernte (*Easily-learned-easily-remembered-Heuristik*; vgl. auch Koriat, Ackerman, Lockl & Schneider, 2009a; s. Abschnitt 3.3.2). Auch experimentelle Manipulationen der Lernzeit wurden als extrinsischer Hinweisreiz genutzt. In diesem Fall fielen jedoch sowohl JOLs als auch die spätere Erinnerungsleistung bei länger gelernten Items größer aus, da hier nicht die Teilnehmer selbst bestimmen konnten, welche Stimuli für sie schwer bzw. leicht zu lernen sind (Koriat et al., 2006).

Ein direkter Vergleich der Nutzung intrinsischer und extrinsischer Hinweisreize zeigt, dass erstere bei JOLs, die unmittelbar nach dem Lernvorgang abgegeben werden, deutlich mehr Einfluss haben als extrinsische (z.B. Dunlosky & Nelson, 1994; Koriat, 1997; Koriat & Ma'ayan, 2005). In seinem Erklärungsansatz geht Koriat (1997) davon aus, dass ein einzelnes JOL in erster Linie auf der Grundlage eines Vergleichsprozesses innerhalb der Items abgegeben wird und sich je nach Zeitpunkt das Gewicht einzelner Hinweisreize zugunsten anderer verschiebt. So sind z.B. auch bei verzögerten JOLs neben den intrinsischen Hinweisen vermehrt Einflüsse aus dem Lernvorgang oder Informationen aus einem Probeabruf präsent, was als ein wichtiger Grund dafür gesehen werden kann, dass verzögert abgefragte JOLs in der Regel deutlich genauer sind als unmittelbare.

Befunde zu mnemonischen Hinweisreizen sind eng verknüpft mit solchen zu intrinsischen und extrinsischen Informationen. So zeigten Begg et al. (1989) beispielsweise, dass der Einfluss der Gegenständlichkeit oder der Häufigkeit verschiedener Stimuli auf die Höhe der JOLs von der daraus resultierenden Leichtigkeit der Verarbeitung vermittelt wurde (vgl. auch Koriat, 1997). Undorf und Erdfelder (2011) belegten, dass die Leichtigkeit der Verarbeitung auch dann als (mnemonischer) Hinweisreiz für JOLs herangezogen wird, wenn die Itemschwierigkeit kontrolliert wird. Auch andere mnemonische Hinweisreize, wie z.B. die Vertrautheit mit den Items, werden von extrinsischen Faktoren, wie z.B. der Häufigkeit, mit der ein Stimulus präsentiert wird, beeinflusst (z.B. Kelley & Lindsay, 1993).

Der Vorteil des Cue-utilization-Ansatzes ist darin zu sehen, dass dieses Rahmenmodell eine Fülle von möglichen Einflussfaktoren auf JOLs bündelt und der experimentell beobachteten flexiblen Nutzung verschiedener Grundlagen für metakognitive Urteile Rechnung trägt (z.B. Koriat & Ma'ayan, 2005). Zudem sind hier sowohl erfahrungsbasierte, d.h. Informationen, die aus dem aktuellen Lernvorgang abgeleitet werden, als auch theoriebasierte, also potenziell bewusste Annahmen über eigenes Wissen oder Fähigkeiten, zusammengefasst. Im Gegensatz zur Direct-access-Hypothese sind hier demnach auch unterschwellig ablaufende Prozesse als Hinweisreize eingeschlossen. Obwohl der Ansatz von Koriat (1997) prinzipiell für JOLs

konzipiert wurde, liegt die Übertragung auf andere metakognitive Urteile nahe. So ist davon auszugehen, dass zu anderen Zeitpunkten im Lernprozess, so also bei der Abgabe von EOLs oder SUs, die Gewichtung der einzelnen Hinweisreize im Vergleich zu JOLs unterschiedlich ausfällt, in dem Sinne, dass z.B. bei EOLs vermehrt intrinsische, bei SUs eher extrinsische und mnemonische Hinweisreize herangezogen werden. Als kritisch ist zu sehen, dass der Cue-utilization-Ansatz keine Aussagen über die exakte Nutzung bzw. die Gewichtung verschiedener Hinweisreize trifft. Es bleibt also unklar, welcher Anteil eines konkreten JOL in einer konkreten Lernsituation auf welchen Hinweisreiz zurückzuführen ist, sodass auch die Trennung zwischen den drei Arten von Hinweisreizen als unscharf gelten muss. Weiterhin wird in dem Ansatz zwar eine Vielzahl von Einflussfaktoren berücksichtigt, dennoch kann die Zusammenstellung nicht als erschöpfend angesehen werden. So wird z.B. das Ausmaß, zu dem ein (Probe-)Abruf und das sich daraus ergebende JOL auf einer detaillierten und bewussten Erinnerung an Details aus dem Lernvorgang beruht (*Recollection*; z.B. McCabe & Soderstrom, 2011; Metcalfe & Finn, 2008), nicht berücksichtigt, sondern nur das Ausmaß der Vertrautheit (*Familiarity*; Koriat, 1997). Diese Unterscheidung ist jedoch im entwicklungspsychologischen Kontext besonders bei Überwachungsleistungen bei älteren Erwachsenen relevant, da es hier Belege gibt, dass sich in dieser Altersgruppe vermehrt Schwierigkeiten finden, „Recollection“-Prozesse zu überwachen (z.B. Daniels, Toth & Hertzog, 2009; vgl. Abschnitt 3.1.3.2). Trotz dieser Kritikpunkte ist der Ansatz von Koriat (1997) als wegweisend zu sehen, da er in der Lage ist, vielfältige empirische Beobachtungen zu den Grundlagen von metakognitiven Urteilen zu integrieren, prinzipiell erweiterbar ist und – vergleichbar zu Ergebnissen aus anderen Forschungsrichtungen, die gezeigt haben, dass menschliches Urteilen durch das Heranziehen von Heuristiken fehleranfällig sein kann (z.B. Tversky & Kahneman, 1974) – berücksichtigt, dass auch metakognitive Urteile in der Mehrheit der Fälle nicht vollständig planvoll und überlegt, sondern nach Verfügbarkeit von Hinweisreizen getroffen werden. Auch neuere Erklärungsansätze für das Zustandekommen von metakognitiven Urteilen – so z.B. das *Isomechanism Framework* von Dunlosky und Tauber (2014) – gehen davon aus, dass jede metakognitive Aussage aus den verfügbaren Hinweisreizen abgeleitet wird und diese je nach Kontext oder Zeitpunkt des Urteils variieren können.

JOLs beim Lernen von komplexeren Materialien. Auch beim Lernen aus Texten und aus anderen komplexeren Materialien wie Filmen können während oder nach dem Lernvorgang Leistungsvorhersagen getroffen werden. Diese Urteile werden zumeist unter dem Begriff *Metacomprehension judgments* gefasst, da sie in der Regel nicht nur eine Einschätzung der

zukünftigen Lern- und Erinnerungsleistung, sondern auch des Grades des Verständnisses erfassen (vgl. Thiede, Griffin, Wiley & Redford, 2009). Zur besseren Vergleichbarkeit mit den Befunden beim Paar-Assoziations- bzw. Listenlernen wird jedoch im Folgenden auch beim Lernen aus Texten für prospektive Lernurteile der Begriff JOL verwendet. Im Vergleich zu Überwachungsurteilen bei weniger komplexen Materialien fallen JOLs bei Texten typischerweise deutlich niedriger aus. In einer frühen Studie von Glenberg und Epstein (1985) wurden die Teilnehmer aufgefordert, für kurze Texte jeweils vorherzusagen, wie sicher sie sind, auf der Grundlage der Inhalte eine Frage zum zentralen Thema des Textes zu beantworten. Die Korrelationen zwischen dieser Vorhersage und der Korrektheit der nachfolgenden Antwort waren nicht von Null verschieden; die Probanden konnten also ihre Testleistung nicht vorhersagen. In späteren Überblicksarbeiten (Dunlosky & Lipko, 2007; Maki, 1998a) lag die über mehrere Studien gemittelte Vorhersagegenauigkeit zwar etwas höher, mit Korrelationskoeffizienten um .25 jedoch immer noch im niedrigen Bereich. Aus diesem Grund wurden auch in der Metacomprehension-Forschung Anstrengungen unternommen, die Überwachungsgenauigkeit beim Lernen aus Texten zu steigern. Als ein wesentlicher Grund für die niedrigen Genauigkeiten wird häufig das Abfrageformat genannt. In vielen Studien korrespondiert die JOL-Frage nur in geringem Maß mit der Erinnerungsfrage. Zum Teil wird lediglich ein globales Lern- bzw. Verständnisurteil erhoben, welches dann mit der Leistung in einzelnen, detaillierten Aufgaben zum Text in Relation gesetzt wird (z.B. Glenberg & Epstein, 1985). Eine Studienreihe der Arbeitsgruppe um Dunlosky zeigte, dass sich die relative Genauigkeit der Überwachungsurteile durch eine spezifischere Formulierung der Vorhersagen vor allem dann verbesserte, wenn die Probanden explizit aufgefordert wurden, einen Probeabruf zu machen, und wenn sie zusätzlich eine Rückmeldung über die Korrektheit dieses Abrufes erhielten (Dunlosky, Rawson & McDonald, 2002; Dunlosky, Rawson & Middleton, 2005; Rawson & Dunlosky, 2007). Weitere Steigerungen der Vorhersage-Genauigkeiten konnten durch verzögert abgegebene Zusammenfassungen der Texte bzw. durch die verzögerte Generierung von Schlüsselwörtern erreicht werden (Anderson & Thiede, 2008; Thiede, Dunlosky, Griffin & Wiley, 2005). Das heißt also, je vergleichbarer die Situation bei der Abgabe der JOLs mit der Erinnerungsabfrage ist und je tiefer das Lernmaterial verankert wurde, desto höher werden auch die Genauigkeiten. Schließlich konnte auch über das zweifache Lesen eines Textes eine Steigerung der JOL-Genauigkeit erreicht werden. Besonders bei schwierigen Texten erleichtert das erneute Lesen zum einen die Bildung eines Situationsmodells, zum anderen werden kognitive Ressourcen frei, die dann der Überwachung dienen können (Dunlosky, 2005; Rawson, Dunlosky & Thiede, 2000).

Vergleichbar mit den theoretischen Annahmen über die Grundlagen von JOLs beim Paar-Assoziationslernen liegen inzwischen auch für Metacomprehension-Urteile Studien über zugrunde liegende Hinweisreize und deren diagnostische Validität vor. Nach der Theorie von Kintsch (1998) sollten Personen, die sich bei der Überwachung des Textverständnisses nach Hinweisreizen richten, die auf das Situationsmodell Bezug nehmen, höhere Genauigkeiten erzielen als Personen, die eher auf Hinweisreize aus der Textoberfläche achten. Erste Ergebnisse zeigen, dass z.B. Probanden mit Leseschwierigkeiten eher auf der wörtlichen Ebene verarbeiten und überwachen; Interventionen wie das „Concept-Mapping“, die Hinweisreize aus dem Situationsmodell salient machen, konnten in dieser Gruppe auch die JOL-Genauigkeit steigern (Thiede, Griffin, Wiley & Anderson, 2010; vgl. auch T. D. Griffin, Wiley & Thiede, 2008). Ergänzend führten Zhao und Linderholm (2008) an, dass mögliche Hinweisreize für das Zustandekommen von Metacomprehension-Urteilen auch vorgefertigte Annahmen über die eigene Leistungsfähigkeit in einem bestimmten Themenbereich sein können, die dann – vergleichbar mit den intrinsischen Hinweisreizen sensu Koriat (1997) – als Anker für die Urteile dienen. Diese Erwartungen können besonders bei Personen mit einem hohen Vorwissen im abgefragten Leistungsbereich wirksam werden und unter Umständen die Vorhersagegenauigkeit verschlechtern.

Insgesamt lässt sich festhalten, dass sich JOL-Genauigkeiten beim Lernen aus Texten schwerer verbessern lassen als solche zu Paarassoziationen. Zusammenfassungen, das Generieren von Schlüsselwörtern, das zweimalige Lesen des Textes, besonders aber eine hohe Korrespondenz zwischen der Art der Urteilsabfrage und der Abfrage der Lernleistung stellen Ansätze dar, die Überwachungsleistungen zu steigern. Inwiefern diese Möglichkeiten die Grundlagen von Metacomprehension-Urteilen beeinflussen, konnte bislang nicht geklärt werden, da diese insgesamt noch wenig erschlossen sind.

3.1.3.2 Entwicklungspsychologische Befunde

Entwicklung im Vor- und Grundschulalter. Während allgemeinspsychologische Arbeiten zu JOLs bei Erwachsenen in großer Zahl vorliegen, sind entwicklungspsychologische Studien deutlich seltener.

Erst seit kurzem gibt es Befunde zu JOLs im Vorschulalter. Ähnlich wie EOLs fallen auch JOLs bei jüngeren Kindern deutlich überoptimistisch aus (z.B. Lipowski, Merriman & Dunlosky, 2013; Lipko, Dunlosky & Merriman, 2009). So gaben beispielsweise in einer Stichprobe von

Vorschülern bei unmittelbar abgegebenen JOLs fast 100% der Kinder an, sich später alle Items korrekt merken zu können. Eine verzögerte JOL-Abgabe reduzierte diese Rate auf etwa 50%, doch erst die Erhebung der JOLs in einem *Forced choice-Format* („Welches dieser beiden Bilder wirst du dir besser merken?“) trug dazu, die Überwachungsgenauigkeit zu erhöhen (Lipowski et al., 2013). Auch bei JOLs kann nicht davon ausgegangen werden, dass ein allgemeines Überwachungsdefizit vorliegt, da auch Vorschulkinder in der Lage sind, ihre letzte Erinnerungsleistung korrekt wiederzugeben (Lipowski et al., 2013), d.h. auch hier scheint der motivationale Einfluss des Wunschenkens wirksam zu sein (Schneider, 2015; s. aber Lipko et al., 2009).

Ab dem Grundschulalter liegt die JOL-Genauigkeit auf einem deutlich höheren Niveau als im Kindergartenalter; Alterstrends im Sinne einer Steigerung der Vorhersage zeigen sich nicht (z.B. Schneider, Visé, Lockl & Nelson, 2000) oder nur in geringem Ausmaß (z.B. Koriat & Shitzer-Reichert, 2002). Dabei scheinen JOLs im Grundschulalter ähnlichen Mechanismen zu unterliegen wie bei Erwachsenen. So fanden Schneider et al. (2000) bei sechs-, acht- und zehnjährigen Kindern in einem Paar-Assoziationsdesign ebenfalls deutliche Belege für den Delayed JOL-Effekt: Die mittleren Gammakorrelationen stiegen über die Altersgruppen hinweg von .18 für unmittelbar abgegebene JOLs auf .73 für verzögerte JOLs; es zeigten sich keine signifikanten Alterseffekte. Als zugrunde liegende Wirkmechanismen sind die gleichen Faktoren anzunehmen wie bei Erwachsenen (s.o.). Auch für den Cue-utilization-Ansatz wiesen Koriat und Shitzer-Reichert (2002) bei Kindern der zweiten und vierten Jahrgangsstufe ähnliche Befunde wie bei Erwachsenen nach: So nahm im Lauf mehrerer Lerndurchgänge der Einfluss der Übung auf die JOL-Genauigkeit als extrinsischer Hinweisreiz zu. Dieser Effekt war bei den Zweitklässlern etwas stärker ausgeprägt, da bei den Viertklässlern Deckeneffekte in der JOL-Genauigkeit vorlagen. Der Einfluss der intrinsischen Hinweisreize auf die Höhe der JOLs – hier operationalisiert durch die Itemschwierigkeit – nahm jedoch im Unterschied zu den Befunden bei Erwachsenen über die Lerndurchgänge hinweg vor allem bei den jüngeren Probanden zu, d.h. die Salienz der Itemschwierigkeit schien bei dieser Gruppe erhalten zu bleiben.

Auch bei komplexeren Lernmaterialien wie dem Erinnern von Informationen aus Filmen zeigten sich für Kinder JOL-Genauigkeiten in vergleichbarer Höhe wie bei Erwachsenen (z.B. Roebbers, von der Linden, Schneider & Howie, 2007). In Bezug auf mögliche Verbesserungen von Vorhersagen bei komplexeren Materialien zeigte sich bei Kindern im Alter von sieben bzw. neun Jahren kein positiver Einfluss einer vorgeschalteten eigenen Zusammenfassung der Filminhalte (von der Linden et al., 2011). Auch beim Lernen von Texten hatte das verzögerte

Generieren von Schlüsselwörtern nur bei Teilnehmern der siebten und achten Jahrgangsstufe, nicht aber bei Viertklässlern eine Steigerung der JOLs zur Folge (de Bruin, Thiede, Camp & Redford, 2011). Bei Grundschulern scheinen also Überwachungsurteile beim Lernen von komplexeren Materialien deutlich schwieriger zu beeinflussen zu sein als bei Erwachsenen.

Entwicklung im Jugendalter. Ähnlich wie EOLs umfassen Studien zu JOLs sehr selten jugendliche Teilnehmer. Eine Arbeit von Koriat, Ackerman, Lockl und Schneider (2009b), die JOLs beim Paar-Assoziationslernen bei Teilnehmern im Alter zwischen 7 und 12 Jahren erhob, zeigte einen leichten alterskorrelierten Anstieg der Überwachungsgenauigkeiten. Auch für den Vergleich von Probanden der ersten bis achten Klasse ergaben sich bei Hoffmann-Biencourt, Lockl, Schneider, Ackerman und Koriat (2010) höhere Gammakorrelationen für die älteren Schüler als für die jüngeren. Eine andere Studie, die unmittelbare JOLs bei Schülern zwischen 6 und 15 Jahren erfasste, berichtete keine Maße der Resolution oder Kalibrierung, sodass deren Ergebnisse hier nicht eingeordnet werden können (Paulus, Tsalas, Proust & Sodian, 2014). Selbst wenn also die Entwicklung der Überwachungsfähigkeit schon im Grundschulalter ein recht hohes Niveau erreicht, gibt es Hinweise darauf, dass leichte Verbesserungen bis ins Jugendalter nicht auszuschließen sind. Es muss jedoch berücksichtigt werden, dass die wenigen vorliegenden Studien keine Anpassung der Lernaufgaben zwischen den Altersgruppen vorgenommen hatten, sodass aus den Anstiegen in der JOL-Genauigkeit nicht grundsätzlich auf eine Verbesserung der Überwachungsleistung geschlossen werden kann.

Entwicklung im höheren Erwachsenenalter. Für ältere Erwachsene liegen deutlich mehr Studien mit JOLs vor, die mehrheitlich dafür prechen, dass Leistungsvorhersagen, die nach dem Lernvorgang getroffen werden, in ihrer Genauigkeit recht konstant bleiben.

Schon Lovelace und Marsh (1985) verglichen JOLs von jüngeren und älteren Probanden beim Lernen von Wortpaaren und fanden bezüglich der relativen Genauigkeit (Resolution) keine Unterschiede zwischen den beiden Altersgruppen, d.h. die Teilnehmer waren generell in der Lage, zwischen später erinnerten und später vergessenen Paaren zu differenzieren. Lediglich in der absoluten Genauigkeit (Kalibrierung) zeigten sich Alterseffekte, in dem Sinne, dass die älteren Probanden die Schwierigkeit der Aufgabe eher unterschätzten, was die Autoren mit der geringen Erfahrung der älteren Teilnehmer mit dieser Art von Aufgaben erklärten.

Etwas spätere Studien, die Gammakorrelationen als Maß für die Überwachungsleistung einsetzten, zeigten für Paar-Assoziationsparadigmen ebenfalls vergleichbare Ergebnisse bei jüngeren und älteren Erwachsenen. Für unmittelbar abgegebene JOLs wurden bei Hertzog, Kidder, Powell-Moman und Dunlosky (2002) beispielsweise Korrelationen um .50 sowohl für jüngere als auch für ältere Erwachsene gefunden. Zudem wurden auch für den Delayed JOL-Effekt bei älteren Probanden Nachweise erbracht. So steigerten sich z.B. bei Connor et al. (1997) die Gammakorrelationen von etwa .50 für die unmittelbaren JOLs auf etwa .80 bis .90 für verzögerte JOLs.

In Bezug auf die zugrunde liegenden Wirkmechanismen existieren aktuell einige Befunde, die im Wesentlichen den Cue-utilization-Ansatz (Koriat, 1997) auch bei älteren Erwachsenen belegen. Bereits 1982 verfolgten Rabinowitz, Ackerman, Craik und Hinchley einen Ansatz, der in Koriats theoretisches Modell eingeordnet werden kann: Vergleichbar mit dem Vorgehen, das für Experiment 4 (s. Abschnitt 6.5) der vorliegenden Arbeit gewählt wurde, erhielt je die Hälfte der Probanden der beiden Altersgruppen die Instruktion, interaktive Bilder aus den dargebotenen Wortpaaren zu bilden. Unabhängig von der Altersgruppe profitierten die Teilnehmer von der Anwendung dieser Elaborationsstrategie. Es zeigten sich jedoch keine Transfereffekte auf die JOL-Genauigkeit; vielmehr lagen bei den Probanden in der Instruktionsbedingung die Erinnerungsraten auch bei solchen Items hoch, deren JOLs niedrig ausfielen. Rabinowitz et al. (1982) schlossen daraus, dass Personen nicht in der Lage sind, während des Überwachungsvorgangs den differenziellen Einfluss verschiedener Enkodierstrategien im Vergleich zum Einfluss anderer Faktoren, wie z.B. der Itemschwierigkeit, zu berücksichtigen. Eine neuere Arbeit von Hertzog et al. (2010), die 285 Personen der gesamten Altersspanne zwischen 18 und 80 Jahren untersuchte, zeigte jedoch, dass sich in allen Altersstufen auch ohne die explizite Instruktion, eine Interactive imagery-Strategie zu verwenden, der spontane Gebrauch dieser Strategie sowohl auf die Erinnerungs- als auch auf die Überwachungsleistung positiv auswirkte. Auch hier war der Einfluss der Itemschwierigkeit auf die JOLs größer, jedoch fiel der Einfluss der Elaborationsstrategie auf die Erinnerungsleistung größer aus. Sowohl bei Rabinowitz et al. (1982) als auch bei Hertzog et al. (2010) wurden unmittelbare JOLs erfasst; deshalb liegt die Vermutung nahe, dass die intrinsischen Hinweisreize hier in beiden Altersgruppen deutlich präsenter waren als der Einfluss der Qualität der Enkodierung, sodass eine Replikation mit verzögerten JOLs sinnvoll erscheint. Es stellt sich in diesem Zusammenhang weiterhin die Frage, inwiefern eine etwas ausführlichere Instruktion in der Interactive imagery-Strategie, die den Teilnehmern auch die Gelegenheit gibt, diese einzuüben, die Effekte auf die JOL-Genauigkeit steigern kann. Einen

weitere Einfluss auf die JOL-Genauigkeit, der speziell bei älteren Erwachsenen von Bedeutung ist, hat die Tatsache, ob Lern- und Erinnerungsprozesse auf einem bewussten Abrufprozess (Recollection) oder auf der reinen Vertrautheit (Familiarity) mit dem Stimulus beruhen (s.o.). Generell fallen JOLs für Vorgänge, die auf der Basis von „Recollection“-Prozessen stattfinden, genauer aus als solche, die sich auf Vertrautheitsprozesse stützen. Es zeigte sich jedoch auch, dass ältere im Vergleich zu jüngeren Erwachsenen mehr Schwierigkeiten haben, „Recollection“-Prozesse akkurat zu überwachen und sich verstärkt auf die Vertrautheit mit dem Lernmaterial berufen (Daniels et al., 2009). Es scheint also sinnvoll zu sein, „Recollection“-Prozesse bei älteren Erwachsenen zu fördern. Auch dies kann durch die Instruktion einer Elaborationsstrategie geschehen, da durch die in der Lernsituation stattfindende Anreicherung des Lernmaterials mit Details Familiarity-Prozesse in den Hintergrund gerückt werden sollten.

Die Forschung zu JOLs beim Lernen von Texten steht für die Gruppe der älteren Erwachsenen noch in den Anfängen. Eine erste Studie von Olin und Zelinski (1997) erfasste beim Lesen von Texten sowohl eine Einschätzung der Leichtigkeit der Verarbeitung als auch des Textverstehens sowie eine Leistungsvorhersage. Diese Einschätzungen wurden jedoch nicht mit der Leistung beim Wiedererkennen einzelner Sätze verglichen, sodass streng genommen keine Aussage über die Qualität der Überwachungsleistung möglich ist. Denn die Ratings decken sich hinsichtlich der erfassten Ebene (Situationsmodell vs. Textoberfläche) nicht mit der späteren Leistungsabfrage. Olin und Zelinski (1997) fanden, dass die Korrelation zwischen den Einschätzungen und der Wiedererkennensleistung sich zwischen jüngeren und älteren Teilnehmern insgesamt nicht unterschied; bei den jüngeren Erwachsenen wurde die Leistung jedoch besser durch die Einschätzung des Textverstehens vorhergesagt, bei den älteren durch die Einschätzung der Verarbeitungsleichtigkeit. Dies wird von den Autoren dahingehend interpretiert, dass die Beurteilung der Verarbeitungsleichtigkeit eine besser zugängliche Heuristik darstellt, auf die besonders ältere Erwachsene zurückgreifen, um ihre kognitiven Ressourcen ökonomisch einzusetzen. Neuere Studien zeigen ebenfalls keine oder nur geringe Unterschiede in der JOL-Genauigkeit beim Lernen von Texten zwischen jüngeren und älteren Erwachsenen (Miles & Stine-Morrow, 2004; Shake, Noh & Stine-Morrow, 2009; Stine-Morrow, Shake, Miles & Noh, 2006). Unterschiede in den Gammakorrelationen zugunsten jüngerer Erwachsener ergaben sich vor allem dann, wenn die JOL-Fragen im Vergleich zum späteren Erinnerungsabruf sehr unspezifisch erfragt wurden (Miles & Stine-Morrow, 2004) oder wenn während der Bearbeitung mit freier Lernzeiteinteilung die Schwierigkeit des Textmaterials stark variierte, sodass eine konstante Anpassung von Lernzeit und Überwachungsstrategie erforderlich war (Shake et al., 2009). Insgesamt gesehen lagen die

Gammakorrelationen bei itembasierten Vorhersagen im niedrigen bis mittleren Bereich, was auch der in anderen Studien mit jüngeren Erwachsenen gefundenen Ausprägung entspricht. Eine Studie, die Metacomprehension bei jüngeren und älteren Teilnehmern mit einem einzigen Globalurteil erfasste, fand für die älteren Erwachsenen eine Gammakorrelation mit der späteren Erinnerungsleistung um Null, für jüngere Erwachsene um .4 (Baker, Dunlosky & Hertzog, 2010). Dies lässt sich damit erklären, dass ältere Erwachsene für ein Gesamturteil eher nicht diagnostische Hinweisreize wie z.B. die Vertrautheit mit dem Thema des Textes heranziehen, anstatt sich auf ihr Textverständnis zu stützen. Zusammenfassend lässt sich festhalten, dass die Metacomprehension-Leistung bei älteren Erwachsenen dann vergleichbar mit der Leistung jüngerer Erwachsener ausfällt, wenn die Urteile auf Itemebene und in der Formulierung analog zur Erinnerungsabfrage erfasst werden und wenn die Aufgabenschwierigkeit konstant gehalten wird.

3.1.3.3 Fazit

Insgesamt gesehen konnten die ursprünglich gefundenen mittelhohen Genauigkeiten in den JOLs vor allem durch die zeitlich verzögerte Erfassung der Urteile immens gesteigert werden (Nelson & Dunlosky, 1991). Dies gilt besonders für Aufgaben, die die freie Erinnerung erfassen, jedoch in etwas geringerem Maße auch für Aufgaben zum Wiedererkennen (z.B. Dunlosky & Nelson, 1997). Versuche, die insgesamt niedrigeren JOL-Genauigkeiten beim Lernen aus Texten zu verbessern, waren bislang weniger erfolgreich. Als vielversprechend hat sich herausgestellt, die Formulierung der Leistungsvorhersage und die Leistungsabfrage möglichst identisch zu gestalten (Dunlosky, Rawson & Hacker, 2002) – ein Ansatz, der auch in den eigenen Studien mit komplexeren Materialien umgesetzt werden soll. Weiterhin scheint besonders bei anspruchsvollen Texten das zweimalige Lesen zu einer Verbesserung der Überwachungsleistung zu führen (Rawson et al., 2000).

Mit dem Cue-utilization-Ansatz steht ein wegweisendes Modell zur Verfügung, um die theoretischen Grundlagen von JOLs zu klassifizieren (Koriat, 1997). Es fehlen bisher jedoch genauere Aussagen darüber, welche Art von Hinweisreizen in welchem Ausmaß auf die Höhe des Urteils Einfluss nimmt. Weiterhin sollte das Modell, gerade um Grundlagen von Überwachungsprozessen im höheren Erwachsenenalter spezifischer zu untersuchen, um die Unterscheidung zwischen „Recollection“- und „Familiarity“-basierten Urteilen erweitert werden (vgl. Daniels et al., 2009).

Im Entwicklungsverlauf zeigt sich, dass allgemeinspsychologische Effekte zu JOLs im Wesentlichen auch für andere Altersgruppen gelten. Die starke Überschätzung der eigenen Lernleistung, die noch bei Vorschulkindern vorliegt, reduziert sich zur Mitte der Grundschulzeit hin deutlich. Bis zum jüngeren Erwachsenenalter scheint es insgesamt kaum oder nur geringe Steigerungen zu geben; da Studien mit Jugendlichen allerdings selten sind, können zu dieser Altersgruppe noch keine weitreichenden Schlussfolgerungen gezogen werden (für einen Überblick s. Schneider, 2015). Im höheren Erwachsenenalter scheint die Überwachungsleistung in den JOLs zumindest unter günstigen Bedingungen konstant zu bleiben (für einen Überblick s. Dunlosky & Metcalfe, 2009). Hier fehlen jedoch Befunde zu komplexeren Materialien, um die Allgemeingültigkeit dieser Beobachtung bestätigen zu können.

3.1.4 Feeling-of-Knowing-Urteile

Auch wenn Feeling-of-Knowing-Urteile (FOKs) als einziges der Maße im Modell von Nelson und Narens (1990; 1994) nicht in der vorliegenden Arbeit erhoben werden, soll trotzdem der Vollständigkeit halber kurz auf deren Ausprägung und Grundlagen im Zusammenhang mit allgemeinspsychologischen Befunden sowie deren Entwicklung über die Lebensspanne eingegangen werden.

3.1.4.1 Allgemeinspsychologische Befunde

FOKs werden etwas abweichend von den anderen metakognitiven Überwachungsurteilen EOLs, JOLs und SUs erfasst, da hierfür zunächst der Zustand zwischen Erinnern und Vergessen, dessen intensivste Ausprägung dem Phänomen „Es liegt mir auf der Zunge“ entspricht, induziert werden muss. Noch bevor Flavell (1971) den Begriff Metagedächtnis prägte, entwickelte Hart (1965) ein Untersuchungsparadigma, das im Wesentlichen auch noch heute angewandt wird, um diesen Überwachungsvorgang messbar zu machen. Er ließ Probanden zunächst Fragen aus dem Bereich Allgemeinwissen im offenen Format beantworten. Danach sollten sie für die Aufgaben, die nicht gelöst werden konnten, mit einem Ja-Nein-Urteil einschätzen, ob sie diese aus einer Reihe von möglichen Antworten erkennen könnten. Schließlich folgte für die nicht gelösten Aufgaben eine Wiedererkennungsaufgabe, bei der die richtige Antwort aus vier Vorschlägen ausgewählt werden musste (*Recall-judgment-recognition-Paradigma*). Hart (1965) fand bei jüngeren Erwachsenen recht genaue FOK-

Urteile, die Teilnehmer waren also – wenn die Aufgaben insgesamt ausreichend schwierig waren – in der Lage, vorherzusagen, welche aktuell nicht erinnerten Items sie wahrscheinlich wiedererkennen werden. Im Gegensatz zu den Berechnungen von Hart (1965), der „Treffer“ und „falsche Alarme“ in den Vorhersagen mit Methoden der Signal-Entdeckungstheorie analysierte, wurden später vor allem die relative Genauigkeit mithilfe von Gammakorrelationen sowie zum Teil auch Kalibrierungskurven bestimmt (z.B. Koriat, 1993). Außerdem wurde häufig nicht nur semantisches, sondern auch episodisches Wissen, hier also in der Untersuchungssituation neu erworbenes Wissen, als Grundlage für FOKs herangezogen (z.B. Schwartz, Pillot & Bacon, 2014).

Die berechneten Gammakorrelationen bewegten sich zumeist im mittleren Bereich, also z.B. bei .55 bei Koriat (1993) für das Lernen von sinnlosen Silben, bei .34 für das Lernen von Fakten (Schwartz et al., 2014) oder zwischen .29 und .33 für Fragen zum Allgemeinwissen (Kelemen et al., 2000). Personen sind also generell in der Lage, die potenzielle Erinnerbarkeit von aktuell nicht verfügbarer Information einzuschätzen, es besteht jedoch Raum zur Verbesserung dieser Urteile.

Die Frage nach den Grundlagen von FOKs führt ebenfalls zu einer Arbeit von Koriat (1993). Er stellte den bis dato vorherrschenden *Trace-based view* auf FOKs in Frage, der davon ausgeht, dass Personen direkt auf die im Gedächtnis verfügbaren Informationen zurückgreifen, um ihre Einschätzungen abzugeben. Stattdessen postulierte Koriat (1993) vergleichbar zum Cue-utilization-Ansatz, dass auch FOKs auf inferentiellen Prozessen beruhen, also die Menge an mit dem Zielreiz assoziierter Information die Höhe der FOKs bestimmt (*Trace-accessibility view*). Es ist also vor allem die Quantität der verfügbaren Informationen und weniger die Korrektheit derselben, die auf die FOKs Einfluss hat. Dieser Ansatz konnte mehrheitlich auch experimentell untermauert werden (z.B. Schwartz, 1994; Schwartz et al., 2014; Thomas, Bulevich & Dubois, 2012).

3.1.4.2 Entwicklungspsychologische Befunde

Entwicklung im Kindes- und Jugendalter. Frühe Studien zu FOKs im Kindes- und Jugendalter gingen nach Harts Paradigma vor und fanden schon bei Vorschulkindern grundlegende Überwachungsfähigkeiten in diesem Bereich (Cultice, Somerville & Wellman, 1983) und Anstiege in der Genauigkeit bis ins Jugendalter (Wellman, 1977b; Zabrocky & Ratner, 1986). Bereits 1988 wurden diese Studien jedoch von Butterfield, Nelson und Peck

methodisch kritisiert. So kann bei FOKs, die in Form von Ja-Nein-Urteilen abgegeben werden, nicht geklärt werden, ob und inwiefern die Schwelle für diese Urteile altersabhängig ist. Weiterhin werden absolute FOK-Maße von Butterfield et al. (1988) als problematisch angesehen, da ihre Güte von der Schwierigkeit der Wiedererkennungsaufgabe abhängt. Aus diesen Gründen erfassten die Autoren in ihrer Studie FOKs zusätzlich als relative Urteile, indem die Probanden jeweils für zwei Items im Vergleich angeben sollten, welches sie später mit größerer Wahrscheinlichkeit wiedererkennen würden; aus diesen Einschätzungen und den tatsächlichen Leistungen wurden dann Gammakorrelationen berechnet. Es zeigte sich, dass die sechsjährigen Probanden den zehnjährigen in diesem Design in ihrer mittleren FOK-Genauigkeit sogar überlegen waren. Eine Replikation dieses Vorgehens in einer deutschen Stichprobe mit Kindern im Alter von 7, 8, 9 und 10 Jahren fand jedoch keinen Nachweis für diesen umgekehrten Alterstrend. Hier zeigten sich altersunabhängig niedrige, aber signifikant von Null verschiedene Gammakorrelationen (Lockl & Schneider, 2002a), sodass insgesamt davon ausgegangen werden kann, dass es im Verlauf der Grundschuljahre und darüber hinaus keinen bedeutsamen Anstieg der FOK-Genauigkeiten gibt. Lockl und Schneider (2002a) untersuchten in der genannten Studie weiterhin die Grundlagen für FOK-Urteile und fanden auch in dieser Altersgruppe Belege für Koriats (1993) Trace-accessibility view.

Entwicklung im höheren Erwachsenenalter. Auch für ältere Erwachsene existieren bereits einige frühe Studien. J. L. Lachman, Lachman und Thronesbery (1979) fanden mit dem Vorgehen nach Hart (1965) für absolute FOK-Urteile keine Unterschiede in deren Qualität zwischen jüngeren und älteren Erwachsenen. Auch bei Butterfield et al. (1988) ergab die Erfassung von relativen FOKs bei Fragen zum Allgemeinwissen vergleichbare mittlere Gammakorrelationen bei jüngeren und älteren Teilnehmern (vgl. auch Marquié & Huet, 2000). Im Kontrast dazu zeigte sich bei neueren Arbeiten zum episodischen Gedächtnis jedoch, dass hier ältere Erwachsene deutlich niedrigere Gammakorrelationen in ihren FOK-Urteilen erreichen als jüngere (z.B. Souchay, Isingrini & Espagnet, 2000; Souchay, Moulin, Clarys, Tacconnat & Isingrini, 2007). In ihrem Erklärungsansatz gehen die Autoren davon aus, dass die Tendenz älterer Erwachsener, bei der Abgabe von metakognitiven Überwachungsurteilen verstärkt die Vertrautheit mit dem Stimulusmaterial als Hinweisreiz heranzuziehen, bei neu erworbenen Informationen im Vergleich zu Fragen zum Allgemeinwissen wenig valide für die tatsächliche Erinnerungsleistung ist. Es wirken sich also altersbedingte Schwierigkeiten im episodischen Gedächtnis und die mangelnde Fähigkeit, alternative Hinweisreize heranzuziehen, auf die Qualität der FOKs aus. Besonders bei älteren Erwachsenen scheint es

demnach wichtig zu sein, hier zwischen Aufgaben zum semantischen und episodischen Gedächtnis zu trennen.

3.1.4.3 Fazit

Zusammenfassend lässt sich festhalten, dass das *Recall-judgment-recognition-Paradigma* (Hart, 1965) grundsätzlich immer noch zur Erhebung von FOKs Anwendung findet. Diese grundsätzlich andere Erfassungsmethode im Vergleich zu den anderen Überwachungsmaßen macht eine direkte Vergleichbarkeit der Urteile untereinander schwierig. Ab dem Beginn der Grundschulzeit scheint die Qualität der FOKs keinen wesentlichen Veränderungen mehr zu unterliegen (Lockl & Schneider, 2002a) und auch im höheren Erwachsenenalter zumindest für Aufgaben, die das semantische Gedächtnis betreffen, stabil zu bleiben (Butterfield et al., 1988). Die zitierten Untersuchungen zeigen jedoch, dass in jeder Altersstufe noch Verbesserungen in den zumeist mittelhoch ausfallenden FOK-Genauigkeiten möglich sind. Ein Ansatzpunkt für Steigerungen der Überwachungsqualität in den FOKs könnte in einer Verknüpfung von Trace-based und Trace-accessibility view (Koriat, 1993) liegen, z.B. indem Probanden instruiert werden, die verfügbaren Gedächtnisspuren hinsichtlich ihrer Korrektheit explizit zu prüfen.

3.1.5 Sicherheitsurteile

Sicherheitsurteile (SUs) stellen – im Gegensatz zu den anderen bisher vorgestellten Maßen – rein retrospektive metakognitive Urteile dar. Sie erfassen die Einschätzung, ob eine gegebene Antwort korrekt ist; dies erfolgt meist mit der Frage „Wie sicher sind Sie, dass diese Antwort richtig ist?“.

3.1.5.1 Allgemeinpsychologische Befunde

SUs wurden im allgemeinpsychologischen Bereich vorwiegend in alltagsnahen Kontexten erfasst, besonders im Zusammenhang mit Fragestellungen zur Verlässlichkeit des Augenzeugengedächtnisses. Ergebnisse aus Studien, in denen die relative Genauigkeit von SUs (Resolution) bestimmt wird, zeigen, dass Probanden in der Lage sind, zwischen richtigen und falschen Antworten zu unterscheiden. So lagen z.B. bei Koriat und Goldsmith (1996) die mittleren Gammakorrelationen bei Aufgaben zum Allgemeinwissen in der freien Erinnerung bei .87 und in der Wiedererkennensbedingung bei .68. Die höheren Genauigkeiten bei freier

Wiedergabe im Vergleich zum Wiedererkennen konnten auch beim impliziten Lernen von Informationen zu einem Film repliziert werden (Robinson, Johnson & Robertson, 2000). Dieser Unterschied kann einmal durch den Einfluss von korrektem bzw. falschem Raten in Rekognitionsaufgaben zustande kommen (vgl. Thiede & Dunlosky, 1994), zum anderen können beim Wiedererkennen wichtige Hinweisreize aus dem Abrufprozess wie z.B. die Lebhaftigkeit des Abrufs (*vividness*) nicht genutzt werden (Robinson et al., 2000). Bei Studien zum Lernen von Wortpaaren (Dunlosky & Hertzog, 2000) sowie zur Beantwortung von offenen und geschlossenen Fragen zu einer Film-Episode (Nolan & Markham, 1998) zeigten sich ähnlich hohe Genauigkeiten wie bei Koriat und Goldsmith (1996). Für das Lernen aus Texten ergaben sich Gammakorrelationen im mittleren bis hohen Bereich, wobei der Unterschied zwischen auditiv und schriftlich dargebotenen Texten zu vernachlässigen ist (Dunlosky, Rawson & Hacker, 2002; Pressley & Ghatala, 1988). Eine neuere Studie zum Lernen aus anspruchsvollen Texten zeigte jedoch, dass der Zeitpunkt der SU-Abgabe und die Art der Erinnerungsfrage einen deutlichen Einfluss auf die relative Genauigkeit hat (Mengelkamp & Bannert, 2010): Die Autoren fanden etwas höhere, jedoch insgesamt im mittleren Bereich liegende Gammakorrelationen bei SUs, die in Zwischentests und einem Verständnistest am Ende des Lernvorgangs erfasst wurden, als bei SUs, die bei einem Vortest sowie bei einem Transfertest abgegeben werden sollten. Auch wenn die SU-Genauigkeiten beim Lernen aus komplexeren Materialien in der Regel niedriger ausfallen als solche bei Paar-Assoziationsaufgaben, sind sie doch insgesamt deutlich höher als die entsprechenden Gammakorrelationen bei JOLs. Dies liegt zum einen daran, dass Metacomprehension-Urteile prospektiv häufig mit wenig spezifischen Gesamturteilen erfragt werden (z.B. Dunlosky, Rawson & Hacker, 2002; Glenberg & Epstein, 1985), wohingegen SUs nach konkreten Aufgaben zur Erinnerungsleistung erfasst werden. Das heißt, bei der Abgabe der SUs wird die Überwachung dadurch erleichtert, dass es den Probanden genau bewusst ist, auf welchen Inhalt sich das Urteil bezieht. Zum anderen kann davon ausgegangen werden, dass Probe-Abrufe bei JOLs zu komplexeren Materialien weniger diagnostisch für die spätere Erinnerungsleistung sind, da unter Umständen nicht die vollständige Antwort generiert wird (vgl. Begg et al., 1989; Spellman & Bjork, 1992); für die Einschätzung der Sicherheit der Antwort liegt dagegen die gesamte Information aus Einspeicher-, Abruf- und Erinnerungsprozess vor.

Deutlich häufiger als Resolutionsmaße wurde bei SUs die Kalibrierung untersucht. Hierbei wurde recht konsistent gefunden, dass Personen ihre Leistung retrospektiv zu optimistisch einschätzen (*Overconfidence effect*; z.B. D. Griffin & Brenner, 2004; Lichtenstein, Fischhoff & Phillips, 1982). Wenn die SUs sehr hoch ausfallen, ist die Überschätzungsrate am

deutlichsten (Lichtenstein et al., 1982). Gleichermaßen ist eine überoptimistische Einschätzung der eigenen Erinnerungsleistung bei schwierigen Aufgaben besonders ausgeprägt; schon bei Lichtenstein et al. (1982) zeigte sich, dass sich Probanden im Gegensatz dazu bei leichten Aufgaben eher unterschätzen (*Hard-easy effect*). Vor allem mit dem Ziel, die Überschätzung zu reduzieren, gab es bislang einige Ansätze. Koriat, Lichtenstein und Fischhoff (1980) ließen Probanden Argumente für bzw. gegen die gewählte Antwortalternative bei Multiple-Choice-Fragen zum Allgemeinwissen generieren; es zeigte sich, dass das Benennen von Gründen gegen die eigene Antwort die Überschätzung reduzierte (vgl. jedoch Allwood & Granhag, 1996). Noch deutlicher wirkte sich eine Rückmeldung über die erbrachte Leistung auf die SU-Kalibrierung aus. Schon nach einer einzigen Lerneinheit mit anschließendem Feedback erreichten die Teilnehmer einer Studie von Lichtenstein und Fischhoff (1980) eine angemessene Kalibrierungsleistung. Erklärungsansätze für den Overconfidence- bzw. den Hard-Easy-Effekt stammen vorwiegend aus Forschungsrichtungen zu Urteilsprozessen. Tversky und Kahneman (1974) postulierten, dass bei Entscheidungen verschiedene Heuristiken zur Anwendung kommen. Personen verlassen sich zum einen auf die Repräsentativität einer Antwort, d.h. die Wahrscheinlichkeit, dass eine Antwort richtig ist, wird auf Grundlage der Ähnlichkeit zu einem allgemeinen Muster bestimmt. Zum anderen ist bei der Abgabe von SUs die Verfügbarkeitsheuristik relevant: Je leichter und schneller jemandem eine Antwort einfällt, desto höher fallen die SUs aus (vgl. Abschnitt 3.1.4). Gigerenzer (1991) ging dagegen davon aus, dass Personen bei Fragen, deren Antwort sie nicht genau wissen, zur Einschätzung ihrer Sicherheit auf sog. *Probabilistic cues* zurückgreifen. Das heißt, sie prüfen verschiedene Hinweisreize auf ihre Wahrscheinlichkeit und entscheiden sich für die Antwort, die am wahrscheinlichsten ist.

In der kognitionspsychologischen Literatur finden sich in großer Übereinstimmung Ansätze, die als Grundlage für SUs ebenfalls inferentielle, heuristische Prozesse auf der Basis bestimmter Hinweisreize vermuten. Als wichtige Einflussfaktoren gelten dabei z.B. die Leichtigkeit, mit der ein Item abgerufen werden kann (Kelley & Lindsay, 1993), bzw. die Antwortlatenz (z.B. Koriat et al., 2006) oder die Vollständigkeit des Abrufs (Brewer, Sampaio & Barlow, 2005). Auch für SUs existiert ein kognitionspsychologisches Modell von Koriat (2012), welches verschiedene Erklärungsmöglichkeiten für das Zustandekommen dieser Urteile integriert. Das *Self-Consistency Modell* ist vor dem Hintergrund des Cue-utilization-Ansatzes (Koriat, 1997) zu sehen. Demnach sind für SUs vor allem menmonische Hinweisreize relevant, also erfahrungsbasierte, nicht zwangsläufig verbalisierbare Faktoren, die über die Leistung in der entsprechenden Aufgabe Rückschlüsse erlauben (z.B. Zugänglichkeit der

Erinnerung/*accessibility*, Abrufgeschwindigkeit etc). Koriat (2012) schließt damit nicht aus, dass auch analytischere, informationsbasierte Prozesse von Bedeutung sind, bei denen verschiedene Erinnerungsinhalte gegeneinander abgewogen werden; er betrachtet jedoch mnemonische Hinweisreize als deutlich einflussreicher für SUs. Ähnlich wie bei Gigerenzer (1991) wird der Lernende im Self-Consistency Modell als „intuitiver Statistiker“ gesehen. Das heißt, neben Hinweisreizen wie der Antwortlatenz oder der Vertrautheit mit dem Stimulus als Maße der Stärke der internen Repräsentation beruhen SUs ebenfalls auf der Reproduzierbarkeit einer Antwort. Je höher also die Wahrscheinlichkeit ist, dass dieselbe Antwort auch in einem zweiten und dritten Antwortversuch gewählt wird, desto höher fällt auch die Einschätzung der Antwortsicherheit aus. Zusammengefasst werden diese verschiedenen Einflussfaktoren unter dem Begriff „Selbstkonsistenz“, den Koriat (2012) als einen globalen, inhaltsfreien mnemonischen Hinweisreiz sieht, der die Übereinstimmung zwischen verschiedenen Abwägungsprozessen während des Abrufprozesses umfasst. Koriat (2012) bezieht das Self-Consistency Modell vorwiegend auf Fragen zum Allgemeinwissen, sodass es zu klären gilt, in welcher Form derartige Abwägungsprozesse sowohl bei einfacher strukturierten Paar-Assoziationsaufgaben, wie auch bei komplexeren Materialien wie Texten oder Filmen, Einfluss auf die Abgabe der SUs nehmen.

Insgesamt lässt sich festhalten, dass Probanden retrospektiv in der Regel sehr gut in der Lage sind, zwischen richtigen und falschen bzw. nicht erinnerten Antworten zu unterscheiden (Resolution). Bei Aufgaben zu freier Erinnerung gelingt ihnen dies in der Regel besser als bei Rekognitionsaufgaben; bei Paar-Assoziationsdesigns liegen die Genauigkeiten über denen bei komplexeren Materialien. Was die Einschätzung der Gesamtleistung betrifft, zeigen Analysen zur Kalibrierung dagegen, dass diese besonders bei schwierigen Aufgaben häufig übermäßig optimistisch ausfällt. Generell wurden SUs selten mit Resolutionsmaßen untersucht, sodass hier noch Forschungsbedarf besteht. Gerade Studien mit Paar-Assoziationsaufgaben könnten dazu beitragen, retrospektive und prospektive Überwachungsprozesse unmittelbar bezüglich ihrer Grundlagen zu kontrastieren. Denn auch wenn es plausibel erscheint, dass SUs im Vergleich zu anderen metakognitiven Maßen vorwiegend auf mnemonischen Hinweisreizen beruhen, steht hierzu die genaue Überprüfung noch aus.

3.1.5.2 Entwicklungspsychologische Befunde

Entwicklung im Vor- und Grundschulalter. Auch Studien mit Kindern kommen mehrheitlich aus dem Themenfeld Augenzeugengedächtnis und setzen dementsprechend häufig komplexere Materialien wie Texte oder Filme ein.

Eine der wenigen Studien, die mit Kindergartenkindern durchgeführt wurde, zeigte, dass schon Dreijährige grundsätzlich in der Lage sind, zwischen richtigen und falschen Antworten zu unterscheiden; diese Differenzierungsfähigkeit stieg jedoch bis zum Alter von fünf Jahren an (Lyons & Ghetti, 2011). Die bereits im Kapitel zu EOLs (s. Abschnitt 3.1.2) zitierte Studie von Visé und Schneider (2000) erfasste ebenfalls SUs sowie die Erinnerung an die letzte Leistung; der Befund, dass schon die vier- und sechsjährigen Teilnehmer retrospektiv ihre Leistung recht adäquat einschätzen konnten, macht deutlich, dass in dieser Altersgruppe kein grundlegendes Defizit in der metakognitiven Überwachung vorliegen kann. Dennoch lässt sich im Verlauf der Grundschulzeit eine Steigerung in der Qualität der SUs beobachten. So konnten z.B. Pressley, Levin, Ghatala und Ahmad (1987) zeigen, dass neun- bis elfjährige Kinder ihre Antworten bei einem Wortschatztest retrospektiv genauer einschätzten als sechs- bis achtjährige Kinder. Dies traf sowohl auf die globale Einschätzung über alle Items hinweg als auch – in geringerem Ausmaß – auf SUs auf Itemebene zu. Auch für das Lernen mit komplexeren Materialien, wie z.B. Filmepisoden, konnten Verbesserungen in der SU-Genauigkeit über das Grundschulalter hinweg nachgewiesen werden. Eine genauere Analyse bei Roebers (2002) ergab, dass diese Steigerung vor allem durch die niedrigeren SUs nach falschen Antworten zustande kommt (vgl. auch Roebers & Howie, 2003). Da in Studien zu SUs im Kindesalter häufig der Anwendungsbezug zum Augenzeugengedächtnis besteht, liegt der Fokus hier auch auf der besonderen Suggestibilität von Kindern. Es kommen also z.B. zusätzlich irreführende Fragen zum Einsatz. Die Ergebnisse zeigen, dass es Kindern im Grundschulalter sehr schwer fällt, zwischen korrekten (d.h. die falsche Suggestion wird abgelehnt) und falschen (d.h. der falschen Suggestion wird zugestimmt) Antworten bei diesem Frageformat zu differenzieren (z.B. Roebers, 2002; Roebers & Howie, 2003). Dieser Befund ist besonders ausgeprägt, wenn ausschließlich irreführende Fragen präsentiert werden (Roebers & Howie, 2003). Auch im Zusammenhang mit der Suggestibilität von Kindern gibt es Hinweise auf Alterseffekte. So waren bei von der Linden und Roebers (2006) die neunjährigen Teilnehmer im Gegensatz zu den siebenjährigen grundsätzlich in der Lage, bei irreführenden Fragen zwischen richtigen und falschen Antworten zu differenzieren.

Im Gegensatz zu den Studien mit Erwachsenen (s.o.) weisen Ergebnisse aus Arbeiten mit Grundschulern darauf hin, dass die Qualität von SUs bei Rekognitionsaufgaben höher ist als bei Aufgaben mit freien Antworten (z.B. Roebbers, 2002). Da das Wiedererkennen gerade für jüngere Kinder eine kognitiv weniger beanspruchende Aufgabe ist, sind hier schon Sechsjährige in der Lage, richtige von falschen Antworten zu unterscheiden (vgl. auch Brown, 1978). Eine adäquate Einschätzung der eigenen Leistung beim freien Erinnern gelang dagegen bei Roebbers, Gelhaar und Schneider (2004) erst der Gruppe der Sieben- bis Achtjährigen.

Weiterhin existieren auch Studien mit Grundschulern, die darauf abzielen, das Differenzierungsniveau bzw. die Genauigkeit der SUs zu steigern. In dieser Altersgruppe wurden dabei vor allem soziale Einflussfaktoren sowie die Wirkung von Feedback untersucht. In einer Studie von Schwarz und Roebbers (2006) profitierten die achtjährigen Teilnehmer in Bezug auf die Genauigkeit ihrer SUs nur dann, wenn ein erwachsenes Modell niedrige SUs angab, die Zehnjährigen dagegen verbesserten sich in der SU-Genauigkeit unabhängig von der Höhe der Sicherheit des Erwachsenen (vgl. jedoch von der Linden & Roebbers, 2006). Bezogen auf den Einfluss einer Rückmeldung fand sich bei einer Stichprobe von Zwölfjährigen nur dann eine Abnahme der Überschätzung in den SUs, wenn die Rückmeldung im Gegensatz zur gegebenen Antwort stand (*diskonfirmatorisches Feedback*; Allwood, Jonsson & Granhag, 2005).

Anders als im allgemeinspsychologischen Bereich existieren für Kinder recht wenige Studien, die Kalibrierungsmaße einsetzen. Keast, Brewer und Wells (2007) ließen Kinder im Alter von 11 Jahren sowie Erwachsene nach der Präsentation einer Video-Episode aus einer Reihe von Portraits den „Schuldigen“ identifizieren. Eine Analyse mit Kalibrierungskurven ergab, dass – wengleich auch die erwachsenen Probanden eine Tendenz zur Überschätzung zeigten – diese bei den Kindern deutlich größer ausgeprägt war. Dagegen zeigte eine Reanalyse der Daten von Roebbers und Howie (2003; Howie & Roebbers, 2007) für die nicht irreführenden Fragen, dass hier auch schon die Achtjährigen eine angemessene Kalibrierung aufwiesen. Wenn den Probanden jedoch nur irreführende Fragen gestellt wurden, fand sich für alle Altersgruppen eine Unterschätzung der eigenen Leistung. Diesen Befund erklärten die Autoren u.a. damit, dass die Abgabe eines niedrigen SU nach irreführenden Fragen eine niederschwellige Möglichkeit darstellt, dem Testleiter zu vermitteln, dass die Frage nicht beantwortet werden kann. Insgesamt gesehen kann jedoch geschlossen werden, dass auch bei Grundschulern eher eine Tendenz zur Überschätzung der eigenen Gedächtnisleistung auftritt, da das Format mit irreführenden Fragen eine recht spezifische Fragestellung darstellt, die nicht unbedingt auf andere Untersuchungsdesigns, die SUs einschließen, übertragbar ist.

Außerhalb der Forschung zum Augenzeugengedächtnis sind SUs im Kindesalter wenig untersucht worden. Vor allem Studien zum Paar-Assoziationslernen sind rar. Eine Ausnahme bildet die Arbeit von Roderer und Roebers (2010), bei der sieben- und neunjährige Probanden die Aufgabe hatten, jeweils das japanische Bildsymbol zu einem deutschen Wort zu lernen. Auch hier zeigte sich ein Anstieg in der Differenzierungsleistung mit dem Alter, der in niedrigeren SUs nach falschen Antworten begründet war. Zudem unterschieden nur die Neunjährigen in ihren SUs zwischen schwierigen und unbeantwortbaren Items. Die Ergebnisse aus Studien mit komplexeren Materialien scheinen demnach prinzipiell auf Paar-Assoziationsaufgaben übertragbar zu sein.

Bezüglich der Grundlagen von retrospektiven Urteilen wurde im Kindesalter vorwiegend die Antwortlatenz als mnemonischer Hinweisreiz diskutiert. So fand sich bei der Beantwortung von Fragen zum Allgemeinwissen ein negativer Zusammenhang zwischen der Zeit, die benötigt wurde, um sich für eine von zwei Antworten zu entscheiden, und der Höhe der SUs. Sowohl die Stärke dieses Zusammenhangs als auch die Validität der Antwortlatenz als Hinweisreiz nahmen mit dem Alter (von der zweiten bis zur fünften Klasse) zu (Koriat & Ackerman, 2010).

Entwicklung im Jugendalter. Ähnlich wie bei den anderen Überwachungsmaßen existieren auch für SUs kaum Studien, die Jugendliche einschließen. Für den Übergang von der Grund- zur weiterführenden Schule fanden Krebs und Roebers (2012), dass die elf- bis zwölfjährigen Teilnehmer bei Fragen zu einem Film im Vergleich zu den neun- bis zehnjährigen besser zwischen korrekt erinnerten und vergessenen Items differenzierten. In den Kalibrierungsmaßen ergab sich für alle Altersgruppen eine Überschätzung der eigenen Leistung, die bei den älteren Probanden etwas geringer ausgeprägt war. Interessanterweise zeigten sich auch im Vergleich von leistungsstarken vs. leistungsschwachen Probanden ähnliche Vorteile zugunsten ersterer, sodass davon ausgegangen werden muss, dass die gefundenen Alterseffekte zum Teil mit der kognitiven Leistungsfähigkeit konfundiert sind. Nichtsdestotrotz ist davon auszugehen, dass auch im Jugendalter noch Anstiege in der Überwachung der Antwortsicherheit zu finden sind. Denn auch in einer Arbeit zur visuellen Wahrnehmung von Weil et al. (2013) wurde eine Zunahme der SU-Genauigkeit (hier anhand von ROC-Kurven) bis ins späte Jugendalter nachgewiesen.

Entwicklung im höheren Erwachsenenalter. Zur Qualität von SUs im höheren Erwachsenenalter existiert eine Reihe von Studien mit zum Teil gemischten Befunden (vgl. Hertzog & Dunlosky, 2011). In der im Zusammenhang mit JOLs (s. Abschnitt 3.1.3) bereits erwähnten Studie von Lovelace und Marsh (1985) wurden auch retrospektive Urteile der Erinnerungsleistung erhoben. Im Gegensatz zu den JOLs ergaben sich hier keine Unterschiede zwischen den jüngeren und den älteren Erwachsenen: Beide Gruppen waren gleichermaßen in der Lage, zwischen korrekten und falschen Antworten beim Erinnern von Paarassoziationen zu unterscheiden. In neueren Arbeiten, die die SU-Resolution beim Lernen und freien Erinnern von Itempaaren oder Wortlisten untersuchen, fanden sich zum Teil Unterschiede zwischen jüngeren und älteren Erwachsenen in dem Sinne, dass die älteren Teilnehmer etwas niedrigere Gammakorrelationen aufwiesen als die jüngeren (Dodson, Bawa & Krueger, 2007; Kelley & Sahakyan, 2003). Bei Rekognitionsaufgaben und Versuchsdesigns zum impliziten Lernen wie auch bei Fragen zum Allgemeinwissen bzw. komplexeren, alltagsnäheren Aufgaben ergaben sich mehrheitlich vergleichbare Gammakorrelationen bei jüngeren und älteren Teilnehmern (Dodson et al., 2007; Hines, Touron & Hertzog, 2009; Marquié & Huet, 2000; Pliske & Mutter, 1996; Touron & Hertzog, 2004; Wahlheim, Dunlosky & Jacoby, 2011). Hinsichtlich der Kalibrierungsleistung existieren ebenfalls uneinheitliche Befunde: Crawford und Stankov (1996) fanden bei Intelligenztestaufgaben eine höhere Überschätzung der eigenen Leistung bei älteren Erwachsenen im Vergleich zu jüngeren; andere Studien ergaben jedoch keine Alterseffekte bezüglich der Kalibrierung (Kelley & Sahakyan, 2003; Pliske & Mutter, 1996). Insgesamt gesehen lassen sich die bislang vorliegenden Ergebnisse zur Qualität von SUs bei älteren Erwachsenen insofern zusammenfassen, als Personen dieser Altersgruppe grundsätzlich in der Lage sind, ihre eigene Erinnerungsleistung retrospektiv adäquat einzuschätzen. Es scheinen jedoch besonders bei Aufgaben zur freien Erinnerung, aber auch bei bestimmten Formen von Wiedererkennensparadigmen Verzerrungen in der Leistungseinschätzung von älteren Erwachsenen vorzuliegen. Ähnlich wie bei Kindern wurde auch bei Seniorenstichproben die Suggestibilität durch den Einsatz von irreführenden Fragen bzw. Antwortalternativen untersucht. Bei Kelley und Sahakyan (2003) zeigte sich in diesem Zusammenhang, dass ältere Erwachsene besonders bei Item-Paaren, bei denen das durch Anfangs- und Endbuchstaben markierte Zielwort (z.B. DO_ _ _ R) zwar eine Assoziation (DOCTOR) mit dem Stimulus (NURSE) nahe legt, jedoch in der Lernphase mit einem nicht zusammenhängenden Wort präsentiert wurde (DOLLAR), mehr Fehler machen und höhere Sicherheitenangaben als jüngere Erwachsene. Auch in einem Paradigma von Shing, Werkle-Bergner, Li und Lindenberger (2009) wies die Gruppe der älteren Erwachsenen deutlich mehr

high confidence errors auf als die anderen Altersgruppen (Kinder, Jugendliche, jüngere Erwachsene). Hier wurden den Probanden bei der Rekognition nicht nur bereits gelernte und neue Wortpaare vorgelegt, sondern auch aus den gelernten Items neu zusammengesetzte Wortpaare. Ergebnisse aus Augenzeugen-Gedächtnis-Studien bestätigen dieses Muster: Karpel, Hoyer und Toglia (2001) stellten nach einer Bildpräsentation eines Diebstahls sowohl irreführende als auch gültige Fragen und fanden ebenfalls bei älteren Erwachsenen im Vergleich zu jüngeren erhöhte SUs nach falsch-richtigen Antworten. Auch bei der Darbietung eines Videos über einen Überfall gaben die älteren Teilnehmer besonders nach falschen Antworten auf Suggestivfragen hohe SUs ab (Dodson & Krueger, 2006). Es wurden verschiedene Erklärungen für diese Befunde diskutiert. Zum einen könnten Defizite im *Source monitoring*, also Schwierigkeiten darin, zu erinnern und zu überwachen, in welchem Zusammenhang eine Information gelernt wurde, verantwortlich sein. Aufgrund dessen unterscheiden ältere Erwachsene weniger gut zwischen Information, die während des Versuchs gelernt wird, und bereits bekannter Information und sind anfälliger für Interferenzen aus späteren, konkurrierenden Ereignissen (vgl. Dunlosky & Metcalfe, 2009; Karpel et al., 2001). In eine ähnliche Richtung geht die Erklärung von Kelley und Sahakyan (2003), die postulierten, dass bei älteren Erwachsenen die genaue Erinnerung an Details (*recollection*) erschwert ist und sie sich dementsprechend vermehrt an der Vertrautheit (*familiarity*) mit den vorliegenden Reizen orientieren, wenn es um die Einschätzung der Sicherheit in eine Antwort geht (vgl. auch Shing et al., 2009). Auch auf hirnebene ließen sich Hinweise darauf finden, dass ältere Erwachsene besonders bei hoher Sicherheit bzw. bei Fehlern mit hoher Sicherheit eine veränderte neuronale Aktivität im Vergleich zu jüngeren Erwachsenen zeigen. In einer bildgebenden Studie von Chua, Schacter und Sperling (2009) zeigte sich z.B. bei jüngeren Erwachsenen eine höhere Aktivität im medialen Temporallappen bei Antworten mit hoher Sicherheit, bei den älteren Teilnehmern war dies nicht der Fall. Shing et al. (2009) diskutierten ihre Ergebnisse im Zusammenhang mit altersbedingten Veränderungen in Hippocampus-Strukturen.

3.1.5.3 Fazit

Insgesamt gesehen weist die Forschung zu SUs vor allem im Bereich der Augenzeugengedächtnis-Forschung eine breite Datenbasis auf. Hier zeigte sich, dass Probanden verschiedenen Alters grundsätzlich in der Lage sind, zwischen korrekten und falschen Gedächtnisinhalten zu unterscheiden. Im Vergleich zu JOLs liegen die mittleren SU-

Genauigkeiten bei komplexeren Lernmaterialien deutlich höher. In der Regel sind die retrospektiven Leistungseinschätzungen bei Erwachsenen genauer, wenn es um Aufgaben zur freien Erinnerung geht, Grundschüler scheinen bei Rekognitionsaufgaben besser abzuschneiden.

Außerhalb dieses eher angewandten Kontextes weist die Befundlage zu SUs noch Lücken auf. So fehlt es beispielsweise an Studien, die komplexere Materialien mit einer intentionalen Lerninstruktion und ohne irreführende Fragen einsetzen – ein Umstand, der besonders bei Kindern und älteren Erwachsenen zu Verschlechterungen der Überwachungsleistung führt. Weiterhin sind für die Altersgruppe der jüngeren Erwachsenen weitere Studien zu Resolutionsmaßen nötig, um die vorhandenen Befunde einordnen zu können. Im Entwicklungsverlauf ist es sinnvoll, auch bei Grundschulern SUs in Paar-Assoziationsparadigmen besser zu untersuchen, gerade auch, um in dieser Altersgruppe die retrospektive Überwachung mit prospektiven Maßen vergleichen zu können. Im Jugendalter fehlt es insgesamt an Studien zu SUs, um Entwicklungsprozesse bis zum jüngeren Erwachsenenalter nachvollziehen zu können. Da die Befundlage für die älteren Erwachsenen eher uneinheitlich und gerade die Rolle von „Recollection“- und Vertrautheitsprozessen bislang ungeklärt ist, ist es nötig, die Grundlagen für SUs und den Einfluss möglicher Kontextfaktoren gerade in dieser Altersgruppe genauer zu untersuchen.

3.2 Lernzeitallokation als Beispiel für metakognitive Kontrollprozesse

Neben den Überwachungsvorgängen gelten die sog. Kontrollvorgänge als zweite Komponente des prozeduralen Metagedächtnisses nach Nelson und Narens (1990, 1994). Bei Kontroll- bzw. Steuerungsprozessen greift das Individuum strategisch in den aktuell stattfindenden Lernvorgang ein (vgl. Koriat, Ackerman et al., 2014). Nelson und Narens (1990, 1994) postulierten verschiedene Maße für Kontrollvorgänge während der Lern-, Speicher- und Abrufphase. Da in der eigenen Arbeit lediglich in den Experimenten 5 und 6 erfasst wird, wie Probanden in Abhängigkeit von der Höhe der JOLs und des Anreizniveaus ihre Lernzeit einteilen (*Lernzeitallokation*), und andere Maße für Kontrollvorgänge keine Anwendung finden, sollen an dieser Stelle dementsprechend nur allgemein- sowie entwicklungspsychologische Ergebnisse aus diesem Bereich dargestellt werden. Andere Kontrollprozesse, wie die Itemauswahl oder Entscheidungen über die Zurückhaltung einer Antwort oder über die Beendigung des Abrufvorgangs, bleiben demnach unberücksichtigt.

Die Erfassung der Lernzeitallokation wird in der einschlägigen Literatur nicht vollkommen einheitlich gesehen. So unterscheiden Dufresne und Kobasiwaga (1989) beispielsweise zwischen der *hinreichenden* und der *differenziellen Lernzeitallokation*: Bei Aufgaben zur hinreichenden Lernzeitallokation geht es darum, inwiefern Personen in der Lage sind, zu entscheiden, wann ein Lernstoff ausreichend gelernt wurde; Aufgaben zur differenziellen Lernzeitallokation beziehen sich auf die Verteilung von Lernzeit in Abhängigkeit von Aufgabencharakteristika wie der Itemschwierigkeit. Da in den Experimenten 5 und 6 der vorliegenden Arbeit zwar einerseits die Formulierung der Instruktion auf eine hinreichende Zuweisung der Lernzeit abzielt („Lerne die Wortpaare so lange, wie du brauchst, um sie in etwa einer halben Stunde zu erinnern“), andererseits aber die Fragestellung und das Design der Studie sowie die Auswahl der Items schwerpunktmäßig auf differenzielle Aspekte abzielen, sollen im Folgenden gleichermaßen Ergebnisse aus beiden Bereichen vorgestellt werden.

3.2.1 Allgemeinpsychologische Befunde

Die frühen Studien zur Einteilung von Lernzeit auf Items verschiedener objektiver Schwierigkeit fanden übereinstimmend, dass schwierigere Items länger gelernt werden als einfachere (z.B. Le Ny, Denhiere & Le Taillanter, 1972; Zacks, 1969), sodass grundsätzlich davon ausgegangen werden kann, dass Erwachsene in der Lage sind, ihre Lernprozesse anhand von Eigenschaften des Lernmaterials zu regulieren. Auch etwas spätere Studien zum Zusammenhang zwischen der subjektiven Einschätzung der Item- bzw. Lernschwierigkeiten (zumeist durch JOLs erfasst) und der nachfolgenden Lernzeit fanden in der Regel negative Korrelationen: Items mit niedrigen JOLs wurden also länger gelernt und umgekehrt (z.B. Mazzoni & Cornoldi, 1993; Mazzoni, Cornoldi & Marchitelli, 1990). Die theoretische Erklärung für diese Befunde lieferte das *Diskrepanz-Reduktions-Modell* (Dunlosky & Hertzog, 1998), das besagt, dass Probanden sich eine individuelle *Norm of Study* setzen und dann entsprechend dieses Standards solange lernen, bis die Diskrepanz zwischen Soll- und Ist-Zustand maßgeblich verkleinert wurde. Bei schwierigen Items ist für diese Diskrepanzreduktion folglich deutlich mehr Lernaufwand nötig als bei einfachen.

Eine differenziertere Betrachtung dieses Zusammenhangs in komplexeren Untersuchungsdesigns zeigte jedoch, dass sich dieser nicht in allen Kontexten nachweisen lässt. Thiede und Dunlosky (1999) sowie Son und Metcalfe (2000) konnten z.B. zeigen, dass eine experimentelle Manipulation der „Norm of Study“ die Korrelation zwischen subjektiver Itemschwierigkeit und Lernzeit auch umkehren kann. Bei externer Setzung eines wenig

anspruchsvollen Ziels bzw. hohem Zeitdruck verwendeten die Teilnehmer deutlich mehr Lernzeit als auf leichter eingeschätzte Items. Thiede und Dunlosky (1999) erweiterten entsprechend das Diskrepanz-Reduktionsmodell um eine übergeordnete Ebene der Planung und postulierten, dass Individuen beim Lernen ihr jeweiliges Ziel ökonomisch, d.h. mit möglichst kleinem Aufwand erreichen möchten und demnach ihre Lernzeit strategisch auf die schwierigeren oder die leichteren Inhalten einteilen. Die Autoren wiesen jedoch auch darauf hin, dass nicht alle Versuchsdesigns bei den Teilnehmern auch planendes Verhalten hervorrufen.

Der Fokus auf die Itemschwierigkeit, die in den meisten Studien zur Lernzeitallokation als einflussgebender Faktor vorrangig berücksichtigt wird, muss also auf andere Aspekte ausgeweitet werden. In einer Weiterentwicklung des Modells zum selbstregulierten Lernen von Thiede und Dunlosky (1999) schlugen Ariel, Dunlosky und Bailey (2009) vor, die Itemschwierigkeit lediglich als einen Einflussfaktor auf die Lernzeiteinteilung zu sehen. Weiterhin ist in ihrem Modell der *Agenda-Based Regulation* nicht die Diskrepanzreduktion, sondern das Erfüllen einer „Agenda“ als wesentlich anzusehen. Beim Absolvieren dieses Programms wägen Individuen demnach verschiedene Faktoren, die zur Zielerreichung wichtig sind, gegeneinander ab. Experimentell zeigten Ariel et al. (2009), dass die Belohnungsstruktur einer Aufgabe maßgeblich und mehr als die differenzielle Itemschwierigkeit beeinflusst, welchen Inhalten Probanden wieviel Lernzeit widmen. Die Autoren zeigten für studentische Stichproben, dass die Teilnehmer unabhängig von der Item-Schwierigkeit denjenigen Items mehr Lernzeit zuwies, die mit höherer Wahrscheinlichkeit in der Testphase abgefragt wurden bzw. mit einem höheren Punktwert als Lernanreiz versehen waren.

Insgesamt gesehen zeigen die genannten Befunde also, dass Erwachsene in der Regel schwierigeren Inhalten mehr Lernzeit widmen; dieses Muster wurde für konkrete wie abstrakte Wortpaare, aber auch für komplexere Materialien wie verschiedene Textformen belegt. Wenn jedoch der interne Plan (die „Agenda“) eines Probanden spezifischere Ziele vorgibt, wie dies z.B. bei einer Zeitbegrenzung, der Veränderung der Abfragewahrscheinlichkeit, verschiedenen Punkten für die einzelnen Items oder der Setzung eines einfachen Zieles geschieht, verändern diese ihre Lernzeiteinteilung strategisch, so dass die Itemschwierigkeit nicht mehr den einzigen Einflussfaktor darstellt.

3.2.2 Entwicklungspsychologische Befunde

Entwicklung im Grundschulalter. Eine der ersten und wichtigsten Studien zur Lernzeitallokation im Grundschulalter beim Paar-Assoziationslernen stellt die Arbeit von Dufresne und Kobasigawa (1989) dar. Hier wurden Kindern im Alter zwischen 6 und 12 Jahren semantisch zusammenhängende sowie nicht zusammenhängende Bildpaare mit der Instruktion, diese bis zur perfekten Beherrschung zu lernen, präsentiert. Obwohl in den begleitenden Interviews zum deklarativen Metagedächtnis die meisten Kinder der jüngsten Altersgruppe grundlegendes Wissen zum Zusammenhang zwischen Lernzeit, Itemschwierigkeit und Lernleistung aufwiesen, differenzierten erst die 10-Jährigen in ihren Lernzeiten zwischen leichten und schwierigen Items; bei der Gruppe der 12-Jährigen war dieser Unterschied noch größer ausgeprägt. In einer Folgestudie von Kobasigawa und Metcalf-Haggert (1993), in der die Itemschwierigkeit durch die Verwendung von vollkommen unbekanntem Begriffen als schwierigen Items sehr salient gemacht wurde, verfügten bereits die sechsjährigen Probanden grundsätzlich über die Fähigkeit, in ihren Lernzeiten zwischen einfacheren und schwierigeren Items zu unterscheiden; auch hier waren jedoch die Drittklässler den Erstklässlern überlegen. Methodisch verbesserte Replikationen von Lockl und Schneider (2002b, 2003) bei Schülern der ersten, zweiten und dritten Klasse belegten ebenfalls den Alterstrend in dem Sinne, dass die jüngeren Probanden zwar grundsätzlich zwischen leichten und schwierigen Bildpaaren unterscheiden können, dieses Wissen jedoch nicht im gleichen Maße wie die älteren Kinder für die Gestaltung der Selbstregulationsprozesse nutzen. Vergleichbar mit der Arbeit von Ariel et al. (2009) variierten Lockl und Schneider (2004) im Folgenden neben der Itemschwierigkeit weitere Kontextfaktoren, um Situationen zu identifizieren, die Kindern die angemessene Lernzeiteinteilung erleichtern. Neben hoch assoziativen und niedrig assoziativen Bildpaaren kam in ihrer Studie mit sieben- und neunjährigen Kindern eine Instruktionsbedingung, in der die Genauigkeit, sowie eine Bedingung, in der die Lerngeschwindigkeit betont wurde, zum Einsatz. Weiterhin wurde bei der Hälfte der Probanden ein Anreiz von fünf Cent für jedes richtig erinnerte Bildpaar gesetzt. Es zeigte sich, dass nur die älteren Probanden in der Genauigkeitsbedingung länger lernten als die in der Geschwindigkeitsbedingung, was sich jedoch nicht in unterschiedlichen Erinnerungsleistungen widerspiegelte. Die Belohnung hatte in keiner der beiden Altersgruppen einen Effekt auf Lernzeit oder Leistung. Jüngere Kinder scheinen also weitgehend unabhängig von der Gestaltung der Lernbedingungen Schwierigkeiten zu haben, metakognitives Wissen über eine Aufgabe in entsprechende Regulationsprozesse umzusetzen. Es ist jedoch zu vermuten, dass sich in der älteren Gruppe der Effekt der Belohnung auf die differenzielle Lernzeiteinteilung gezeigt hätte, wenn diese

innerhalb der Aufgabe variiert worden wäre, d.h. wenn jedem Probanden sowohl Items mit als auch solche ohne Anreiz dargeboten worden wären. Denn auf diese Weise könnte die Belohnung deutlich salienter werden.

Insgesamt lässt sich festhalten, dass sich im Lauf der Grundschulzeit die strategische Allokation von Lernzeit auf Items verschiedener Schwierigkeitsgrade deutlich verbessert. Zu Beginn der Elementarstufe verfügen die meisten Kinder zwar über ein grundlegendes Wissen über den Zusammenhang zwischen Lernzeit, Itemschwierigkeit und Erinnerungsleistung; sie nutzen dieses jedoch noch nicht zur Regulation ihres Lernprozesses. Komplexere Variationen des Lernvorgangs durch angepasste Instruktionen oder Anreize wirken sich bei jüngeren Kinder nicht, bei älteren Kindern nur sehr bedingt auf die Lernzeitallokation aus.

Entwicklung im Jugendalter. Die frühen Studien zur differenziellen Lernzeitallokation bei Jugendlichen befassten sich ausschließlich mit dem Lernen aus Texten. Da die Lernzeiteinteilung in den eigenen Studien lediglich in Paar-Assoziationsparadigmen untersucht wird, sollen die Ergebnisse an dieser Stelle stark verkürzt dargestellt werden. Grundsätzlich zeigte sich über das Jugendalter hinweg eine Verbesserung in der Verteilung von Lernzeit auf wichtige im Gegensatz zu unwichtigeren Passagen (z.B. Brown & Smiley, 1978). Auch hier ergab sich bei den Fünftklässlern, dass diese zwar prinzipiell zwischen für das Verständnis wesentlichen und weniger wesentlichen Abschnitten unterscheiden konnten; diese Wichtigkeitseinschätzungen differenzierten sich bis zur siebten Klasse jedoch noch feiner aus (Brown & Smiley, 1977). Studien zur Lernzeitallokation in Paar-Assoziationsdesigns, die Jugendliche einschlossen, gab es bis vor wenigen Jahren nicht. Erst neuere Studien in der Arbeitsgruppe um Asher Koriat und Wolfgang Schneider begannen, diese Lücke zu schließen. So wiesen z.B. Hoffmann-Biencourt et al. (2010) nach, dass sich die differenzielle Lernzeiteinteilung auf leichtere und schwierigere Items erst ab der dritten Klasse zeigte; der Unterschied war bei den Teilnehmern der fünften bis zur achten Klasse noch ausgeprägter. Koriat, Ackerman et al. (2014) verwendeten leichte und schwierige Items, die zu je 50% mit einem Lernanreiz von 1 bzw. 5 Punkten markiert waren, und maßen die entsprechend verwendete Lernzeit. Fünft-, Sechst- und Neuntklässler verwendeten in fast allen Telexperimenten mehr Lernzeit auf die höher dotierten Items; dies schlug sich jedoch nicht in allen Fällen in einer für diese Items verbesserten Erinnerungsleistung wider. Da sowohl die Studie von Hoffmann-Biencourt et al. (2010) als auch die von Koriat, Ackerman et al. (2014) hohe Relevanz für die Frage, inwiefern Überwachungs- und Steuerungsprozesse wechselseitig

zusammenhängen, hat, sollen sie in Abschnitt 3.3 ausführlicher vorgestellt werden. Zusammenfassend lässt sich für das Jugendalter sagen, dass – trotz Fortschritten im Vergleich zum Grundschulalter – auch über diese Alterspanne hinweg noch Verbesserungen in der Selbstregulation, zumindest, was die Lernzeitallokation betrifft, zu beobachten sind. Die spärliche Studienlage lässt jedoch zum aktuellen Zeitpunkt noch keine weitreichenden Schlüsse zu.

Entwicklung im höheren Erwachsenenalter. In der Gruppe der älteren Erwachsenen divergieren die aktuellen Forschungsergebnisse bezüglich der Frage, inwiefern diese Gruppe sich in ihrer Lernzeitallokation von jüngeren Erwachsenen unterscheidet. So boten Dunlosky und Connor (1997) älteren und jüngeren Probanden semantisch unzusammenhängende Wortpaare dar und untersuchten, wie sich die subjektiv (anhand von JOLs) eingeschätzte Schwierigkeit der jeweiligen Items in einem Trial auf die Lernzeit im nächsten Trial auswirkte. Auch wenn die Korrelationen zwischen JOLs und Lernzeit in beiden Altersgruppen signifikant von Null verschieden waren, erreichten die jüngeren Erwachsenen doch deutlich höhere Werte als die älteren Erwachsenen. Die Interpretation, dass Senioren die Ergebnisse aus Überwachungsprozessen weniger als jüngere Erwachsene zur Steuerung von Lernvorgängen nutzen, also eine Art „Nutzungsdefizit“ aufweisen, konnte in nachfolgenden Studien jedoch nicht bestätigt werden. So erbrachten Hines et al. (2009) mit einem im Vergleich zu Dunlosky und Connor (1997) leicht veränderten Versuchsdesign den Nachweis, dass sich ältere und jüngere Erwachsene nicht in der Lernzeitallokation unterscheiden. Zwar verwendeten auch sie nicht assoziierte Itempaare, maßen aber die Wiedererkennens- und nicht die Erinnerungsleistung, zogen als Maß für die subjektive Schwierigkeit ebenfalls die SUs aus dem vorherigen Lerndurchgang heran und passten die Ausgangsleistung der beiden Gruppen durch verschieden lange Präsentationszeiten im ersten Durchgang an. Price, Hertzog und Dunlosky (2010) schließlich zogen zur Bestimmung der subjektiven Schwierigkeit EOLs heran und variierten die objektive Itemschwierigkeit in drei Abstufungen (leicht, mittel, schwer). Die Autoren stellten den Probanden die Aufgabe, sich selbst ein Lernziel zu setzen. Dies fiel bei den älteren Teilnehmern niedriger aus; dementsprechend widmete diese Gruppe den leichten und mittelschweren Items mehr Lernzeit, wohingegen die jüngeren Erwachsenen die schwierigen Items länger lernten. Das heißt, obwohl Unterschiede in der Lernzeiteinteilung zwischen beiden Altersgruppen bestanden, konnten diese auf eine vergleichbare Strategie („Erreiche dein individuelles Lernziel“) zurückgeführt werden. In einem zweiten Experiment verglichen Price et al. (2010) zusätzlich den Einfluss von Anreizen. Den Wortpaaren wurden

hohe, mittlere oder niedrige Punktwerte zugeordnet. Es zeigte sich, dass die älteren Erwachsenen in ihrer differenziellen Lernzeitzuweisung deutlich weniger sensitiv auf die Anreize reagierten als die jüngeren Erwachsenen. Die Autoren interpretierten dies im Kontext des Modells der Agenda-Based Regulation (Ariel et al., 2009; s.o.) in dem Sinne, dass die älteren Probanden auch im zweiten Experiment vorrangig das Ziel verfolgten, schwierige Items zu vermeiden und leichte bevorzugt zu lernen. Es kann jedoch aufgrund dieses Ergebnisses nicht davon ausgegangen werden, dass ältere Erwachsene grundsätzlich Schwierigkeiten haben, verschiedene Punktwerte von einzelnen Items beim Lernvorgang zu berücksichtigen. Studien aus der Arbeitsgruppe um Castel (Castel, Benjamin, Craik & Watkins, 2002; Castel, Humphreys et al., 2011) belegten, dass ältere Erwachsene beim Lernen von Einzelworten durchaus in der Lage sind, Anreizwerte als Grundlage für die Steuerung des Lernprozesses heranzuziehen. Es scheint also so zu sein, dass ältere Erwachsene sich in vergleichbarer Weise wie jüngere Erwachsene übergeordnete Ziele für ihr Lernergebnis setzen und entsprechend ihre Ressourcen auf dieses Ziel hin einsetzen; das Ziel wird jedoch an die selbst eingeschätzte Leistungsfähigkeit angepasst, die bei älteren Erwachsenen niedriger ausfällt als bei jüngeren. Unterschiede zwischen den Altersgruppen scheinen dahingehend zu bestehen, dass es älteren Erwachsenen u.U. schwerer fällt, mehrere Faktoren (z.B. Itemschwierigkeit und Anreizwert) gleichzeitig bei der Anpassung der Lernzeit zu berücksichtigen sowie differenziell auf verschiedene Itempaare zu reagieren, die sich nicht in ihrer normativen Schwierigkeit unterscheiden (Dunlosky & Connor, 1997).

3.2.3 Fazit

Insgesamt gesehen sind Personen ab etwa der Mitte der Grundschule in der Lage, ihre Lernzeit strategisch einzuteilen. Jüngere Kinder können zwar schwierigere von leichteren Items unterscheiden, setzen dies jedoch weniger in der Steuerung ihres Lernverhaltens um. Das heißt, die Kontrollkomponente des prozeduralen Metagedächtnisses scheint sich etwas später zu entwickeln als die Überwachungskomponente. Überdies zeigen Studien zum Jugendalter, dass auch Verbesserungen noch über eine längere Altersspanne stattfinden. Defizite in der strategischen Lernzeitallokation, die zum Teil bei älteren Erwachsenen gefunden wurden, müssen vor dem Hintergrund betrachtet werden, dass diese Altersgruppe im Vergleich zu jüngeren Erwachsenen unter Umständen auch andere Ziele aufweist und somit ihre Ressourcen innerhalb dieser Zielsetzung auch ökonomisch einteilt. Inwiefern Ergebnisse aus Überwachungsprozessen in die Steuerung der Lernzeit einfließen und inwiefern auch

umgekehrt Kontroll- auf Überwachungsvorgänge einwirken, soll im nächsten Kapitel beleuchtet werden.

3.3 Zusammenhang zwischen Überwachungs- und Kontrollprozessen

Wie intuitiv einsichtig und inzwischen vielfach experimentell bestätigt, stellen Überwachungs- und Kontrollvorgänge keine voneinander unabhängigen Prozesse dar. Während bis vor etwa zehn Jahren mehrheitlich davon ausgegangen wurde, dass Informationen aus der Überwachung beeinflussen, wie Lernvorgänge gesteuert werden (z.B. Nelson & Leonesio, 1988), gibt es heute vermehrt auch Befunde dahingehend, dass Kontrollprozesse Überwachungsurteile beeinflussen können (z.B. Koriat et al., 2006). Belege für beide Wirkrichtungen in verschiedenen Altersstufen und aktuelle Befunde über komplexere wechselseitige Zusammenhänge sollen im Folgenden dargestellt werden.

3.3.1 „Monitoring-affects-control“-Modell

Viele der Studien, die zum Ergebnis haben, dass Überwachungsprozesse Kontrollvorgänge beeinflussen, wurden bereits in Abschnitt 3.2 vorgestellt. Deshalb sollen diese hier lediglich in verkürzter Form noch einmal aufgegriffen und unter dem übergeordneten Gesichtspunkt des Zusammenhangs zwischen beiden Prozessen reflektiert werden. Es geht hierbei also nicht vorrangig um die Frage, wie Lernzeit, Itemauswahl oder andere Kontrollprozesse differenziell gesteuert werden, sondern darum, wie der Zusammenhang zu Überwachungsurteilen ausfällt. Bereits Hart (1965; s. Abschnitt 3.1.4) betonte im Zusammenhang mit FOK-Urteilen, dass diese den Suchprozess determinieren. Besagt das FOK, dass ein Item mit hoher Wahrscheinlichkeit im Gedächtnis gespeichert und potenziell erinnerbar ist, wird der Abrufprozess fortgesetzt; ist dies nicht der Fall, so wird die Suche beendet. Nelson und Leonesio (1988) prägten in ihrer Arbeit für diesen Zusammenhang den Begriff *Monitoring-affects-control-Hypothese* (auch: *MC-Modell*) und berichteten ebenfalls Befunde zu jüngeren Erwachsenen, die diesen Ansatz bestätigen. Für EOLs und – zu einem geringeren Maße – auch für FOKs fanden die Autoren substantielle negative Korrelationen zwischen der Höhe der Überwachungsurteile und der selbst eingeteilten Lernzeit. Items, die als schwieriger eingeschätzt wurden, lernten die Teilnehmer demnach länger und umgekehrt. Dieser meist mittelhohe Zusammenhang zwischen Überwachungsurteil und nachfolgender Lernzeit konnte auch für JOLs vielfach nachgewiesen werden (z.B. Mazzoni & Cornoldi, 1993; Thiede & Dunlosky, 1999; für einen Überblick s. Son

& Metcalfe, 2000). Somit kann grundsätzlich als gesichert gelten, dass Personen die Kontrolle ihres Lernprozesses nach Rückmeldungen aus Überwachungsvorgängen gestalten. Fälle, in denen subjektiv schwierigere Items weniger lang gelernt bzw. mit geringerer Wahrscheinlichkeit zum Wiederlernen ausgewählt wurden, können damit erklärt werden, dass hier entweder äußere Anreize, z.B. durch die Vergabe von verschiedenen hohen Punktwerten für die einzelnen Items (z.B. Ariel et al., 2009), oder die externe Setzung eines anderen Ziels, als alle Items perfekt zu lernen (z.B. Thiede & Dunlosky, 1999), Einfluss nahmen.

Die bereits erwähnte Studienreihe von Lockl und Schneider (2002b, 2003, 2004; s. Abschnitt 3.2.2) belegte bei der Darbietung von leichten und schwierigen Items ebenfalls signifikante Zusammenhänge zwischen JOLs und Lernzeiten. Diese fielen z.B. in Lockl und Schneider (2003) zu Beginn der Grundschulzeit noch recht gering aus (Gammakorrelation von .22 für Erstklässler), die Drittklässler wiesen bereits mittelhohe Werte auf (.40). Weiterhin zeigte eine recht neue Studie von Destan, Hembacher, Ghetti und Roebers (2014), dass sechs- und siebenjährige Kinder im Vergleich zu fünfjährigen mehr Zeit auf Items mit niedrigeren JOL-Werten aufwendeten. Das heißt, der Anstieg des Zusammenhangs zwischen Überwachungs- und Kontrollvorgängen findet bereits zu Beginn der Grundschulzeit statt und wird im Verlauf des Grundschulalters enger. Bei der einzigen Studie in diesem Kontext mit älteren Erwachsenen, die Gammakorrelationen zwischen Überwachungsurteil und Kontrollmaß berichtet, handelt es sich um die von Dunlosky und Connor (1997): Hier erreichten die älteren Probanden Korrelationen um .20 (Studie 1), die jüngeren um .40. Aus diesen Studien kann jedoch nicht geschlossen werden, dass ältere Erwachsene Rückmeldungen aus Kontrollprozessen generell weniger gut berücksichtigen. Die niedrigeren Korrelationen können auch durch die vergleichsweise höheren Anforderungen an die älteren Erwachsenen oder durch altersabhängig andere individuelle Zielsetzungen erklärt werden (vgl. Kapitel 3.2.2).

Insgesamt gesehen ist die empirische Basis für die Monitoring-affects-control-Hypothese (Nelson & Leonesio, 1988) recht robust. In einer Zusammenfassung dieser Befunde konstatierten Koriat et al. (2006), dass allen Situationen, in denen Informationen aus Überwachungsvorgängen für die Anpassung der Steuerungsprozesse genutzt werden, gemeinsam ist, dass diese Lernvorgänge *zielorientiert* sind (vgl. Diskrepanz-Reduktionsansatz von Dunlosky & Hertzog, 1998, sowie Agenda-Based Regulation von Ariel et al., 2009; s. Abschnitt 3.2.1). Diese Zielorientierung kann durch die Instruktion und durch das verwendete Material (starke Unterschiede in der normativen Itemschwierigkeit, externes Ziel, unterschiedliche Anreizwerte für die einzelnen Items) verschieden stark betont werden. Für zielorientierte Lernvorgänge konnte für Kinder ab der Grundschulzeit, aber auch für ältere

Erwachsene belegt werden, dass sie in diesen Fällen ausgehend von Informationen aus den Überwachungsvorgängen ihre Steuerungsprozesse gestalten.

3.3.2 „Control-affects-monitoring“-Modell

Obwohl also zielorientierte Lernvorgänge, bei denen die „Monitoring-affects-control“-Hypothese zum Tragen kommt, in den meisten Altersgruppen – mit Berücksichtigung altersbedingter Unterschiede in der Stärke des Zusammenhangs – gut belegt sind, wiesen Koriat et al. (2006) darauf hin, dass nicht alle Lernprozesse auf diese Weise ablaufen. Die Autoren stellten fest, dass die Lernzeit, die im Modell von Nelson und Narens (1990, 1994) ja ein Maß für die Steuerung ist, in bestimmten Settings auch als Hinweisreiz für die Höhe von Überwachungsurteilen (hier: JOLs) genutzt werden kann. Dies ist vor allem bei Lernvorgängen der Fall, in denen sich die Probanden die Lernzeit selbstständig einteilen. In Abgrenzung zu den oben beschriebenen zielorientierten Lernprozessen nennen Koriat et al. (2006) die Vorgänge, bei denen die Lernzeit bzw. allgemeiner gesprochen, der „Lernaufwand“ (*memorizing effort*) als Grundlage für die JOLs herangezogen wird, *datenorientiert (data-driven)*. Das heißt, der Lernende widmet den einzelnen Items so viel Zeit, wie diese „einfordern“ und nutzt diese Zeit als Hinweisreiz (*Memorizing effort-Heuristik*). Je weniger Lernzeit demnach in ein Item investiert werden muss, desto leichter bleibt es im Gedächtnis und desto höher wird also das JOL ausfallen. Entsprechend wurde die zugrunde liegende Heuristik in einem späteren Artikel auch als *Easily learned, easily remembered-Heuristik* bezeichnet (Koriat et al., 2009a). Der Wirkzusammenhang zwischen Überwachungs- und Kontrollvorgängen vollzieht sich hier also im Vergleich zur „Monitoring-affects-control“-Hypothese in umgekehrter Reihenfolge (*Control-affects-monitoring-Hypothese; CM-Modell*). Wenn das CM-Modell gilt, sollten folglich die JOLs mit ansteigender Lernzeit kleiner werden; wenn das MC-Modell gilt, sollten lang gelernte Items mit höheren JOLs einhergehen.

Empirische Belege für das CM-Modell in datenorientierten Lernprozessen erbrachten Koriat und Ma'ayan (2005) sowie Koriat et al. (2006) zunächst für studentische Stichproben. Sie wiesen beim Lernen von hoch und niedrig assoziativen Wortpaaren mit selbst eingeteilter Lernzeit negative Korrelationen zwischen JOLs und Lernzeit von im Mittel $-.21$ nach; bei festgelegter Lernzeit fanden die Autoren eine positive Korrelation zwischen JOLs und Lernzeit von im Mittel $.11$ (Koriat & Ma'ayan, 2005). Bei Koriat et al. (2006) wurden in der Folge (ebenfalls bei leichten und schwierigen Itempaaren) Werte bis $-.42$ (Experiment 1) für den Zusammenhang zwischen JOLs und Lernzeit im CM-Modell gefunden. Einem möglichen

Kritikpunkt an den gewählten Studiendesigns, nämlich die Tatsache, dass die Korrelation zwischen JOLs und Lernzeit durch die vorgegebene intrinsische Itemschwierigkeit mediiert wird, widmete Koriat (2008) sich in einer Folgestudie: Den Probanden wurden nun Wortpaare, die keine inhaltliche Verknüpfung aufwiesen, zum Lernen vorgelegt. Auch hier zeigten sich kleine, aber dennoch signifikante negative Zusammenhänge zwischen JOLs und Lernzeit. Auch wenn der Lernaufwand durch die Anzahl der Lerndurchgänge mit jeweils festgelegter Lernzeit erhoben wurde, konnte das CM-Modell bestätigt werden.

Entwicklungspsychologische Studien zur Gültigkeit des CM-Modells sprechen – bei selbst eingeteilter Lernzeit – für eine Zunahme des Zusammenhangs zwischen Lernaufwand und JOLs. Koriat et al. (2009a) ließen Schüler der ersten bis sechsten Klasse assoziierte und nicht assoziierte Wortpaare lernen und fanden für die Teilnehmer ab der dritten Klasse signifikante Korrelationen zwischen JOLs und Lernzeit (zwischen $-.18$ und $-.29$; Experiment 1). Auch eine alternative Auswertungsmethode, bei der nicht die Korrelation berechnet, sondern die Items individuell nach dem Mediansplit der Lernzeiten in länger und kürzer gelernte Paare unterteilt und verglichen wurden, erbrachte nur bei den älteren Kindern ab der dritten Klasse signifikante Unterschiede in den JOLs. In einer zweiten Studienreihe (Koriat et al., 2009b) wurde der Lernaufwand mittels der Anzahl der Lerndurchgänge erfasst. Beim Lernen von leichten und schwierigen Wortpaaren zeigten hier sowohl die Zweit- als auch die Viertklässler mittelhohe negative Zusammenhänge zwischen Lerndurchgängen und JOLs, die sich zwischen den beiden Altersgruppen nicht unterschieden (Experiment 1). In einem zweiten Experiment kamen nur schwierige Wortpaare zum Einsatz. Hier zeigten nur die Viertklässler Nachweise für Selbstregulationsvorgänge nach dem CM-Modell. Weitere Untersuchungen in einer breiteren Altersspanne ergaben, dass beim Lernen von leichten und schwierigen Bildpaaren zwar schon die jüngeren Teilnehmer (erste bis vierte Klasse) niedrige negative Korrelationen zwischen JOLs und Lernzeit aufwiesen, diese jedoch bei den älteren Schülern (fünfte bis achte Klasse) deutlich größer ausfielen. Eine Auswertung nach dem Mediansplit der Lernzeit bestätigte diese Befunde (Hoffmann-Biencourt et al., 2010). Insgesamt ist an den entwicklungspsychologischen Studien kritisch zu sehen, dass keine Anpassung der Basisschwierigkeit der Aufgaben erfolgte. Teilnehmer aller Altersgruppen bekamen die gleiche Anzahl an Wortpaaren präsentiert, obwohl daraus Altersunterschiede in den Erinnerungsraten resultierten (Hoffmann-Biencourt et al., 2010; Koriat et al., 2009a, 2009b). So könnten Unterschiede in der Ausprägung des CM-Modells in den verschiedenen Altersgruppen z.T. auch daraus resultieren, dass den jüngeren Teilnehmern durch die größere Anforderung der Aufgabe die Nutzung von adäquaten Hinweisreizen noch schwerer fiel. Eine Anpassung der Aufgaben-Schwierigkeit ist besonders

dann von Bedeutung, wenn eine große Altersspanne untersucht wird, wie es in den eigenen Studien der Fall ist. Weiterhin existieren bislang kaum Ergebnisse mit Jugendlichen sowie keine Studien, die über das jüngere Erwachsenenalter hinausgehen, sodass über die Altersverläufe bislang nur eingeschränkt Aussagen getroffen werden können.

3.3.3 Wechselseitige Zusammenhänge zwischen Überwachungs- und Kontrollprozessen

Es kann also zum einen als gesichert gelten, dass Überwachungsprozesse die Selbstregulation beim Lernen beeinflussen. Zum anderen zeigen die dargestellten Befunde, dass auch der umgekehrte Wirkzusammenhang in datenorientierten Lernsituationen (zumindest ab etwa der Mitte der Grundschulzeit) greift. Schon im Artikel von 2006 wiesen Koriat et al. darauf hin, dass sich beide Einflussrichtungen nicht gegenseitig ausschließen, sondern auch während eines einzigen Lernvorganges simultan auftreten können. So mag sich beispielsweise ein Student, der für eine Prüfung lernt, dazu entscheiden, einzelnen Lerninhalten mehr Zeit als anderen zu widmen; dies sowohl deshalb, weil diese Inhalte besonders schwer (Datenorientierung), aber auch, weil sie besonders wichtig für den späteren Prüfungserfolg sind (Zielorientierung). Diesen wechselseitigen Zusammenhang von MC- und CM-Modell überprüften die Autoren experimentell, indem sie beim Lernen von Wortpaaren mit selbst eingeteilter Lernzeit der Hälfte der Items einen Anreizwert von einem Punkt, der anderen Hälfte von drei Punkten zuwies (Experiment 5). Auf diese Weise kann eine zielorientierte Zuweisung von Lernzeit dann beobachtet werden, wenn die Probanden den Items mit höheren Werten im Vergleich zu denen mit niedrigeren Werten mehr Lernzeit widmen. Innerhalb derselben Belohnungskategorie sollten jedoch datenorientierte Prozesse zu beobachten sein, insofern als hier Items mit niedrigeren Lernzeiten höhere JOLs zugewiesen werden sollten. Diese Hypothesen zur simultanen Verknüpfung von CM- und MC-Modell konnten Koriat et al. (2006) bei jüngeren Erwachsenen belegen. Der flexible Wechsel zwischen daten- und zielorientiertem Vorgehen innerhalb eines Lernprozesses setzt voraus, dass die Verbindung zwischen Überwachungs- und Steuerungsprozessen jeweils über verschiedene zugrunde liegende Attributionen vermittelt wird. Koriat und Nussinson (2009) machten diesen Attributionsprozess explizit, indem sie die Teilnehmer ihrer Studie baten, den Korrugator zu kontrahieren, um das Gefühl der Anstrengung zu veranschaulichen. In Experiment 1 sollten die Probanden diese Anstrengung auf daten-orientierte Prozesse beziehen, in Experiment 2 auf zielorientierte Vorgänge. Im Vergleich zu einer Kontrollgruppe zeigte sich, dass der

anstrengungsbezogene Gesichtsausdruck bei daten-orientierten Prozessen zu geringeren JOLs führte, bei zielorientierten Prozessen dagegen zu höheren JOLs. In einer neueren Studie beeinflussten Koriat, Nussinson und Ackerman (2014) den Attributionsprozess auf eine andere Weise: Um datenorientierte Prozesse hervorzurufen, sollten die Probanden den Lernaufwand einschätzen, den ein Item „einfordert“; um zielorientierte Prozesse hervorzurufen, sollten die Probanden einschätzen, für welchen Lernaufwand sie sich bei einem Item entschieden, d.h. wie viel Lernzeit sie aktiv zuwiesen. Die Ergebnisse dieses Experiments, bei dem nicht zusammenhängende Wortpaare mit freier Lernzeit memoriert werden sollten, belegten die Bedeutung der jeweiligen Attributionsprozesse.

Für das Kindes- und Jugendalter gibt es bislang erst eine einzige Studie, die sich mit der simultanen Verknüpfung von CM- und MC-Modell beschäftigt (Koriat, Ackerman et al., 2014). Hier wurde repliziert, dass Fünft- und Sechstklässler grundsätzlich datenorientiertes Lernverhalten zeigen können. Sie waren jedoch nicht in der Lage, bei der Darbietung leichter und schwieriger Items, die je zur Hälfte mit hohen und niedrigen Punktwerten markiert waren, flexibel zwischen CM- und MC-Modell zu wechseln, sondern zeigten in diesem Fall nur eine datenorientierte Vorgehensweise, d.h. die Itemschwierigkeit wurde als Grundlage für die JOLs genutzt, nicht aber der Anreizwert. Wenn jedoch die Zielorientierung salient gemacht wurde, indem nur nicht assoziative Wortpaare dargeboten und mit verschiedenen hohen Punktwerten versehen wurden, konnte auch hier das MC-Modell nachgewiesen werden. Gleichmaßen fanden sich auch deutliche Belege für das CM-Modell, wenn die Anreizwerte konstant gehalten wurden. Neuntklässler waren in dieser Studie besser in der Lage, zwischen MC- und CM-Modell simultan zu wechseln. Auch für diese Fragestellung existieren bislang keine Studien, die ältere Erwachsene einschließen. Es ist hier allerdings zu erwarten, dass diese Altersgruppe etwas weniger sensitiv auf eine Betonung der Zielorientierung mit Anreizen reagiert (vgl. Price et al., 2010), was insgesamt den flexiblen Umgang mit daten- und zielorientierten Lernprozessen erschweren sollte.

3.3.4 Fazit

Die Befunde zu ziel- und datenorientierten Lernvorgängen zeigen, dass Personen prinzipiell ab Beginn des Grundschulalters in der Lage sind, Informationen aus Überwachungsprozessen für die Steuerung des Lernens zu nutzen. Dieser Zusammenhang wird jedoch im Lauf der Grundschulzeit deutlich enger. Studien mit älteren Erwachsenen weisen das MC-Modell auch

in dieser Altersgruppe nach. Ob der Zusammenhang hier wieder etwas geringer ausfällt, muss noch genauer geklärt werden.

Datenorientiertes Lernen scheint sich insgesamt etwas später, das heißt, ca. ab der dritten Klasse zu entwickeln. Bei Viertklässlern wurde das CM-Modell auch dann gefunden, wenn die Itemschwierigkeit konstant gehalten wurde. Der Wirkzusammenhang scheint also im Verlauf der späten Grundschul- und Sekundarschulzeit noch enger zu werden. Befunde zu älteren Erwachsenen liegen bislang nicht vor.

Das simultane Vorliegen von ziel- und datenorientiertem Lernen scheint über Attributionsvorgänge vermittelt zu werden: Beim zielorientierten Lernen wird eine hohe Anstrengung mit einem guten Lernerfolg in Verbindung gebracht, bei datenorientierten Lernen dagegen mit anspruchsvollem Lernmaterial. Der flexible Wechsel zwischen beiden Modellen scheint eher komplex zu sein und konnte erst bei Schülern ab der neunten Klasse vollständig belegt werden. Allerdings existieren bislang nur wenige Befunde, insbesondere keine zu älteren Erwachsenen.

3.4 Zusammenhang zwischen prozeduralem Metagedächtnis und Gedächtnisleistung

Wenngleich der Zusammenhang zwischen Überwachungs- und Steuerungsleistungen und der Gedächtnisleistung nicht im Mittelpunkt dieser Arbeit steht, sollen dennoch ausgewählte Befunde dazu vorgestellt werden, da dies die praktische Relevanz von prozeduralen Metagedächtnisvorgängen unterstreicht. Auch hier werden die Ergebnisse getrennt nach Überwachungs- und Kontrollvorgängen vorgestellt.

3.4.1 Überwachungsleistung und Gedächtnisleistung

Die frühen Studien zum Zusammenhang zwischen Überwachungs- und Testleistung erbrachten widersprüchliche Ergebnisse. So resümierten Cavanaugh und Perlmutter (1982), dass kein starker empirischer Bezug zwischen beiden Maßen zu beobachten sei; Maki und Berry (1984) wiesen dagegen nach, dass Probanden, deren Gedächtnisleistung oberhalb des Medians lag, auch eine höhere Überwachungsgenauigkeit zeigten als solche, deren Leistung unterhalb des Medians angesiedelt war. Die Vergleichbarkeit dieser Ergebnisse wird dadurch erschwert, dass sich die Metagedächtnisforschung zu diesem Zeitpunkt noch nicht auf einheitliche Modelle und

Operationalisierungen bezog. Eine Metaanalyse von Schneider (1989) ergab für das Kindes- und Jugendalter mittelhohe Korrelationen für verschiedene Überwachungsmaße und die Testleistung von .39 bis .44, woraus grundsätzlich gefolgert werden kann, dass ein substanzieller Zusammenhang vorhanden ist. Auch etwas neuere Studien bestätigten diesen Befund bei jüngeren Erwachsenen. So fand Thiede (1999) beim Lernen von Itempaaren, dass die spätere Erinnerungsleistung durch die Überwachungsleistung signifikant vorhergesagt wurde. Auch für das Lernen von Texten zeigte sich dieses Muster: Bei Thiede et al. (2003) erreichten Teilnehmer, die durch die Instruktion, Schlüsselwörter zum Text zu generieren, bessere JOL-Genauigkeiten aufwiesen, auch bessere Gedächtnisleistungen. Die Vorteile von Personen, die gute Überwachungsfähigkeiten aufweisen, konnten auch in einem angewandten Setting bei Studenten gezeigt werden (Nietfeld, Cao & Osborne, 2005): Hier gingen Verbesserungen in der Überwachungsleistung nach einem Semester Training auch mit besseren akademischen Leistungen einher. Bei älteren Erwachsenen fanden Dunlosky, Kubat-Silman und Hertzog (2003), dass eine Unterweisung in Selbsttests als Möglichkeit, die metakognitive Überwachung zu verbessern, positive Effekte auf die spätere Gedächtnisleistung hat. Insgesamt zeigen diese ausgewählten Befunde, dass eine gute Überwachungsleistung in der Regel mit einer höheren Gedächtnisleistung einhergeht, und dass dies für die in den eigenen Studien untersuchte Altersspanne (7 bis 80 Jahre) sowie in einer Vielzahl von Designs (Labor- und Feldstudien; Paar-Assoziationslernen und Textlernen) der Fall zu sein scheint.

Ähnlich wie beim deklarativen Metagedächtnis (s. Abschnitt 2.1.3) geht man auch bei der metakognitiven Überwachung nicht ausschließlich von einem direkten kausalen Einfluss auf die Erinnerungsleistung aus. Stattdessen wird postuliert, dass die gute Nutzung von Informationen aus den Überwachungsvorgängen die adäquate Steuerung von Lernprozessen erleichtert, was schließlich die Gedächtnisleistung verbessert (z.B. Thiede, 1999). Das heißt, die „Monitoring-affects-control“-Hypothese müsste um einen dritten Bestandteil – „*affects performance*“ – erweitert werden. Studien, die den Zusammenhang zwischen Lernzeitallokation und Gedächtnisleistung untersuchen, sollen im nächsten Abschnitt vorgestellt werden.

3.4.2 Kontrollprozesse und Gedächtnisleistung

Wenn man also der „Monitoring-affects-control“-Hypothese folgt, sollte es für einen effektiven Lernvorgang (d.h. für gute Gedächtnisleistungen) sinnvoll sein, nur solchen Items viel Lernaufwand zu widmen, die schwierig sind bzw. in einem ersten Lerndurchgang schlecht

beherrscht wurden. Dies wurde z.B. von Nelson, Dunlosky, Graf und Narens (1994) bei jüngeren Erwachsenen untersucht: In der ersten Bedingung wurden computergesteuert diejenigen Items zum Wiederlernen ausgewählt, die in einem ersten Lerndurchgang niedrige JOLs erhalten hatten; in der zweiten Bedingung wählte der Computer diejenigen Items aus, die vorher als gut gelernt eingeschätzt wurden; in der dritten Bedingung schließlich konnten die Teilnehmer selbst auswählen, welche Items sie noch einmal lernen wollten. Erwartungsgemäß waren die Testleistungen in der ersten und dritten Bedingung, d.h. unter guter (Selbst-) Regulation, am besten. Auch Thiede (1999) fand, dass nicht die Teilnehmer, die am meisten Items zum erneuten Lernen auswählten, die besten Leistungen aufwiesen, sondern diejenigen, die die beste Selbstregulation, hier operationalisiert als Korrelation zwischen JOLs und Lernzeit, aufwiesen. Zusammenfassend ist also eine strategische Einteilung des Lernaufwands für das Erreichen guter Testleistungen zielführend. Das reine Verlängern der Lernzeit führt dagegen häufig zu einem sog. *Labor-in-vain-effect*; d.h. auch sehr lang gelernte Inhalte werden dann nicht besser memoriert, da vermutlich in der zusätzlichen Lernzeit keine effektiveren Strategien zum Einsatz kommen (Nelson & Leonesio, 1988).

Die entsprechenden Befunde bei Grundschulkindern decken sich weitgehend mit denen bei Erwachsenen. Eine ältere Studie von Masur, McIntyre und Flavell (1973) zur Itemauswahl ergab, dass Erstklässler die Items zum erneuten Lernen eher zufällig wählten; die Drittklässler zeigten dagegen eher die Tendenz, nicht erinnerte Items zum wiederholten Einprägen auszusuchen, profitierten davon jedoch kaum in ihrer Gedächtnisleistung. Einfachere Designs zur Lernzeitallokation fanden jedoch sowohl bei Erst- als auch bei Drittklässlern eine verbesserte Erinnerungsleistung bei höheren Selbstregulationskennwerten (Lockl & Schneider, 2003). Auch Hinweise auf den Labor-in-vain-Effekt konnten schon bei Neunjährigen gefunden werden (Lockl & Schneider, 2004). In Bezug auf ältere Erwachsene zeigt die bereits erwähnte Studie von Dunlosky et al. (2003), dass eine durch das Training verbesserte Überwachungsleistung es den Teilnehmern erleichtert, noch unzureichend gelernte Items zum erneuten Lernen auszuwählen, und sich dadurch die Testleistung verbessert.

3.4.3 Fazit

Der vermutete Zusammenhang, dass Informationen aus Überwachungsprozessen genutzt werden, um Steuerungsprozesse anzupassen, was schließlich in einer besseren Testleistung resultiert, kann insgesamt für die verschiedenen Altersgruppen als bestätigt gelten. Das heißt, die Beziehung zwischen Überwachungs- und Gedächtnisleistung ist nicht direkt, sondern

erfolgt vermittelt über Kontrollvorgänge. Dieser Wirkzusammenhang entspricht also in etwa der Korrelation zwischen deklarativem Metagedächtniswissen und Erinnerungsleistung, die ebenfalls vermittelt, und zwar über die Strategienutzung, erfolgt (s. Abschnitt 2.1.3). Es muss jedoch festgehalten werden, dass der Bezug zwischen Überwachung, Kontrollprozessen und Gedächtnisleistung nicht für sehr lang gelernte Items gilt (Labor-in-vain-Effekt). Außerdem kann dieser Zusammenhang nur bei zielorientierten Lernprozessen gefunden werden (vgl. Abschnitt 3.3.3). Beim daten-orientierten Lernen ist die Erinnerungsleistung in der Regel für solche Items besser, die kurz gelernt werden und entsprechend hohe Überwachungsurteile erhalten (z.B. Hoffmann-Biencourt et al., 2010; Koriat, Ackerman et al., 2014).

4. EXKURS: EXPERTISE

Da sich zwei Experimente der vorliegenden Arbeit mit dem Einfluss von Expertenwissen auf die metakognitive Überwachungsleistung befassen (Studien 2 und 3) und das Thema Expertise zu umfassend ist, um es angemessen in den Kapiteln zu den Überwachungsmaßen abzubilden, soll es im folgenden Exkurs ausführlicher dargestellt werden. Zu diesem Zweck sollen kurz die Begrifflichkeiten geklärt, die Entwicklung von Expertise über die Lebensspanne sowie der Zusammenhang von Expertise mit Gedächtnis- und vor allem Metagedächtnisleistungen behandelt werden.

4.1 Begriffsklärung und Operationalisierung

In der einschlägigen Literatur wird der Begriff *Expertise* nicht einheitlich verwendet. So kann sich Expertentum zum einen auf das Wissen über ein bestimmtes Thema, zum anderen aber auch auf praktische Fertigkeiten (z.B. eine Sportart ausführen) beziehen (vgl. Chi, 2006; Gruber, 2010). Weiterhin wird Expertise einerseits als allgemeine Fähigkeit beschrieben, andererseits als eingegrenzt auf eine Domäne (Gruber, 2010). In den meisten Definitionen ist jedoch das Ergebnis des Expertenstatus enthalten, nämlich, dass die entsprechende Person (in ihrem Gebiet) dauerhaft hervorragende Leistungen erbringt (z.B. Posner, 1988). Im Gegensatz dazu handelt es sich bei einem *Novizen* um eine Person, die im Wortsinn „neu“ in der Domäne ist, und – zumindest bislang – keine überdurchschnittliche Leistungsfähigkeit und Erfahrung aufweist (Chi, 2006; Gruber, 2010). Häufig kommt – gerade in entwicklungspsychologischen Studien – das sog. *Experten-Novizen-Paradigma* zum Einsatz, d.h. die Leistungen von Experten und Novizen verschiedener Altersgruppen werden miteinander kontrastiert. Im Gegensatz zu Studien, bei denen die Entwicklung von Expertise retrospektiv erfasst wird, erlaubt dieses Vorgehen, Vorwissenseffekte von Effekten des chronologischen Alters zu trennen (z.B. Schneider, Gruber, Gold & Opwis, 1993).

In der vorliegenden Arbeit bezieht sich der Begriff des Experten rein auf wissensbezogene Komponenten in einer einzigen Domäne (Fußball). Diese Eingrenzung auf den Bereich Vorwissen wurde deshalb gewählt, da interessiert, wie Personen mit unterschiedlich ausgeprägtem deklarativen Faktenwissen sich auch hinsichtlich der Überwachungsleistungen in Gedächtnisaufgaben unterscheiden. Außerdem wird der Terminus Experte in dieser Arbeit sehr weit gefasst und vorrangig in seiner Abgrenzung zum Novizen verwendet; Fußballexperten sind demnach Personen, die im Vergleich zu Fußballnovizen in einem Test mehr

domänenspezifisches Vorwissen aufweisen. Diese Definition erlaubt es, recht alltagsnah den Einfluss des Vorwissens auf die Metagedächtnisleistung zu untersuchen.

4.2 Besonderheiten von Expertenwissen

Der Kernbestandteil von Expertisedefinitionen, nämlich das bereichsspezifische Vorwissen, welchem auch in der vorliegenden Arbeit die zentrale Rolle zukommt, weist einige Besonderheiten auf. Es wird angenommen, dass Expertenwissen durch eine lang andauernde Auseinandersetzung mit der Domäne erworben wird, demnach also trainierbar ist. Außerdem wird davon ausgegangen, dass dieses Wissen zu herausragenden Problemlösefertigkeiten, geringen Fehlerquoten und insgesamt hoher Effizienz beiträgt (Gruber, 2007). Vielfach wurde beobachtet, dass Experten im Vergleich zu Novizen nicht nur über eine deutlich größere Wissensbasis verfügen, sondern, dass diese sie auch zu besonderen Gedächtnisleistungen befähigt (z.B. Simon & Chase, 1973; Schneider et al., 1993; Schneider, Körkel & Weinert, 1989). Das semantische Netzwerk von Experten ist in der jeweiligen Domäne also insgesamt weiter ausgebaut, besser strukturiert und enger verschaltet als das von Novizen, sodass einzelne Informationen besser organisiert und leichter zugänglich sind (vgl. Schneider & Stumpf, 2007). Beim Abruf von Informationen gehen Experten auch insofern effizienter vor, als sie einzelne Elemente in größere Zusammenhänge (sog. *Chunks*) bringen können. Personen mit einer großen Wissensbasis verfügen demnach über weitaus mehr und größere Chunks als Personen mit einem geringen Vorwissen. Diese Verarbeitung in größeren Mustern verschafft Experten einen Leistungsvorteil gegenüber Novizen (Simon & Chase, 1973). Auch neue Informationen können demnach besser in das bestehende Netzwerk eingefügt werden (vgl. Schneider & Stumpf, 2007, sowie zusammenfassend die *Skilled memory theory* von Chase & Ericsson, 1981, 1982).

4.3 Entwicklung von Expertise

Im Verlauf der Expertiseforschung wurden verschiedene Annahmen formuliert, wie die Entstehung hervorragender Leistungen bzw. eines großen, domänenbezogenen Wissensschatzes im Verlauf der Entwicklung erklärt werden kann. Da diese wesentlich für das Verständnis von Expertise sind, sollen einige davon in Kürze vorgestellt werden, bevor auf die Entwicklung im Kindes-, Jugend- und höheren Erwachsenenalter eingegangen wird.

Modelle des Expertiseerwerbs. Eines der ältesten Modelle zur Ausbildung von Expertise stammt von Fitts und Posner (1967). Sowohl auf kognitive als auch auf handlungs-bezogene Fertigkeiten anwendbar, beschreibt es die zunehmende Automatisierung von Tätigkeiten. Zunächst, in der *kognitiven Stufe*, muss die zu erlernende Tätigkeit verstanden werden, es muss eine innere Repräsentation der Fertigkeit erstellt werden und Wichtiges von Unwichtigem während der Ausführung getrennt werden. Hier ist also deklaratives Wissen über die Fertigkeit entscheidend. In der zweiten, der *assoziativen Stufe*, laufen die kognitiven Prozesse bereits schneller ab; das deklarative wird in prozedurales Wissen überführt. In der *automatisierten Stufe* schließlich vollziehen sich die Prozesse, die für die Ausführung der Fertigkeit bzw. die Bewältigung der Aufgabe notwendig sind, weitgehend unbewusst und automatisch, sodass steuernde Eingriffe nur noch bei größeren Anpassungen an die Aufgabe oder Problemen notwendig sind.

Bezüglich des Ausmaßes der Erfahrung nahmen Ericsson, Krampe und Tesch-Römer (1993) an, dass in den meisten Domänen mindestens zehn Jahre intensiver Auseinandersetzung mit dem Bereich notwendig sind, um Spitzenleistungen zu erreichen (*10-Jahres-Regel*). Die Autoren wiesen jedoch darauf hin, dass in diesen zehn Jahren anstrengungsbetonte Übungs- und Lernprozesse stattfinden müssen (*Deliberate Practice-Modell*). Es geht also nicht nur um quantitative, sondern vor allem auch um qualitative Merkmale. Zum anstrengungsbetonten Üben gehören z.B. regelmäßige Rückmeldungen über den Leistungsstand, Wiederholungen, die Anwendung von Strategien und Übungsmethoden, herausragende Lehrer, Fehlerdiagnosen etc. Obwohl die Relevanz von qualitativ hochwertigen und zeitintensiven Übungsprozessen auch empirisch bestätigt werden konnte (z.B. Gruber, Degner & Lehmann, 2004), ist bei diesem Ansatz kritisch zu sehen, dass die Rolle von angeborenen Fähigkeitsunterschieden wie beispielsweise der intellektuellen Begabung nur eine sehr untergeordnete Rolle zukommt.

Angeborene Fähigkeitsunterschiede wirken sich im Sinne des sog. *Schwellenwertmodells* ebenfalls auf die Ausbildung hervorragender Fertigkeiten aus: Erst, wenn das Fähigkeitsniveau einer Person oberhalb eines Grenzwertes im überdurchschnittlichen Bereich liegt, bestimmen nicht kognitive Einflüsse wie anstrengungsbetontes Üben, ob insgesamt eine Höchstleistung erreicht wird (Schneider, 2000). Neben den angeborenen Fähigkeiten werden weiterhin auch das Alter, in dem mit der Auseinandersetzung mit dem Expertisebereich begonnen wird, oder Persönlichkeitseigenschaften als wichtige Einflussfaktoren auf die Ausbildung von Höchstleistungen diskutiert (vgl. Hambrick et al., 2014).

Entwicklung im Kindes- und Jugendalter. Da in den eigenen Studien die Wissensbasis und nicht praktische Fertigkeiten als Maß für die Expertise relevant ist, sollen im Folgenden sowohl Studien zu kognitiver Expertise als auch Studien zu (i.d.R. experimentell induziertem) Vorwissen vorgestellt werden. Hierbei geht es zunächst darum, wie Vorwissen bzw. Expertise mit Gedächtnisleistungen zusammenhängen, bevor weiter unten auf die Auswirkungen auf das Metagedächtnis eingegangen wird.

Schon bei Kindern kann beobachtet werden, dass Vorwissen die Erinnerung erleichtert und Einspeicher- und Abrufvorgänge verbessern kann (vgl. Schneider, 2015). So zeigte bereits Lindberg (1980), dass Drittklässler Collegestudenten in ihrer Erinnerungsleistung überlegen waren, wenn es sich bei den Stimuli um den Kindern bekannte Charaktere aus Cartoons oder Filmen handelte, jedoch nicht, wenn es sich um eine normierte Wortliste handelte. Die Kinder gingen beim Lernen der ersten Liste strategischer vor als die jüngeren Erwachsenen, indem sie die Items inhaltlich zusammenfassten. Ähnliche Effekte konnten ebenfalls für andere Domänen (z.B. Star Wars: Means & Voss, 1985; Baseball: Recht & Leslie, 1988) gefunden werden. Auch im Bereich Fußball, der für die vorliegende Arbeit von besonderer Bedeutung ist, konnte im Vergleich von Dritt-, Fünft- und Siebtklässlern, die als Fußballexperten bzw. -novizen klassifiziert wurden, nachgewiesen werden, dass die üblichen Alterseffekte in der Erinnerungsleistung durch die Vorwissenseffekte eliminiert wurden (Schneider et al., 1989). Die Autoren boten den Probanden einen leicht verständlichen Text über ein Fußballspiel dar und erfassten neben der Erinnerungsleistung und dem fußballbezogenen Vorwissen auch die intellektuellen Fähigkeiten sowie Maße der metakognitiven Überwachung. Bezüglich der Gedächtnisleistung zeigte sich, dass die Experten der dritten Klasse den Novizen der siebten Klasse überlegen waren. Weiterhin ergaben sich keine bedeutsamen Unterschiede zwischen Experten mit höheren und niedrigeren intellektuellen Fähigkeiten. Auf die Ergebnisse zum Metagedächtnis soll in Abschnitt 4.4 eingegangen werden.

Bei Kindern und Jugendlichen kann – ähnlich wie bei Erwachsenen – bereits beobachtet werden, dass sich Vorwissen nicht nur in quantitativer, sondern auch in qualitativer Hinsicht auf die Erinnerungsleistung auswirkt. So zeigte eine Replikation und Erweiterung der berühmten Schachstudie von Chi (1978), dass sowohl die Kinder als auch die Erwachsenen unter den Schachexperten die Figuren in einer ähnlichen Reihenfolge auswählten, wenn es darum ging, die Positionen zu rekonstruieren, was für die Nutzung vergleichbarer Chunks spricht (Schneider et al., 1993).

Insgesamt gesehen scheinen also bei Kindern und Jugendlichen grundsätzlich die gleichen Vorteile von gut strukturiertem und vernetztem Vorwissen auf die Gedächtnisleistung wie bei Erwachsenen zu gelten. Dennoch müssen besonders bei jüngeren Kindern bedeutsame Ausnahmen diskutiert werden: So existieren Befunde, dass diese Altersgruppe nicht immer von ihrem Vorwissen profitiert. Elischberger (2005) beispielsweise induzierte bei fünf- und sechsjährigen Kindern Vorwissen über eine Geschichte, die später präsentiert wurde. Zwar verbesserten diese Informationen erwartungsgemäß die Erinnerungsleistung, indem Speicher- und Abrufvorgänge erleichtert wurden. Das Vorwissen machte die Kinder jedoch auch anfälliger für den Einfluss von suggestiv formulierten Fragen. Suggestionen, die mit dem Vorwissen kompatibel waren, wurden häufiger bestätigt, inkompatible dagegen häufiger abgelehnt (vgl. auch Lipson, 1982).

Entwicklung im höheren Erwachsenenalter. Während viele kognitive Funktionen, wie z.B. die Verarbeitungsgeschwindigkeit oder die Fähigkeit, assoziative Verknüpfungen zu bilden, im Alter nachlassen, bleibt die Wissensbasis in der Regel intakt bzw. nimmt noch weiter zu (Hertzog, Kramer, Wilson & Lindenberger, 2008). Dieses in umgrenzten Bereichen erworbene Wissen zu nutzen und zu vermehren, wird als ein wesentlicher Aspekt erfolgreichen Alterns gesehen (vgl. Baltes & Baltes, 1990).

Die positiven Auswirkungen einer reichen Wissensbasis in einer bestimmten Domäne wurden für ältere Erwachsene vielfach belegt. So bestanden in einer Studie von Hultsch und Dixon (1983) keine Altersunterschiede in der Erinnerungsleistung zwischen jüngeren und älteren Probanden, wenn die älteren Erwachsenen ein höheres Vorwissen aufwiesen, d.h. wenn die dargebotenen Texte von Personen handelten, die vor einigen Jahrzehnten berühmt gewesen waren. In ähnlicher Weise erinnerten sich bei Toth, Daniels und Solinger (2011) die Senioren an mehr berühmte Schauspieler der 50er-Jahre, während die jüngeren Erwachsenen bei den Schauspielern der 90er-Jahre überlegen waren. Auch in noch alltagsnäheren Aufgaben konnte die Wirkung von Vorwissen auf die Gedächtnisleistung belegt werden: Beim Lesen von Kochrezepten wiesen diejenigen älteren Erwachsenen mit dem größeren Vorwissen die größere Leseeffizienz (operationalisiert durch Lesezeit/Erinnerungsleistung) auf; bei den jüngeren Erwachsenen ergaben sich diesbezüglich keine Unterschiede (Soederberg Miller, 2009). Auch hier wurden demnach altersbezogene Verschlechterungen durch den Einbezug von Vorwissen kompensiert. Auch erwartungsgemäße Preise für Lebensmittel erinnerten ältere Erwachsene genauso gut wie jüngere (Castel, 2005). Ebenso kann im Beruf erworbene Expertise

Alterseffekte eliminieren: Castel (2007) untersuchte das assoziative Defizit, also die Schwierigkeit älterer Erwachsener, neue semantische Verknüpfungen zu bilden (Naveh-Benjamin, 2000), anhand von willkürlich zugeordneten Zahlen und Objekten. Ältere Erwachsene, die als Buchhalter gearbeitet hatten, zeigten hier kein assoziatives Defizit; ihr gutes Gedächtnis für Zahlen schien also ausreichend Ressourcen für die Bildung von Assoziationen bereitzustellen.

Trotz dieser vielfältigen Belege für einen Ausgleich der kognitiven Leistung zwischen jüngeren und älteren Erwachsenen durch die Wissensbasis kann nicht davon ausgegangen werden, dass sich das Vorwissen bei älteren Erwachsenen in jedem Fall positiv auswirkt. So verbesserte bei einer komplexeren Gedächtnisaufgabe – dem Anhören eines fingierten Radioberichts eines Baseballspiels – das Vorwissen zwar insgesamt die Erinnerungsleistung; die Unterschiede zwischen den jüngeren und den älteren Probanden wurden dadurch jedoch nicht vollständig ausgeglichen (Hambrick & Engle, 2002). Weiterhin gibt es – vergleichbar mit den Ergebnissen bei Kindern (s.o.) – Situationen, in denen sich bei älteren Erwachsenen das Vorwissen negativ auf die Leistung auswirkt: So weisen ältere Erwachsene häufiger falsche Erinnerungen auf, wenn diese mit ihrem Vorwissen übereinstimmen. In der bereits erwähnten Studie von Castel (2005) waren ältere Erwachsene deutlich schlechter als jüngere in der Lage, sich nach oben oder unten gravierend abweichende Preise für Lebensmittel zu merken. Außerdem wirkte sich Vorwissen negativ aus, wenn Objekte einer Kategorie präsentiert wurden (z.B. Musikinstrumente) und in einer Liste zum Wiedererkennen dann Instrumente vorkamen, die in der ursprünglichen Liste nicht verwendet wurden, oder wenn es darum ging, sich falsche Multiplikationsaufgaben einzuprägen (für einen Überblick s. Umanath & Marsh, 2014). Das heißt, der kompensatorische Effekt von Vorwissen und bereichsspezifischer Expertise kommt bei älteren Erwachsenen nur dann zum Ausdruck, wenn die Aufgabe diesem nicht widerspricht.

4.4 Zusammenhang von Expertise und Metagedächtnisleistung

Während die Vorteile von Vorwissen und Expertise auf die Gedächtnisleistung in den meisten Situationen und Altersgruppen klar belegt sind, ist die Studienlage nicht völlig einheitlich, wenn es um die Auswirkungen auf die metakognitive Überwachungsleistung geht.

Auf der einen Seite gibt es Ergebnisse, die dafür sprechen, dass Personen von einem hohen Vorwissen auch bezüglich ihrer Überwachungsqualität profitieren. Bei jüngeren Erwachsenen wurde dies beispielsweise von de Bruin, Rikers und Schmidt (2007) für den Bereich Schach

gezeigt. Die Autoren fanden, dass die Schachexperten sowohl besser darin abschnitten, Spielzüge eines Endspiels korrekt vorherzusagen, als auch höhere JOL-Genauigkeiten erreichten, wenn sie gefragt wurden, wie sicher sie sind, einen ähnlichen Spielzug in Zukunft vorherzusagen. Bei den Schachnovizen waren die Gammakorrelationen für die JOL-Genauigkeiten nicht von Null verschieden. Auch für die Baseballdomäne wurde beim Lesen von Texten gezeigt, dass die Probanden mit höherem Vorwissen in der absoluten, hier jedoch nicht in der relativen Genauigkeit der Metacomprehension-Urteile, denen mit wenig Vorwissen überlegen waren (T. D. Griffin, Jee & Wiley, 2009). Insgesamt gesehen war in dieser Studie eine Tendenz zur Unterschätzung zu beobachten, die bei den Baseballexperten weniger stark ausgeprägt war. Für SUs fanden Nietfeld und Schraw (2002), dass diese beim Lösen von Wahrscheinlichkeitsaufgaben bei Collegestudenten mit höherem Statistikvorwissen genauer ausfielen als bei einer Vergleichsgruppe mit weniger Vorwissen. In allen drei Studien wurde das bessere Abschneiden der Experten im Vergleich zu den Novizen in den Überwachungsurteilen jeweils damit erklärt, dass diesen mehr kognitive Ressourcen zur Verfügung stehen, da die Verarbeitung der eigentlichen Aufgabe deutlich automatisierter abläuft (de Bruin et al., 2007; Nietfeld & Schraw, 2002). Weitere Gründe für die Überlegenheit von Experten sind zum einen die Tatsache, dass Überwachungsurteile von Experten auf einer breiteren Wissensbasis beruhen (T. D. Griffin et al., 2009) und dass sie insgesamt weniger Fehler begehen als Novizen (de Bruin et al., 2007).

Auch für die Altersgruppe der Kinder und Jugendlichen existieren Belege, dass Expertenwissen sich positiv auf die Überwachungsleistung auswirkt. In der bereits erwähnten Studie von Schneider et al. (1989) mit Dritt-, Fünft- und Siebtklässlern waren die Experten den Novizen auch hinsichtlich der metakognitiven Maße überlegen: Sowohl ihre globale Leistungsvorhersage als auch die FOK-Urteile waren in ersterer Gruppe genauer. Alterseffekte im Sinne einer besseren metakognitiven Leistung bei den älteren Schülern im Vergleich zu den jüngeren zeigten sich nur für die FOKs, nicht aber für die Leistungsvorhersagen.

Neben diesen Belegen gibt es jedoch ebenfalls Befunde, die zeigen, dass Vorwissen nicht in allen Fällen zu besseren metakognitiven Leistungen führt, sondern dass Experten bisweilen ihre eigenen Leistungen überschätzen. Bereits Oskamp (1965) fand bei klinischen Psychologen, dass deren Überschätzung in den SUs bei der Beurteilung von Fallvignetten umso größer war, je mehr Erfahrung sie vorweisen konnten. Auch bei Glenberg und Epstein (1987) zeigte sich ein vergleichbares Muster. Sie untersuchten die Überwachungsgenauigkeit von Studenten mit verschiedenen Vorwissensgraden in Musik oder Physik anhand von einschlägigen Texten. Hier zeigte sich ein inverser Zusammenhang von bereichsspezifischem Vorwissen und

metakognitiver Überwachungsleistung. Das heißt, je höher das Vorwissen war, desto geringer fiel die Genauigkeit aus. Beim Lösen von Physikaufgaben fanden Veenman und Elshout (1999), dass die Experten den Novizen nur bei den mittelschweren, nicht aber bei den schwierigen Aufgaben in ihrer metakognitiven Überwachung (hier gemessen durch Protokolle des lauten Denkens) überlegen waren. Eine Erklärung hierfür ist, dass „Wissen“ und „Sicherheit“ (*confidence*) assoziativ eng verknüpft sind und dieser Bezug besonders dann zum Tragen kommt, wenn Experten wissen, dass ihr Fachwissen gefragt ist (Son & Kornell, 2010). Weiterhin ist es denkbar, dass sich Experten bei Überwachungsurteilen eher an der Vertrautheit mit den Stimuli orientieren und weniger an der tatsächlichen Leistung (Glenberg & Epstein, 1987). Schließlich beobachteten Son und Kornell (2010), dass Experten seltener die Antwortoption „weiß nicht“ auswählten, was die Überwachungsleistung insgesamt beeinträchtigen kann. Eine Studie, die den Einfluss von Vorwissen auf die Überwachung bei älteren Erwachsenen untersucht, kam zum Ergebnis, dass sich in dieser Altersgruppe ebenfalls Tendenzen zur Überschätzung bei den vertrauten Stimuli zeigen (Toth et al., 2011). Während – wie oben bereits erwähnt – die jüngeren und älteren Probanden jeweils die Schauspieler ihrer Epoche (50er- vs. 90er-Jahre) besser erinnerten, ergab sich nur bei den jüngeren Teilnehmern ein Vorteil des Vorwissens in den JOL-Genauigkeiten. Die älteren Probanden überschätzten sich dagegen stärker bei den Schauspielern der 50er-Jahre. Die Autoren schlossen daraus, dass in dieser Altersgruppe die Neigung, sich auf die Vertrautheit mit den Stimuli zu stützen, besonders ausgeprägt ist.

4.5 Fazit

Insgesamt ist die vorliegende Literatur zum Zusammenhang zwischen Vorwissen und Überwachungsleistung zurzeit noch nicht ausreichend, um allgemein gültige Schlüsse über Entwicklungsprozesse zu treffen. Zwar ist die Korrelation zwischen einem hohen Vorwissen und einer guten Gedächtnisleistung sehr robust. Bezüglich metakognitiver Prozesse scheint es sich bei Expertenwissen bzw. einer soliden Wissensbasis um eine zweiseitige Angelegenheit zu handeln. Je nach Setting können diese sich also sowohl positiv als auch negativ auf die Überwachungsleistung auswirken. An dieser Stelle muss berücksichtigt werden, dass sich die dargestellten Studien stark in ihren Designs, den verwendeten Maßen, der Schwierigkeit der Aufgaben sowie darin, ob Expertise explizit erfragt wurde, unterscheiden. Außerdem existieren in diesem Bereich nur wenige Studien zu Entwicklungsverläufen und keine Befunde, bei denen eine größere Altersspanne verglichen wurde.

5. ZUSAMMENFASSUNG

Im theoretischen Teil dieser Arbeit wurde im ersten Kapitel zunächst eine Klärung der Begriffe Metagedächtnis und Metakognition vorgenommen. Dies geschah vor dem Hintergrund der historischen Entwicklung der Konzepte. Obwohl weiterhin Unschärfen in der Definition der Termini bestehen und sich das Konzept des Metagedächtnisses als sehr offen für Erweiterungen erweist, kann die Unterscheidung zwischen einer deklarativen Komponente, die verbalisierbares Wissen über Gedächtnisprozesse beinhaltet, und einer prozeduralen Komponente, die Überwachungs- und Steuerungsprozesse während eines spezifischen Lernvorgangs umfasst, als allgemein anerkannt gelten.

Auch die vorliegende Arbeit orientiert sich an dieser Differenzierung. Da der Schwerpunkt in den Studien auf den prozeduralen Aspekten liegt, wurde das deklarative Metagedächtnis im zweiten Kapitel bezüglich der Erfassungsmethoden, der Entwicklung über die Lebensspanne und des Zusammenhangs mit Gedächtnisleistungen nur kurz behandelt.

Im Zentrum des Theorieteils stand das dritte Kapitel mit Befunden zum prozeduralen Metagedächtnis. Die Gliederung orientierte sich am Modell von Nelson und Narens (1990, 1994), sodass zunächst Überwachungsprozesse ausführlicher vorgestellt wurden. Hinsichtlich der Erfassungsmethoden wurde zwischen Resolutions- und Kalibrierungsmaßen unterschieden; daraus wurde die Entscheidung, in den eigenen Studien die Resolution mittels der Überwachungsgenauigkeit (Gammakorrelationen) und der Differenzierungsfähigkeit zu analysieren, abgeleitet und begründet. Anschließend wurden die vier Überwachungsmaße des genannten Modells, nämlich EOLs, FOKs, JOLs und SUs sowohl aus allgemeinspsychologischer als auch aus entwicklungspsychologischer Sicht diskutiert. Dabei zeigte sich, dass an verschiedenen Stellen noch Forschungsbedarf besteht: So ist z.B. das Jugendalter bislang wenig in Bezug auf Überwachungsleistungen untersucht worden; weiterhin existiert nur eine einzige einschlägige Studie (Shing et al., 2009), die eine lebensspannenübergreifende Stichprobe einschließt. Gerade dies ist von Bedeutung, um Gemeinsamkeiten und Unterschiede in Entwicklungsprozessen herauszustellen. Außerdem sind sowohl Grundlagen für Überwachungsurteile als auch die Wirkung von verschiedenen Kontextfaktoren (z.B. Komplexität des Lernmaterials, Art der Erinnerungsabfrage) nicht in allen Altersgruppen gleichermaßen gut untersucht worden.

Als Beispiel für Kontrollvorgänge wurde die Lernzeitallokation herausgegriffen, da diese auch für die eigene Arbeit von Bedeutung ist, und vor dem Hintergrund der Entwicklung über die

Lebensspanne dargestellt. Ausführlich wurden weiterhin wechselseitige Wirkzusammenhänge zwischen Überwachungs- und Kontrollprozessen diskutiert. Dabei wurde verdeutlicht, dass die Nutzung der Informationen aus Überwachungsvorgängen für die Anpassung von Steuerungsvorgängen bereits für isolierte Altersgruppen recht gut untersucht ist und als bestätigt gelten kann. Die Forschung zum gegenteiligen Zusammenhang steht dagegen erst in den Anfängen; so gibt es für ältere Erwachsene bislang keinerlei diesbezügliche Befunde. Auch erste Ergebnisse zum simultanen Auftreten beider Wirkrichtungen bedürfen noch der Überprüfung an weiter gefassten Altersgruppen, um genauere Rückschlüsse auf Entwicklungsverläufe zuzulassen.

Um die Bedeutung von Überwachungs- und Kontrollvorgängen herauszustellen, wurden am Ende des dritten Kapitels Zusammenhänge mit Gedächtnisleistungen berichtet.

Da der Einfluss des Vorwissens auf das prozedurale Metagedächtnis einen weiteren Schwerpunkt dieser Arbeit bildet, wurden im vierten Kapitel eine Begriffsklärung vorgenommen und verschiedene Formen der Operationalisierung von Expertenwissen vorgestellt. Außerdem wurden Besonderheiten von Expertise sowie die Entwicklung über die Lebensspanne berichtet. Der letzte Abschnitt dieses Kapitels widmete sich dann Befunden zum Zusammenhang zwischen einer breiten Vorwissensbasis und Metagedächtnisleistungen. Auch hier wurden Forschungslücken deutlich: Zum einen sind die existierenden Ergebnisse eher widersprüchlich – es wurden sowohl Vor- als auch Nachteile von Experten gegenüber Novizen in Metagedächtnisleistungen gefunden. Zum anderen fehlen einschlägige entwicklungspsychologische Befunde.

Insgesamt wurde im theoretischen Teil herausgestellt, dass das prozedurale Metagedächtnis trotz seiner hohen Bedeutung für Lernprozesse in jedem Lebensalter nur sehr wenig über größere Altersabschnitte hinweg untersucht wurde. Diese Forschungslücke und damit zusammenhängende offene Fragen zur Wirkung von Kontextfaktoren und zur wechselseitigen Beeinflussung von Überwachungs- und Kontrollvorgängen sollen im empirischen Teil dieser Arbeit angegangen werden.

6. EMPIRISCHER TEIL

6.1 Einführung

Die vorliegende Arbeit verfolgt das Ziel, verschiedene Ansätze zur metakognitiven Überwachung und Steuerung unter besonderer Berücksichtigung der Lebensspannenperspektive zusammenzuführen. Denn bislang existieren kaum Studien, die Entwicklungsverläufe prozeduraler Metagedächtnisprozesse in einer breiten Altersspanne von der Grundschule bis zum höheren Erwachsenenalter untersuchen. Dabei widmen sich die einzelnen Experimente der Arbeit dem Einfluss verschiedener Kontextfaktoren sowie dem wechselseitigen Einfluss von Überwachungs- und Steuerungsprozessen.

Der empirische Teil dieser Arbeit stützt sich theoretisch auf das in Abschnitt 1.2.5 vorgestellte Modell von Nelson und Narens (1990, 1994). Dies ermöglicht eine einheitliche Benennung der einzelnen metakognitiven Maße und eine größtmögliche Vergleichbarkeit mit den Ergebnissen anderer Studien. Es werden jeweils Operationalisierungen von Überwachungs- bzw. Kontrollvorgängen aus dem Modell herausgegriffen und diese – unter Verwendung von einfacheren und komplexeren Lernmaterialien – an einer Stichprobe mit jeweils vier Altersgruppen (Grundschüler der dritten Klasse, Jugendliche der siebten und achten Klassen, jüngere Erwachsene, Senioren) untersucht.

Die Entscheidung für diese Altersgruppen ergab sich aus der Überlegung heraus, dass die Überwachungsfähigkeiten bereits ab der Mitte der Grundschule auf einem recht hohen Niveau, welches nur wenig unter dem jüngerer Erwachsener liegt, ausgebildet sind (vgl. Schneider, 2015 für einen Überblick). Aus diesem Grund sind insgesamt nur wenige Entwicklungsveränderungen bis zum Jugendalter zu erwarten, selbst wenn die Datenlage für diese Altersgruppe zum aktuellen Zeitpunkt noch als ungenügend bezeichnet werden muss (vgl. Paulus et al., 2014). Ein möglicher altersabhängig unterschiedlicher Einfluss der untersuchten Kontextfaktoren wie einem Strategie-Training oder dem Vorwissen sollte demnach nicht auf grundlegende Defizite bei den jüngeren Teilnehmern zurückzuführen sein. Der Einschluss der Gruppe der jüngeren Erwachsenen erfolgte unter dem Gesichtspunkt, dass diese als „Vergleichsstandard“ für die anderen Altersgruppen gelten kann. Denn in diesem Alter befindet sich die fluide kognitive Leistungsfähigkeit auf ihrem Höhepunkt (Craik & Bialystok, 2006) und auch die metakognitiven Überwachungsfähigkeiten verbessern sich später in der Regel nicht mehr (Schneider, 2015). Da insgesamt nur von wenigen Entwicklungsveränderungen in der Überwachungsleistung im mittleren Erwachsenenalter auszugehen ist (Weil et al., 2013),

erschien es angemessen und ökonomisch, beispielhaft für die Untersuchung von Alterungsprozessen im prozeduralen Metagedächtnis lediglich eine Gruppe von älteren Erwachsenen im Alter zwischen 60 und 80 Jahren einzuschließen (vgl. Shing et al., 2009). Auch die Steuerungskomponente als zweite Facette des prozeduralen Metagedächtnisses kann etwa ab der dritten Klasse bereits als recht gut ausgebildet gelten (z.B. Lockl & Schneider, 2003), selbst wenn sich hier noch größere Veränderungen bis ins Jugendalter zeigen (Schneider, 2015), sodass es auch für diesen Bereich als sinnvoll erschien, eine Versuchsgruppe im Jugendalter einzuschließen. Für das höhere Erwachsenenalter ist die Befundlage zu Kontrollprozessen wie z.B. der Lernzeitallokation bislang unklar. Diese Altersgruppe scheint nur unter günstigen Bedingungen in der Lage zu sein, ihre Zeit vergleichbar effizient auf verschiedene Lerninhalte aufzuteilen wie jüngere Erwachsene (Hines et al., 2009). Aus den genannten Gründen lag es sowohl für die Untersuchung von Einflussfaktoren auf die Qualität von Überwachungsprozessen als auch für die vom Zusammenspiel zwischen Überwachungs- und Steuerungsprozessen (Studien 5 und 6) nahe, diese vier Altersgruppen in einem querschnittlichen Design jeweils innerhalb einer Studie miteinander zu vergleichen, um einen Überblick über Entwicklungsverläufe im prozeduralen Metagedächtnis zu erhalten.

Da bisher weitgehend ungeklärt ist, wie sich verschiedene metakognitive Überwachungsprozesse in der Lebensspanne entwickeln, sollen im ersten Experiment die beiden bislang am meisten untersuchten Urteile JOLs und SUs herausgegriffen und sowohl anhand von Paar-Assoziationsaufgaben als auch anhand von komplexeren episodischen Gedächtnisaufgaben zum Lernen von filmischer Information untersucht werden. Dabei interessiert die Frage, inwiefern sich zum einen die recht hohe Stabilität der JOL-Genauigkeit in einzelnen Altersgruppen über eine größere Altersspanne hinweg replizieren lässt. Zum anderen soll erstmals eine einheitliche intentionale Instruktion auch beim Lernen von Filmmaterial verwendet werden. Dies ist besonders für die SUs von Bedeutung, da diese bei komplexeren Materialien bislang vor allem in Studien zum Augenzeugengedächtnis (also beim inzidentellen Lernen) untersucht wurden (z.B. Roebbers, von der Linden et al., 2007).

Einen größeren Schwerpunkt dieser Arbeit bildet die Frage nach dem Einfluss des bereichsspezifischen Vorwissens auf die Überwachungsleistung, die in den Experimenten 2 und 3 untersucht werden soll. Anknüpfend an Studie 1 sollen hier sowohl ein Paar-Assoziationsdesign (Experiment 2) als auch komplexere Materialien, hier in Form einer Fußballgeschichte (Experiment 3), zum Einsatz kommen. Damit soll zur Klärung der in Abschnitt 4.4 dargestellten divergierenden Befunde zur Wirkung von Vorwissen auf die Überwachungsleistung eine breite Datenbasis zur Verfügung gestellt werden – bestehend aus

den drei Überwachungsmaßen EOLs, JOLs und SUs, verschiedenen Lernmaterialien und einer breiten Altersspanne. Außerdem soll untersucht werden, inwiefern sich der Einfluss von Fußballvorwissen auf die Qualität der Überwachung im Laufe des Lebens ändert.

Das Experiment 4 befasst sich damit, wie sich Überwachungsleistungen verbessern lassen. Durch ein kurzes Training einer visuellen Elaborationsstrategie („visual imagery“) soll insbesondere bei Kindern und älteren Erwachsenen, aber auch bei Jugendlichen erreicht werden, dass diese mehr Details aus dem Lernkontext mit ihren Erinnerungen verknüpfen und somit bei der Abgabe der metakognitiven Urteile auf diagnostisch wertvollere Hinweisreize zurückgreifen. Weiterhin soll hier untersucht werden, inwiefern durch die Strategieinstruktion vermehrt „Recollection“-Prozesse (anstelle von Vertrautheitsprozessen) beim Abruf wirksam werden (vgl. Koriat, 1997; McCabe & Soderstrom, 2011).

Experiment 5 und 6 schließlich widmen sich – im Gegensatz zu den übrigen Studien – nicht ausschließlich den Überwachungsurteilen, sondern gehen der Frage nach, wie sich Überwachungs- und Kontrollvorgänge in den verschiedenen Altersstufen wechselseitig beeinflussen. Dazu soll das sog. MC-Modell (vgl. Abschnitt 3.3.1) altersübergreifend repliziert werden (Experiment 5). Gleichzeitig sollen im zweiten Durchgang dieses Experiments erste Belege für das CM-Modell bei älteren Erwachsenen gefunden werden. In Experiment 6 sollen dann Nachweise für das MC- und das CM-Modell innerhalb eines einzigen Durchgangs erbracht werden, indem die Hälfte der Wortpaare bei selbst eingeteilter Lernzeit mit hohen vs. niedrigen Anreizwerten versehen wird (vgl. Koriat, Ackerman et al., 2014). Bislang konnte ein flexibler Wechsel zwischen daten- und zielorientiertem Lernen nur bei Personen ab der neunten Klasse (Koriat, Ackerman et al., 2014) gezeigt werden. Auch hier fehlen Ergebnisse zu älteren Erwachsenen und insbesondere vergleichende Daten über die Lebensspanne.

Alle sechs Experimente wurden im Rahmen eines von der Deutschen Forschungsgemeinschaft geförderten Projektes zur Entwicklung der Metakognition in der Lebensspanne unter der Leitung von Prof. Wolfgang Schneider und Dr. Nicole von der Linden durchgeführt. Es handelte sich jeweils um Querschnittstudien. Insgesamt nahmen 816 Personen an den Experimenten teil, die zwischen September 2012 und Februar 2015 stattfanden. Die Stichprobe für die Pilotstudien umfasste insgesamt 257 Teilnehmer.

6.2 Studie 1: Einfluss der Komplexität der Lernmaterialien auf die Überwachungsleistung

6.2.1 Herleitung der Fragestellung

Wie gerade dargestellt ist die Zielsetzung der Studie 1 eher grundlegender Art: Es geht hierbei darum, erstmals sowohl JOLs als auch SUs in einer die Lebensspanne übergreifenden Stichprobe zu erfassen und dabei die Qualität dieser Urteile bei komplexen im Gegensatz zu einfachen Stimuli zu kontrastieren.

Bezüglich der Qualität der JOLs ist nicht von größeren Entwicklungsunterschieden auszugehen. Bei Kindern ab dem Grundschulalter legen die verfügbaren Studien nahe, dass hier die prospektive Überwachungsleistung bereits auf dem Niveau Erwachsener liegt (z.B. Roebbers, von der Linden, Schneider et al., 2007; Schneider et al., 2000). Es wurden zwar geringfügige Anstiege der JOL-Qualität bis zum Jugendalter nachgewiesen (Hoffmann-Biencourt et al., 2010; Koriat et al., 2009b); diese können jedoch auch darauf zurückzuführen sein, dass in diesen Studien die Aufgaben für Kinder und Jugendliche identisch waren, so dass Verbesserungen in den JOLs auch auf einer erleichterten Verarbeitung bei den älteren Teilnehmern beruhen könnten. Für das höhere Erwachsenenalter zeigen die vorhandenen Befunde recht deutlich, dass hier nicht von Verschlechterungen in der JOL-Qualität auszugehen ist (z.B. Connor et al., 1997; Hertzog et al., 2002). Weiterhin ist davon auszugehen, dass sich der Delayed-JOL-Effekt (s. Abschnitt 3.1.3.1) in den einbezogenen Altersgruppen gleichermaßen zeigt (z.B. Connor et al., 1997; Schneider et al., 2000). In Studie 1 – wie auch in den fünf anderen Experimenten – sollen jeweils verzögerte JOLs erfasst werden, da hier über alle Altersstufen hinweg bessere Überwachungsleistungen als bei unmittelbaren JOLs zu erwarten sind. Dies ist besonders für die Experimente relevant, in denen zusätzliche Effekte auf die JOL-Qualität (vgl. Studie 2 bis 4) untersucht werden sollen. Studien, die direkt den Einfluss von einfachen vs. komplexeren Materialien auf die Güte der JOLs erfassen, existieren bislang nicht. Grundsätzlich ist davon auszugehen, dass Aufgaben zum Paar-Assoziations-Lernen höhere Genauigkeiten bzw. bessere Differenzierungsleistungen hervorbringen als Aufgaben mit Text- oder Filmmaterial (vgl. Dunlosky & Lipko, 2007). Dies sollte für alle Altersgruppen gelten (vgl. von der Linden et al., 2011; Olin & Zelinski, 1997). Es muss an dieser Stelle jedoch berücksichtigt werden, dass gerade die Studien zum Lernen von komplexeren Materialien schwer miteinander und auch schwer mit Aufgaben zum Paar-Assoziationslernen zu vergleichen sind. Wenn z.B. filmische Stimuli im Rahmen von Augenzeugengedächtnis-Paradigmen eingesetzt werden, handelt es sich dabei meist um inzidentelles Lernen. Weiterhin

können häufig die Maße zur Erfassung der Leistungsvorhersagen bei komplexeren Materialien nicht direkt verglichen werden, da es sich hier z.T. um Verständnisurteile oder um globale Vorhersagen handelt und nicht um Fragen, die unmittelbar mit der Art der Erinnerungsabfrage korrespondieren (vgl. Baker et al., 2010; Miles & Stine-Morrow, 2004). In der folgenden Studie sind deshalb die JOL-Fragen zum Film genau wie die Erinnerungsabfrage formuliert, sodass die JOL-Leistung bei beiden Arten von Lernmaterialien vergleichbar ist. Um die Vergleichbarkeit weiter zu erhöhen, wird für beide Stimulitypen eine intentionale Lerninstruktion gewählt. Dies hat auch den Vorteil, dass der Lernvorgang im Experiment ähnlich wie in alltagsnahen Lern- und Erinnerungsprozessen stattfindet.

Auch für SUs liegen bisher – mit einer Ausnahme (Shing et al., 2009) – keine Studien vor, die die gesamte Lebensspanne umfassen. Hier ist davon auszugehen, dass sich etwas größere Alterseffekte als bei JOLs zeigen. Zwar lassen sich schon bei jüngeren Kindern recht genaue SUs beobachten (Visé & Schneider, 2000), jedoch sind Leistungsverbesserungen im Verlauf der Grundschulzeit wahrscheinlich (Pressley, Levin et al., 1987; Roebbers, 2002). Für das Jugendalter bestehen in Bezug auf SUs erhebliche Forschungslücken. Eine der wenigen Studien, die diese Altersgruppe mituntersucht, ist die Arbeit von Krebs und Roebbers (2012), die leichte Anstiege in der SU-Qualität bis zum Beginn der Sekundarstufe verzeichnet. Allerdings wurde auch hier keine Anpassung der Aufgabenschwierigkeit zwischen den Altersgruppen vorgenommen, sodass die Leistungssteigerungen nicht ausschließlich auf eine bessere retrospektive Überwachungsleistung zurückzuführen sind. Studien mit älteren Erwachsenen sind aufgrund der verschiedenen Enkodierbedingungen nur schwer vergleichbar. Bei Studien zum Paar-Assoziationslernen zeigen sich z.T. etwas niedrigere Gammakorrelationen bei älteren im Vergleich zu jüngeren Erwachsenen (z.B. Dodson et al., 2007). Da ältere Erwachsene jedoch besondere Schwierigkeiten haben, wenn suggestive Fragen oder bei Wiedererkennungsparadigmen Disktraktoren mit einem hohen Vertrautheitsgrad enthalten sind, sollte in der eigenen Studie der Leistungsabfall in den SUs bei älteren Erwachsenen eher gering ausfallen (vgl. Kelley & Sahakyan, 2003; Shing et al., 2009). Sowohl bei Kindern als auch bei älteren Erwachsenen sind schlechtere Überwachungsleistungen bei den SUs im Vergleich zu jüngeren Erwachsenen vor allem auf höhere Urteile nach falschen Antworten zurückzuführen (vgl. Roebbers, 2002; Shing et al., 2009). Dies sollte auch in der eigenen Studie bestätigt werden. Während bei Kindern SUs bislang vorrangig in Arbeiten zum Augenzeugengedächtnis (mit zumeist inzidenteller Instruktion) und nur sehr selten in Paar-Assoziationsaufgaben untersucht wurden, existieren bei älteren Erwachsenen bislang keine Studien, die komplexe Materialien umfassen. Aus diesem Grund ist auch für die SUs ein

Vergleich verschiedener Stimuli mit einheitlich intentionaler Lerninstruktion an einer Stichprobe, die eine große Altersspanne umfasst, sinnvoll.

Sowohl für die JOLs als auch – in etwas geringerem Maße – für die SUs wird davon ausgegangen, dass die Überwachungsleistungen über alle Altersgruppen hinweg beim Lernen von Paar-Assoziationen besser ausfallen als beim Lernen mit Filmmaterial. Dies ist zum einen auf die erleichterte Verarbeitung von einfacheren Stimuli zurückzuführen, die mehr kognitive Kapazität für den Überwachungsvorgang bereitstellt (de Bruin et al., 2007). Zum anderen sollten die Probeabrufe, die kurz vor der JOL-Abgabe stattfinden, bei den Bildpaaren eine höhere diagnostische Validität haben als bei den Fragen zum Film, da hier zu beurteilen ist, ob die Antwort erfolgreich abgerufen werden konnte oder nicht (vgl. Dunlosky & Metcalfe, 2009; Spellman & Bjork, 1992).

Konkret soll in Studie 1 also zunächst eine Aufgabe zum Lernen von semantisch assoziierten und nicht assoziierten Bildpaaren zum Einsatz kommen. Kurz nach dem Lernvorgang sowie nach der Erinnerungsabfrage sollen JOLs bzw. SUs erhoben werden. Durch die Anpassung der Aufgabenschwierigkeit zwischen den Altersgruppen durch verschieden lange Präsentationszeiten soll zudem die Qualität der Überwachungsurteile möglichst unabhängig von der Erinnerungsleistung erfasst werden. Zusätzlich soll auch für den Einsatz komplexer Materialien eine einheitliche Datengrundlage geschaffen werden. Erstmals werden JOLs zu Fragen zu einem Film bei älteren Probanden erfasst. Zudem ergänzt das Vorgehen mit einer intentionalen Lerninstruktion vor allem bei den beiden jüngeren Altersgruppen die bisherige Forschung zum Augenzeugengedächtnis. Es werden also JOLs bzw. SUs zu Fragen zu einem Film erhoben. Auch hier soll die Schwierigkeit der Fragen zum Film zwischen den Altersgruppen ausbalanciert werden.

6.2.2 Methodik

6.2.2.1 Design

In der Studie wurde ein 4 (Altersgruppe: Kinder, Jugendliche, jüngere und ältere Erwachsene) x 2 (Komplexität des Lernmaterials: Bildpaare vs. Film) faktorielles Design verwendet. Die Altersgruppe wurde als Zwischensubjektfaktor, die Komplexität des Lernmaterials als Innersubjektfaktor erfasst. Als abhängige Variablen wurden die Erinnerungsleistung sowie die Güte der JOLs und SUs erhoben.

6.2.2.2 Stichprobe

An der Studie nahmen insgesamt $N = 112$ Personen teil. Die Gruppe der Grundschüler umfasste 28 Drittklässler (10 männlich, 18 weiblich), die im Mittel 9.04 Jahre alt waren ($SD = 0.79$; Altersspanne: 8 bis 10). In der Gruppe der Jugendlichen waren 28 Schüler (5 männlich, 23 weiblich) mit einem Durchschnittsalter von 13.07 Jahren ($SD = 0.77$; Altersspanne: 12 bis 14), die je etwa zur Hälfte die siebte bzw. achte Klasse einer Hauptschule bzw. eines Gymnasiums besuchten. Weiterhin nahmen 28 jüngere Erwachsene teil (9 männlich, 19 weiblich), die im Mittel 20.25 Jahre alt waren ($SD = 1.88$; Altersspanne: 18 bis 25). Es handelte sich um Studierende sowie Berufstätige bzw. Auszubildende verschiedener Richtungen. In der Gruppe der älteren Erwachsenen befanden sich 28 Personen (10 männlich, 18 weiblich) im Alter von durchschnittlich 68.39 Jahren ($SD = 5.14$; Altersspanne: 58 bis 77).

Die Leistungen im Zahlennachsprechen (Normwerte des HAWIK- bzw. WIE-Untertests; Petermann & Petermann, 2010; von Aster, Neubauer & Horn, 2006) als Maß für das Kurzzeit- und Arbeitsgedächtnis unterschieden sich nicht signifikant zwischen den Altersgruppen (Grundschüler: $M = 10.19$; $SD = 2.32$; Jugendliche: $M = 10.04$; $SD = 2.70$; jüngere Erwachsene: $M = 9.79$; $SD = 3.48$; ältere Erwachsene: $M = 10.14$; $SD = 2.22$).

Gleiches galt für die Normwerte im Wortschatztest (Petermann & Petermann, 2010; von Aster et al., 2006) als Maß für die verbalen Fertigkeiten. Auch hier unterschieden sich die Mittelwerte der vier Altersgruppen nicht signifikant voneinander (Grundschüler: $M = 11.39$; $SD = 2.67$; Jugendliche: $M = 11.89$; $SD = 2.86$; jüngere Erwachsene: $M = 10.96$; $SD = 3.48$; ältere Erwachsene: $M = 11.07$; $SD = 3.42$).

Des Weiteren zeigten sich keine Unterschiede zwischen den jüngeren und den älteren Erwachsenen hinsichtlich der durchschnittlichen Schul- und Ausbildungszeit: Jüngere Erwachsene besuchten durchschnittlich 13.22 Jahre ($SD = 1.32$) die Schule, Berufsschule bzw. Universität, ältere Erwachsene im Schnitt 13.59 Jahre ($SD = 5.17$). Auch in der subjektiven Einschätzung des persönlichen Gesundheitszustands unterschieden sich die Werte in den beiden älteren Altersgruppen nicht. Jüngere Erwachsene erreichten hier einen Mittelwert von 1.70 ($SD = 0.61$) auf einer Skala von 1 („sehr guter Gesundheitszustand“) bis 5 („sehr schlechter Gesundheitszustand“), ältere Erwachsene von $M = 2.00$ ($SD = 0.54$). Lediglich hinsichtlich der Zahl der täglich eingenommenen Medikamente als weiterem Indikator zur gesundheitlichen Verfassung bei den erwachsenen Versuchsgruppen ergaben sich Unterschiede: Die jüngeren Erwachsenen berichteten, im Durchschnitt 0.29 Medikamente täglich einzunehmen ($SD =$

0.64), bei den älteren Erwachsenen lag dieser Wert bei $M = 1.64$ ($SD = 1.95$; $t(52) = 3.38$; $p < .01$).

6.2.2.3 Versuchsmaterial

Bildpaare. Beim Versuchsmaterial handelte es sich zum einen um 20 Bildpaare (s. Anhang A 1), von denen die Hälfte semantisch assoziiert (z.B. Block – Stift), die andere Hälfte unzusammenhängend war (z.B. Klavier – Banane). Die Bildpaare wurden aus dem Material bereits am Lehrstuhl für Psychologie IV durchgeführter Studien ausgewählt (z.B. Hoffmann-Biencourt et al., 2010). Zur Erfassung der JOLs wurde den Teilnehmern jeweils das linke Bild eines Paares präsentiert mit der Aufgabe vorherzusagen, wie sicher sie seien, in etwa einer halben Stunde das rechte Bild erinnern zu können. Die Urteile wurden auf einer Thermometerskala erfasst, die bereits in anderen Studien erfolgreich eingesetzt wurde (z.B. Koriat et al., 2009a; von der Linden et al., 2011), um vor allem den Drittklässlern den Zugang zu den JOLs zu erleichtern. Vergleichbar mit dem Spiel „Topf schlagen“ wurde den Probanden erklärt, dass „kalte“ Urteile am unteren Ende der Skala für solche Informationen stünden, die auch im Gedächtnis „weiter weg“, also unwahrscheinlicher zu erinnern seien. Bei „wärmeren“ Urteilen am oberen Ende der Skala handele es sich um Inhalte, die gut abgespeichert und abrufbar seien. Beim Erinnerungsabruf und den SUs wurde analog verfahren. Auch hier sollten die Teilnehmer bei Darbietung des linken Bildpaares das rechte erinnern. Anschließend sollte für jedes Paar die Sicherheit der gegebenen Antwort eingeschätzt werden. Da für die Erfassung der JOLs auf eine bestehende Programmierung in Java zurückgegriffen werden konnte, wurden diese auf einer durchgehenden Skala von 0 bis 100 erfasst (s. Abb. 4); die SUs wurden handschriftlich auf einer Skala von 1 („sehr unsicher“) bis 7 („sehr sicher“) erfasst. Das untere Ende der Skala wurde mit einem traurigen, das mittlere mit einem neutralen und das obere Ende mit einem fröhlichen Smiley markiert (s. Abb. 5).

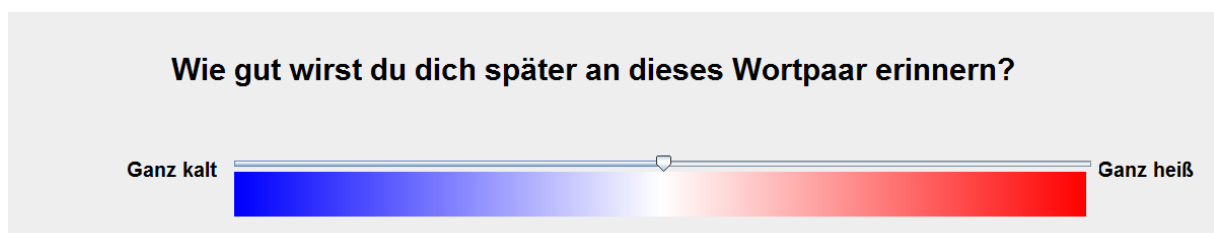



Abbildung 4: Thermometerskala zur Erfassung der JOLs

Wie sicher bist du dir, dass deine Antwort richtig ist?



Sehr unsicher Sehr sicher

Abbildung 5: Smiley-Skala zur Erfassung der SUs

Filmmaterial. Bei dem komplexeren Material handelte es sich um eine knapp sechsminütige Verfilmung einer kindgerechten Zaubershow. Auch dieser Film wurde bereits in einer Studie erfolgreich eingesetzt (Roebbers et al., 2004). Zu dem Film wurden jeder Altersgruppe zwölf offene Fragen gestellt (s. Anhang A 2). Bei den Gedächtnisfragen zum Film wurde ebenfalls zum Teil auf vorhandene Fragen (Roebbers et al., 2004) zurückgegriffen, die durch eigene ergänzt wurden, um vor allem bei den jüngeren Erwachsenen das Schwierigkeitsniveau zu steigern. In jeder Altersgruppe wurden aufgrund einer Pilotierungsstudie mit $N = 41$ Probanden jeweils ein Drittel leichte Fragen (mit einem Anteil richtiger Antworten von 70 bis 90%), ein Drittel mittelschwere Fragen (40 bis 60% richtige Antworten) und ein Drittel schwierige Fragen (10 bis 30% richtige Antworten; vgl. Marquié & Huet, 2000) dargeboten. Zur Erfassung der JOLs wurden die Gedächtnisfragen umformuliert, indem jeder Frage der Einschub „Wie sicher bist du/sind Sie, dass du dich/Sie sich in etwa einer halben Stunde daran erinnern kannst/können, dass...?“ vorangestellt wurde. Auch hier erfolgten die Einschätzungen auf einer Smiley-Skala von 1 („sehr unsicher“) bis 7 („sehr sicher“). Die SUs wurden erhoben, indem die Probanden nach der Beantwortung jeder Frage ebenfalls auf einer siebenstufigen Skala angeben sollten, wie sicher sie sind, dass diese Antwort korrekt ist.

Zusatzvariablen. Als zusätzliche Variablen wurden die Untertests Zahlennachsprechen und Wortschatz aus HAWIK bzw. WIE (Petermann & Petermann, 2010; von Aster et al., 2006) als Maße für das Arbeitsgedächtnis bzw. die verbale Intelligenz erhoben. Weiterhin erhielten die jüngeren und älteren Erwachsenen einen kurzen Fragebogen zu demografischen Angaben und

Gesundheitszustand (s. Anhang B 2), dessen Ergebnisse bei der Stichprobenbeschreibung auszugsweise vorgestellt wurden.

6.2.2.4 Studienablauf

Sämtliche Erhebungen fanden als Einzeltestungen – entweder in ruhigen Räumen der jeweiligen Schule oder in den Testräumen des Lehrstuhls für Psychologie IV – statt. Bei den minderjährigen Teilnehmern lag die schriftliche Einverständniserklärung der Erziehungsberechtigten vor. Bei etwa der Hälfte der Teilnehmer in jeder Altersgruppe wurde zunächst der Versuchsteil zum Lernen von Bildpaaren und anschließend der Teil zum Lernen aus filmischer Information durchgeführt; der anderen Hälfte der Teilnehmer wurde das Testmaterial in umgekehrter Reihenfolge vorgelegt.

Den Probanden wurden an einem Laptop die Bildpaare präsentiert. Den Teilnehmern wurde mitgeteilt, dass sie sich die Items so einprägen sollten, dass sie bei Darbietung des Stimulus das Target frei erinnern können. Vor der Lernphase wurden jedem Teilnehmer zwei Übungspaare als Beispiele präsentiert. Die Lernzeiten unterschieden sich je nach Altersgruppe, um die Lernleistung zwischen den Altersgruppen möglichst konstant zu halten. Auf Grundlage der oben erwähnten Pilotierungsstudie lag diese Lernzeit bei den Grundschulern und den älteren Erwachsenen bei acht Sekunden, bei den Jugendlichen bei sechs Sekunden sowie bei den jüngeren Erwachsenen bei drei Sekunden. Die Reihenfolge der Darbietung der einzelnen Bildpaare wurde zwischen den Teilnehmern variiert. Nach der Lernphase wurde die JOLs wie oben beschrieben erfasst. Anschließend wurde ebenfalls am Computer der Film dargeboten mit der Instruktion, ihn sich konzentriert anzusehen und sich die enthaltenen Informationen einzuprägen, da hinterher Fragen dazu gestellt würden. Es wurde dafür Sorge getragen, dass der Film für alle Teilnehmer akustisch und visuell gut wahrnehmbar war. Anschließend erfolgte die Erfassung der JOLs mit den nach Schwierigkeiten angepassten Fragen für jede Altersgruppe.

Im Intervall zwischen Lernen und Erinnern wurde der Test zum Zahlennachsprechen, der Wortschatztest sowie bei den jüngeren und älteren Erwachsenen der demografische Fragebogen bearbeitet. Anschließend wurde den Teilnehmern das linke Bild jeden Paares präsentiert mit dem Auftrag, das rechte zu ergänzen. Zu jeder Antwort wurde das SU erfasst. Die Teilnehmer hatten hierbei die Option mit „weiß nicht“ zu antworten. Die Bildpaare wurden in jeder Phase des Experiments in der gleichen Reihenfolge präsentiert. Anschließend erfolgten die

Erinnerungsabfrage sowie die Abgabe der SUs für den Film. Auch hier bestand die Möglichkeit, mit „weiß nicht“ zu antworten. Die Reihenfolge der Fragen war ebenfalls für die Lernphase, die Erinnerungsphase sowie die metakognitiven Urteile identisch (s. Anhang B 1).

Als Dank für die Teilnahme erhielten die minderjährigen Teilnehmer kleine Geschenke, die Psychologiestudenten Versuchspersonenstunden und die anderen erwachsenen Teilnehmer zehn Euro.

6.2.2.5 Überblick über die statistischen Auswertungen

Alle Berechnungen wurden mit SPSS 22[®] durchgeführt. Zur Bestimmung der Erinnerungsleistung wurde der Anteil korrekter Antworten bestimmt und in einer Varianzanalyse mit Messwiederholung mit dem Zwischensubjektfaktor Altersgruppe (vierstufig: Grundschüler, Jugendliche, jüngere und ältere Erwachsene) und dem Innersubjektfaktor Komplexität des Lernmaterials (zweistufig: Bildpaare und Film) inferenzstatistisch geprüft.

Zur Auswertung der Überwachungsmaße wurden jeweils zwei Methoden eingesetzt. Die Differenzierungsleistung wurde geprüft, um ein Maß für die absolute Höhe der Urteile zur Verfügung zu haben. Hierfür wurden jeweils die mittleren JOLs bzw. SUs bei richtigen sowie bei falschen Antworten getrennt berechnet. Der Faktor „Korrektheit der Antwort“ wurde dann als weiterer Innersubjektfaktor in die Messwiederholungs-ANOVA aufgenommen. Die abhängigen Variablen bildeten die Höhe der JOLs bzw. SUs. Um die JOLs bei den Bildpaaren unmittelbar mit denen bei den Fragen zum Film vergleichen zu können, wurde die Skala von 0 bis 100 in sieben Stufen umgerechnet (1 – 14: Stufe 1; 15 – 28: 2; 29 – 43: 3; 44 – 57: 4; 58 – 72: 5; 73 – 86: 6; 87 – 100: 7). Als weiteres Maß für die Güte der Überwachungsleistung wurden für jede Person Goodman-Kruskal Gammakorrelationen zwischen der Erinnerungsleistung und der Höhe der Überwachungsurteile berechnet und diese als abhängige Variablen in eine univariate ANOVA mit dem Zwischensubjektfaktor Altersgruppe und dem Innersubjektfaktor Komplexität des Lernmaterials eingesetzt. Da Gammakorrelationen das häufigste Maß für die Analyse der Genauigkeit von Überwachungsindikatoren darstellen, kann hierdurch die Vergleichbarkeit mit anderen Studien erhöht werden (vgl. Nelson, 1984). Auf die Berechnung von Kalibrierungsmaßen wurde aufgrund der geringen Stichprobengröße verzichtet.

Für die Bewertung der Ergebnisse wurde ein Signifikanzniveau von $\alpha = .05$ zugrunde gelegt. Post-hoc-Tests zur Bestimmung von Gruppenunterschieden wurden nach dem Verfahren von

Scheffé berechnet. Zur Einschätzung der Bedeutsamkeit der gefundenen Effekte wurde das partielle Eta² (η^2) herangezogen. Die Effektstärkemaße wurden nach Cohen (1988) interpretiert; eine Effektstärke von $\eta^2 = .02$ stellt dabei einen kleinen Effekt, von $\eta^2 = .13$ einen mittleren und von $\eta^2 = .26$ einen großen Effekt dar.

Vorgeschaltete Varianzanalysen mit den Zwischensubjektfaktor Geschlecht ergaben keine konsistenten Unterschiede zwischen männlichen und weiblichen Teilnehmern. Auch die Reihenfolge der Darbietung hatte keine substantziellen Effekte auf die Ergebnisse. Aus diesem Grund wurden diese beiden Faktoren in der Darstellung der Ergebnisse nicht berücksichtigt.

6.2.3 Ergebnisse

Die Ergebnisse der Studie 1 werden getrennt nach Erinnerungsleistung, JOLs und SUs berichtet. Für beide Materialarten wurden falsch beantwortete und nicht erinnerte Items jeweils zusammengefasst, da in jeder Altersgruppe nur sehr wenige falsche Antworten (insgesamt bei 5% der Bildpaare und bei 20% der Fragen zum Film) gegeben wurden, sodass diese alleine keine ausreichende Grundlage zur Berechnung von Gütemaßen der Überwachungsleistung darstellen.

6.2.3.1 Erinnerungsleistung

Für die Berechnung der Erinnerungsleistung wurde der Anteil korrekt erinnelter Bildpaare bzw. korrekt beantworteter Fragen zum Film herangezogen (s. Tab. 1). Eine ANOVA mit Zwischensubjektfaktor Altersgruppe und Innersubjektfaktor Komplexität des Lernmaterials ergab einen signifikanten Haupteffekt der Komplexität ($F(1,108) = 77.20; p < .001; \eta^2 = .42$) sowie eine signifikante Interaktion zwischen beiden Faktoren ($F(3,108) = 10.24; p < .001; \eta^2 = .22$). Nachgeschaltete Vergleiche zwischen den beiden Materialien mit *t*-Tests nach Altersgruppen getrennt ergaben, dass sich die Erinnerungsleistung nur bei den Grundschulern ($p < .05$) sowie bei den Jugendlichen und den jüngeren Erwachsenen (p 's $< .001$) zwischen Bildpaaren und Fragen zum Film bedeutsam unterschied, nicht aber bei den älteren Erwachsenen.

Tabelle 1: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe und Komplexität des Lernmaterials

	Bildpaare	Fragen zum Film
Grundschüler	73.27 (17.88)	63.99 (11.58)
Jugendliche	78.57 (11.70)	61.61 (14.23)
Jüngere Erwachsene	83.21 (17.33)	51.19 (15.17)
Ältere Erwachsene	71.61 (20.91)	65.77 (16.72)

Standardabweichungen in Klammern.

6.2.3.2 Judgments of Learning

Die prospektive Überwachungsleistung wurde – wie oben beschrieben – zunächst durch die Berechnung und inferenzstatistische Analyse der Differenzierungsleistung bestimmt. Die mittleren JOLs vor richtigen bzw. falschen Antworten sind in Tabelle 2 einzusehen. In der Messwiederholungs-ANOVA mit dem Zwischensubjektfaktor Altersgruppe und den Innersubjektfaktoren Komplexität des Lernmaterials und Korrektheit der Antwort zeigte sich ein signifikanter Haupteffekt des Faktors Altersgruppe ($F(3,93) = 8.51; p < .001; \eta^2 = .22$). Post-hoc-Tests ergaben, dass die Jugendlichen (4.10) etwas niedrigere JOLs abgaben als die Grundschüler (4.69; $p < .05$) und als die älteren Erwachsenen (5.02; $p < .001$). Weiterhin resultierten ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,93) = 829.90; p < .001; \eta^2 = .90$) und eine signifikante Wechselwirkung zwischen Korrektheit der Antwort und Komplexität des Lernmaterials ($F(1,93) = 10.73; p < .01; \eta^2 = .10$). Die Betrachtung der Interaktionsgrafik (s. Abb. 6) zeigt, dass über beide Lernmaterialien hinweg später korrekt erinnerte Inhalte deutlich höhere JOLs (6.02) erhielten als später vergessene/falsch erinnerte (3.11). Dieser Unterschied war bei den Bildpaaren etwas stärker ausgeprägt (6.12 vs. 2.92) als bei den Fragen zum Film (5.91 vs. 3.29).

Als zweite Auswertungsmethode für die Güte der JOLs kam die Berechnung der individuellen Gammakorrelationen zwischen JOL und Erinnerungsleistung für jeden Probanden (s.o.) zum Einsatz. Die Höhe der gemittelten Korrelationen als Maß für die Genauigkeit der Vorhersagen wurde ebenfalls zwischen den einzelnen Altersgruppen verglichen (s. Tab. 3). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Komplexität des Lernmaterials als Innersubjektfaktor ergab keine signifikanten Haupteffekte und keine Interaktionen. Mit t -Tests wurde nachgewiesen, dass die Korrelationen in allen Altersgruppen und Versuchsbedingungen

sehr hoch ausfielen, d.h. deutlich von Null verschieden waren (alle p 's $< .001$); die Probanden trafen also insgesamt sehr genaue Vorhersagen ihrer späteren Leistung.

Tabelle 2: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Komplexität des Lernmaterials

	Bildpaare		Fragen zum Film	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	6.25 (0.70)	3.14 (1.33)	5.93 (0.92)	3.43 (1.33)
Jugendliche	5.64 (0.93)	2.63 (1.31)	5.32 (0.62)	2.81 (1.04)
Jüngere Erwachsene	5.99 (1.25)	2.93 (1.40)	5.93 (0.71)	2.96 (1.14)
Ältere Erwachsene	6.64 (0.53)	2.99 (1.39)	6.48 (0.83)	3.97 (1.16)

Standardabweichungen in Klammern.

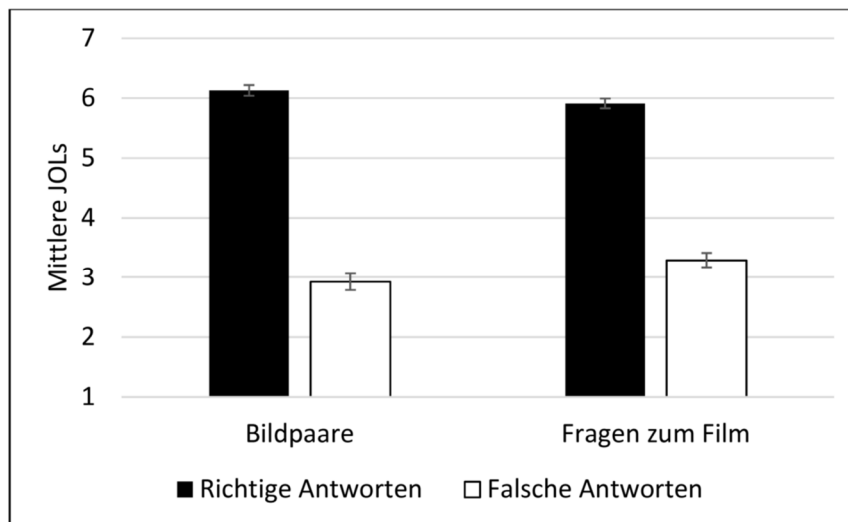


Abbildung 6. Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort und Komplexität des Lernmaterials.

6.2.3.3 Sicherheitsurteile

Die Sicherheitsurteile wurden identisch zu den JOLs ausgewertet (s. Tab. 4). Bezüglich der Differenzierungsleistung ergab die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Komplexität des Lernmaterials und Korrektheit der Antwort als Innersubjektfaktoren einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,84) = 3.98$; $p < .05$; $\eta^2 = .12$). Der

Post-hoc-Test zeigte, dass dieser Effekt durch höhere mittlere SUs bei den älteren Erwachsenen (5.01) im Vergleich zu den Jugendlichen (4.26) zustande kam ($p < .05$). Weiterhin ergab sich ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,84) = 511.60$; $p < .01$; $\eta^2 = .86$) sowie eine signifikante Interaktion zwischen der Korrektheit der Antwort und der Komplexität des Lernmaterials ($F(1,84) = 10.03$; $p < .01$; $\eta^2 = .12$). Hier zeigte sich das gleiche Muster wie bei den JOLs (s.o.): Auch hier waren die mittleren SUs nach richtigen Antworten (6.34) deutlich höher als die nach falschen (2.92). Die Teilnehmer differenzierten bei den Bildpaaren etwas besser zwischen richtigen und falschen Antworten (6.56 vs. 2.75) als bei den Fragen zum Film (6.13 vs. 3.10; s. Abb. 7).

Tabelle 3: Mittlere Gammakorrelationen für JOLs und SUs in Abhängigkeit von Altersgruppe und Komplexität des Lernmaterials

	JOLs		SUs	
	Bildpaare	Fragen zum Film	Bildpaare	Fragen zum Film
Grundschüler	.89 (.29)	.82 (.30)	.86 (.44)	.84 (.49)
Jugendliche	.86 (.39)	.86 (.19)	.76 (.66)	.87 (.17)
Jüngere Erwachsene	.89 (.20)	.84 (.20)	.67 (.67)	.92 (.10)
Ältere Erwachsene	.86 (.41)	.90 (.15)	.84 (.44)	.90 (.14)

Standardabweichungen in Klammern.

Tabelle 4: Mittlere SUs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Komplexität des Lernmaterials

	Bildpaare		Fragen zum Film	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	6.63 (0.47)	2.20 (1.69)	6.38 (0.82)	3.27 (1.73)
Jugendliche	6.53 (0.78)	2.27 (2.18)	5.71 (0.67)	2.54 (1.24)
Jüngere Erwachsene	6.69 (0.28)	3.04 (2.16)	6.07 (0.79)	2.75 (0.92)
Ältere Erwachsene	6.38 (1.30)	3.47 (1.91)	6.34 (0.77)	3.83 (1.20)

Standardabweichungen in Klammern.

Die varianzanalytische Auswertung der mittleren Gammakorrelationen zwischen SUs und Erinnerungsleistung (s. Tab. 3) ergab keine signifikanten Haupteffekte oder Interaktionen. Auch hier zeigten *t*-Tests, dass alle Korrelationen signifikant von Null verschieden waren ($p < .01$ bei jüngeren Erwachsenen bei den Bildpaaren; alle anderen p 's $< .001$).

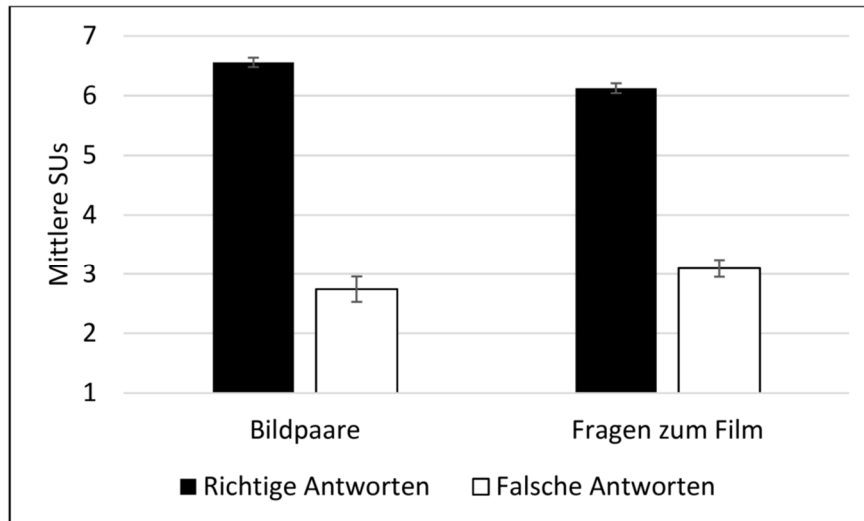


Abbildung 7. Mittlere SUs in Abhängigkeit von Korrektheit der Antwort und Komplexität des Lernmaterials

6.2.4 Diskussion

Die Studie verfolgte mehrere Ziele: Zum einen sollte die bestehende Forschung sowohl für JOLs als auch für SUs auf eine lebensspannenübergreifende Stichprobe übertragen werden. Zum anderen sollten erstmals komplexere Materialien auch an JOLs untersucht werden und ein direkter Vergleich zwischen komplexeren, aber alltagsnäheren, und vergleichsweise einfach strukturierten Paar-Assoziationsaufgaben unternommen werden.

Um die metakognitive Leistung möglichst unabhängig von der altersbedingt unterschiedlichen Lern- und Erinnerungsfähigkeit zu erfassen, wurden die Lernzeiten bei den Bildpaaren bzw. die Schwierigkeiten der Fragen zum Film zwischen den Altersgruppen angepasst. Weiterhin wurde angestrebt, die grundlegende Aufgabenschwierigkeit zwischen den Bildpaaren und dem Film anzugleichen. Die Ergebnisse bezüglich der Erinnerungsleistung zeigen, dass dies weitgehend gelang. Die Altersgruppen unterschieden sich nicht wesentlich in ihren Lernergebnissen. Allerdings lag die durchschnittliche Leistung für die Bildpaare insgesamt höher als bei den Fragen zum Film (16% mehr richtige Antworten bei den Bildpaaren). Dieser Trend zeigte sich

bei allen Altersgruppen mit Ausnahme der älteren Erwachsenen. Da die Erinnerungsleistung jedoch insgesamt im mittleren Bereich liegt, kann davon ausgegangen werden, dass die Anpassung der Aufgabenschwierigkeiten zufriedenstellend ist.

Die Ergebnisse zu JOLs und SUs zeigen insgesamt ähnliche Tendenzen. So waren die Probanden aller Altersgruppen sowohl in der Leistungsvorhersage als auch in der retrospektiven Einschätzung der Korrektheit der Antworten in der Lage, zwischen richtigen und falschen Antworten deutlich zu unterscheiden. Auch die Gammakorrelationen weisen auf gute bis sehr gute Überwachungsleistungen hin. Es zeigten sich nur geringe Alterseffekte in dem Sinne, dass die jugendlichen Teilnehmer – unabhängig von der Richtigkeit der Antwort – etwas niedrigere JOLs als die Grundschüler und die älteren Erwachsenen bzw. etwas niedrigere SUs als die älteren Teilnehmer abgaben. Abgesehen davon ergaben sich keine Einflüsse der Altersgruppe auf die Höhe bzw. die Güte der Überwachungsurteile. Einen weiteren wichtigen Aspekt dieser Studie stellte die Frage nach unterschiedlichen Auswirkungen der Komplexität der Stimuli auf die Überwachungsleistung dar. Hier zeigten sich lediglich für die Differenzierung bedeutsame Effekte: In allen Altersgruppe gelang die Unterscheidung zwischen richtigen und falschen Antworten sowohl für JOLs als auch für SUs bei den Bildpaaren besser als bei den Fragen zum Film. In der Genauigkeit der Überwachung (erfasst durch die Gammakorrelationen) zeigten sich dagegen keine Unterschiede zwischen den beiden Materialien.

Es ist also mit der vorliegenden Studie gelungen, bisherige Befunde erfolgreich auf ein lebensspannenübergreifendes Paradigma auszuweiten. Für die JOLs zeigte sich in Übereinstimmung zu den Ergebnissen anderer Studien, dass bereits Grundschüler in der Lage sind, ihre zukünftige Leistung bei verzögerter Abfrage recht genau vorherzusagen (vgl. Roebers, von der Linden, Schneider & Howie, 2007; Schneider et al., 2000). Verbesserungen in der JOL-Qualität bei Jugendlichen im Vergleich zu Grundschulern, die z.T. in anderen Studien gefunden worden waren (z.B. Hoffmann-Biencourt et al., 2010), konnten in dieser Studie nicht repliziert werden. Es muss jedoch darauf hingewiesen werden, dass durch die altersangepasste Modifikation der Aufgabenschwierigkeit in den vier Altersgruppen Unterschiede in den Überwachungsleistungen, die auf Unterschiede in der kognitiven Beanspruchung zwischen den Altersgruppen bei der Bearbeitung der Aufgaben zurückzuführen sind, hier weitgehend ausgeschlossen werden konnten. Auch in Bezug auf die Entwicklung im höheren Erwachsenenalter konnte analog zu Vorstudien nachgewiesen werden, dass nicht von einer Verschlechterung der JOL-Qualität auszugehen ist (vgl. Connor et al., 1997; Hertzog et al., 2002).

Für die SUs wurden die erwarteten Alterseffekte im Sinne eines leichten Anstiegs vom Grundschulalter bis zum Jugendalter bzw. eines leichten Abfalls vom jüngeren bis zum höheren Erwachsenenalter nicht gefunden. In der Analyse der Differenzierungsleistung verfehlte die Interaktion zwischen Korrektheit der Antwort und Altersgruppe knapp das erforderliche Signifikanzniveau ($F(3,84) = 2.61$; $p = .057$; $\eta^2 = .09$). Dieses Ergebnis deutet darauf hin, dass die älteren Erwachsenen, vor allem durch etwas höhere SUs nach falschen Antworten, tendenziell etwas schlechtere Differenzierungsleistungen aufweisen als die übrigen drei Altersgruppen. Das heißt, zu hohe Sicherheiten nach „falschen Alarmen“ im höheren Erwachsenenalter, wie sie bei Shing et al. (2009) gefunden wurden, zeigten sich in der hier dargestellten Studie nur ansatzweise. Es ist allerdings zu beachten, dass die Aufgaben in der Studie von Shing und Kollegen (2009) für die Teilnehmer deutlich anspruchsvoller waren (Lernen von deutsch-malaysischen Wortpaaren) und die Erinnerungsraten deutlich niedriger lagen als in der eigenen Studie, so dass die Vergleichbarkeit eingeschränkt ist. Fehlende Entwicklungseffekte vom Kindes- zum Jugendalter können auch hier darauf zurückzuführen sein, dass die Aufgabenschwierigkeiten im Gegensatz zu anderen Studien in diesen Altersgruppen angepasst wurden (vgl. Krebs & Roebers, 2012; Roebers, 2002).

Hinsichtlich der Komplexität des Materials sind die Befunde zu JOLs und SUs vergleichbar. In beiden Fällen war die Differenzierungsfähigkeit etwas besser für die Einschätzung der Bildpaare als für die Fragen zum Film. Dies ist dadurch zu erklären, dass die Entscheidung, ob ein aktuell oder zukünftig abrufbarer Gedächtnisinhalt richtig oder falsch ist, bei einfach strukturierten Materialien deutlich einfacher zu treffen ist (vgl. Begg et al., 1989). Bei Fragen zu komplexeren Materialien ergibt sich eine zusätzliche Unsicherheit darüber, ob die vorhandene Erinnerung vollständig ist oder nicht; das heißt, sogenannte Probeabrufe, die vor allem vor der Abgabe von JOLs relevant sind (vgl. Dunlosky & Metcalfe, 2009; Spellman & Bjork, 1992), sind hier weniger zuverlässig als bei Bildpaaren. Es wird hier also unter Umständen mit höherer Wahrscheinlichkeit fälschlicherweise davon ausgegangen, dass nicht korrekt erinnerte Inhalte zutreffend sind. Da in der vorliegenden Studie im Gegensatz zu einigen Arbeiten zur metakognitiven Überwachung beim Textlernen (z.B. Baker et al., 2010; Glenberg & Epstein, 1985) ausschließlich Fragen, die eine kurze, eindeutige Antwort verlangen (z.B. „Welche Farbe hatte der Hase?“), zum Einsatz kamen, ist hier die Unsicherheit in der Überwachung nicht so groß wie in anderen Studien zu komplexeren Materialien. Besonders im Vergleich zur allgemeinen Differenzierungsleistung fiel der Unterschied zwischen den beiden Materialformen nur gering aus; aus diesem Grund schlug sich dieser Effekt auch nicht in den Gammakorrelationen nieder, die insgesamt hoch bis sehr hoch ausfielen. Insgesamt gesehen

scheinen die Ergebnisse aus Paar-Assoziations-Aufgaben und Aufgaben mit komplexeren Materialien wie Filmen oder Texten ähnliche zugrunde liegende Prozesse zu erfassen. Diese Schlussfolgerung kann aber zunächst nur für Designs gelten, bei denen eine intentionale Lerninstruktion vorliegt, die Art der JOL- bzw. SU-Erhebung sehr nah an der Formulierung der Erinnerungsabfrage ist und keine suggestiven Fragen enthalten sind.

Etwas kritisch ist in der vorliegenden Studie die Tatsache zu sehen, dass den Probanden die Möglichkeit gegeben wurde, mit „weiß nicht“ zu antworten. Dies erhöht zwar die Überwachungsleistung, da bereits Grundschüler in der Lage sind, mit recht großer Sicherheit zu entscheiden, welche Antworten zurückgehalten werden sollten (Roebbers, von der Linden & Howie, 2007). Es war jedoch nicht möglich, die Überwachungsurteile bei falschen Antworten einzeln auszuwerten, da diese zu selten auftraten; weiterhin lagen die SUs bei den „weiß nicht“-Antworten grundsätzlich am unteren Ende der Skala, was ebenfalls die statistische Analyse durch die fehlende Varianz erschwerte, sodass für die folgenden Studien entschieden wurde, ein „forced-report“-Paradigma anzuwenden, bei dem jede Frage beantwortet werden muss, ggf. durch Raten (Roebbers, von der Linden & Howie, 2007).

Zusammenfassend lässt sich festhalten, dass die Umsetzung eines lebensspannenübergreifenden Designs gelungen ist. Es konnten Aufgaben von vergleichbarer Schwierigkeit identifiziert und Ergebnisse aus Vorstudien, die einzelne isolierte Altersgruppen vergleichen, repliziert werden. Grundschüler, Jugendliche, jüngere und ältere Erwachsene sind also in der Lage, sowohl pro- als auch retrospektiv zwischen erinnerten und nicht erinnerten Inhalten zu unterscheiden und ihre eigene Gedächtnisleistung adäquat einzuschätzen. Weiterhin konnte gezeigt werden, dass die zugrunde liegenden Prozesse bei Überwachungsvorgängen prinzipiell zwischen einfacher strukturiertem und komplexerem Material vergleichbar zu sein scheinen, so dass Paar-Assoziationsaufgaben einen guten Indikator für allgemeinere metakognitive Leistungen darstellen. Es kann bei den weiteren Studien vorrangig auf einfaches Lernmaterial zurückgegriffen werden, in der Annahme, dass die Ergebnisse sich auf alltagsnähere Fragestellungen verallgemeinern lassen.

Nachdem in dieser ersten Studie ein grundsätzlicher Nachweis dafür erbracht wurde, dass der Vergleich metakognitiver Leistungen über eine breite Altersspanne und bei Materialien verschiedener Komplexität möglich ist, soll in der nächsten Studie ein möglicher Einflussfaktor auf die Überwachungsleistung, nämlich das Vorwissen, in den vier Altersgruppen untersucht werden.

6.3 Studie 2: Einfluss des Vorwissens auf die Überwachungsleistung beim Lernen von Paar-Assoziationen

6.3.1 Herleitung der Fragestellung

Wie die Ergebnisse der Studie 1 zeigen, sind die metakognitiven Überwachungsleistungen bereits bei Grundschulern adäquat ausgeprägt; es ist weiterhin auch nicht von grundlegenden Defiziten im höheren Erwachsenenalter auszugehen. Dennoch ist anzunehmen, dass sich bestimmte Einflussfaktoren auf die Qualität von Überwachungsurteilen in verschiedenen Lebensaltern unterschiedlich auswirken. Ein möglicher Faktor soll in den folgenden beiden Experimenten untersucht werden: Es handelt sich dabei um den Einfluss von Expertenwissen. Wie im Theorieteil dieser Arbeit (s. Abschnitt 4.1) in größerer Ausführlichkeit dargestellt, ist die gängige Untersuchungsmethode für die Auswirkungen von Vorwissensunterschieden das sog. Experten-Novizen-Paradigma, welches auch in der folgenden Studie zum Einsatz kommen soll.

Während die Effekte von Expertenwissen auf viele kognitive Domänen wie beispielsweise Gedächtnisleistungen, die Anwendung von Strategien oder die Leichtigkeit der Verarbeitung in allen Altersgruppen als gut belegt gelten können (z.B. Castel, 2007; Chi, 2006; Schneider et al., 1989; Schneider & Stumpf, 2007), ist die Datenlage für den Bereich der metakognitiven Überwachung weniger eindeutig. Zum einen existieren insgesamt wenig einschlägige Studien. Zum anderen fehlt es im Bereich der Entwicklungspsychologie an Befunden, die eine größere Altersspanne umfassen. Schließlich zeigen die vorhandenen Ergebnisse keine eindeutigen Tendenzen, wie sich Expertise auf die metakognitiven Überwachungsleistungen auswirkt. Für jüngere Erwachsene konnte gezeigt werden, dass Experten im Vergleich zu Novizen höhere JOL- und SU- Genauigkeiten aufwiesen als Novizen (de Bruin et al., 2007; Nietfeld & Schraw, 2002). In einer Studie von T. D. Griffin et al. (2009) traf dies jedoch nur für die absolute Genauigkeit, nicht aber für Resolutionsmaße zu. Auch für Kinder und Jugendliche gibt es Belege, dass in diesen Altersgruppen Experten im Vergleich zu Novizen genauere globale Leistungsvorhersagen sowie genauere FOK-Urteile treffen (Schneider et al., 1989). Besonders in Studien, die Kalibrierungsmaße untersuchen, zeigt sich jedoch häufig auch der Effekt, dass Personen mit einem ausgeprägten Vorwissen eher dazu neigen, ihre eigenen Leistungen zu überschätzen (z.B. Glenberg & Epstein, 1987; Son & Kornell, 2010). Auch bei älteren Erwachsenen zeigte sich in der Studie von Toth et al. (2011) bei der Anwendung von Resolutionsmaßen, dass sich diese Altersgruppe in ihrem Vorwissensbereich im Vergleich zu jüngeren Erwachsenen eher überschätzte.

Da die erwähnten Studien aufgrund ihrer Vielzahl an Domänen (Fußball, Schach, Basketball, Statistik etc.), an Stimuli (Texte, Problemlöseaufgaben etc.), an verwendeten Maßen (globale Leistungsvorhersagen, JOLs, FOKs, SUs, Protokolle lauten Denkens), an Auswertungsmethoden (Resolution, Kalibrierung) sowie an Altersgruppen nur schwer vergleichbar sind, können noch keine allgemeingültigen Schlüsse gezogen werden.

Deshalb sollen in den folgenden beiden Experimenten sowohl einfach strukturierte (Wortpaare) als auch komplexere, alltagsnähere Materialien (Texte) zum Einsatz kommen. Weiterhin soll eine Vielzahl von Überwachungsmaßen (globale Leistungsvorhersage, EOLs, JOLs und SUs) verglichen werden. Die untersuchte Altersspanne schließt wieder Grundschüler der dritten Klasse, Jugendliche der siebten und achten Klassen sowie jüngere und ältere Erwachsene ein. Bezüglich der Domäne fiel die Entscheidung auf das Thema „Fußball“. Die Entscheidung für diesen Bereich ergab sich aus der Überlegung, in diesem Feld sowohl bei Grundschulern als auch bei Senioren Personen identifizieren zu können, die hier ein großes Vorwissen aufweisen.

Im ersten Experiment wurde der Fokus auf grundlegende Überwachungsprozesse gelegt; das Lernmaterial bestand aus fußballbezogenen Wortpaaren. Damit soll auch die Vergleichbarkeit zu anderen Studien mit Paar-Assoziationsaufgaben gewährleistet werden. In den Erinnerungsleistungen sollten die Experten den Novizen in allen vier Altersgruppen deutlich überlegen sein (vgl. Nietfeld & Schraw, 2002; Schneider et al., 1989). Da es den Probanden explizit bekannt war, dass es um den Vergleich zwischen Fußballexperten und -novizen ging, ist davon auszugehen, dass die generelle Tendenz von Experten, ihre Leistungen zu optimistisch einzuschätzen, sich auch in dieser Studie in den Differenzierungsleistungen zeigt (vgl. Son & Kornell, 2010). Es wird erwartet, dass dieser Befund bei den älteren Erwachsenen am stärksten ausgeprägt ist, da hier bei Personen mit einem hohem bereichsspezifischen Vorwissen die Neigung, sich bei der Beurteilung eigener Leistungen auf die Vertrautheit mit den Reizen zu stützen, noch deutlicher zur Geltung kommen sollte (vgl. Shing et al., 2009; Toth et al., 2011).

Im Gegensatz zur ersten Studie wurden hier zusätzlich EOLs erhoben, da davon ausgegangen wird, dass sich in einer Einschätzung, die noch vor dem eigentlichen Lernvorgang stattfindet, die stärksten Unterschiede in der Überschätzung zwischen Experten und Novizen zeigen, da hier vor allem Faktoren wie die Vertrautheit mit dem Material, aber auch Vorannahmen über den eigenen Expertenstatus am stärksten einfließen sollten (Toth et al., 2011; Zhao & Linderholm, 2008). Bezüglich des Vergleichs zwischen den Überwachungsmaßen wird erwartet, dass diese umso genauer werden, je mehr Informationen aus Lern- und Erinnerungsprozessen einfließen. Folglich sollten die SUs die höchste Genauigkeit und die

besten Differenzierungsleistungen hervorbringen, gefolgt von den JOLs und schließlich den EOLs (vgl. Leonesio & Nelson, 1990).

Weiterhin wird aufgrund der vorhandenen Literatur erwartet, dass die Experten den Novizen – trotz eines insgesamt wahrscheinlich zu hohen Optimismus bezüglich der eigenen Leistungen – in ihrer metakognitiven Überwachung überlegen sind, da diesen mehr kognitive Kapazitäten für die Überwachung zur Verfügung stehen und sie insgesamt weniger Fehler machen sollten als die Novizen (z.B. de Bruin et al., 2007; Schneider et al., 1989). Dieser Effekt sollte bei den Experten unter den älteren Erwachsenen jedoch schwächer ausgeprägt sein, und zwar aus dem oben genannten Grund, dass es dieser Altersgruppe schwerer fallen sollte, adäquat zwischen neu gelerntem und vertrautem Material zu unterscheiden.

6.3.2 Methodik

6.3.2.1 Design

In der Studie wurde ein 4 (Altersgruppe: Kinder, Jugendliche, jüngere und ältere Erwachsene) x 2 (Expertise: Experten vs. Novizen) faktorielles Design verwendet. Die Altersgruppe und die Expertise wurden als Zwischensubjektfaktoren erfasst. Als abhängige Variablen wurden die Erinnerungsleistung sowie die Güte der EOLs, JOLs und SUs erhoben.

6.3.2.2 Stichprobe

An der Studie nahmen insgesamt $N = 160$ Personen teil. Darunter befanden sich 40 Drittklässler (30 männlich, 10 weiblich; mittleres Alter: 8.33 Jahre; $SD = 0.53$; Altersspanne: 7 bis 9 Jahre), 40 Jugendliche (19 männlich, 20 weiblich; mittleres Alter: 12.90 Jahre; $SD = 0.74$; Altersspanne: 12 bis 14), 40 jüngere Erwachsene (21 männlich, 19 weiblich; mittleres Alter: 23.53 Jahre; $SD = 3.68$; Altersspanne: 18 bis 32) sowie 40 ältere Erwachsene (23 männlich, 17 weiblich; mittleres Alter: 66.83 Jahre; $SD = 5.38$; Altersspanne: 59 bis 80). Die Teilnehmer stammten aus der gleichen Population wie in der ersten Studie.

Die Leistungen im Zahlennachsprechen (Petermann & Petermann, 2010; von Aster et al., 2006) unterschieden sich nicht zwischen den Experten und den Novizen. Lediglich für den Faktor Altersgruppe ergab sich in einer ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren ein signifikanter Haupteffekt ($F(3,151) = 2.79$; $p < .05$; $\eta^2 = .05$), der sich in Post-hoc-Analysen jedoch nicht mehr in Gruppenunterschieden abbildete. Die

gemittelten Normwerte lauteten wie folgt: Grundschüler: Experten: $M = 10.58$ ($SD = 2.19$); Novizen: $M = 10.47$ ($SD = 2.63$); Jugendliche: Experten: $M = 11.25$ ($SD = 2.79$); Novizen: $M = 11.65$ ($SD = 3.21$); jüngere Erwachsene: Experten: $M = 11.00$ ($SD = 2.96$); Novizen: $M = 9.75$ ($SD = 2.88$); ältere Erwachsene: Experten: $M = 12.11$ ($SD = 2.17$); Novizen: $M = 11.60$ ($SD = 2.28$).

Auch bei den Normwerten im Wortschatztest (Petermann & Petermann, 2010; von Aster et al., 2006) unterschieden sich die Mittelwerte der Experten und Novizen nicht voneinander, jedoch die der vier Altersgruppen ($F(3,151) = 11.68$; $p < .001$; $\eta^2 = .19$). Der Post-hoc-Test ergab, dass die Ergebnisse der jüngeren Erwachsenen (9.19) etwas niedriger lagen als die der Grundschüler (11.95) und der älteren Erwachsenen (11.92). Die gemittelten Normwerte für die einzelnen Gruppen lagen jedoch ausnahmslos im Normbereich: Grundschüler: Experten: $M = 12.84$ ($SD = 2.32$); Novizen: $M = 11.05$ ($SD = 2.37$); Jugendliche: Experten: $M = 10.10$ ($SD = 1.94$); Novizen: $M = 10.95$ ($SD = 2.19$); jüngere Erwachsene: Experten: $M = 9.37$ ($SD = 2.73$); Novizen: $M = 9.00$ ($SD = 1.95$); ältere Erwachsene: Experten: $M = 11.61$ ($SD = 2.79$); Novizen: $M = 12.20$ ($SD = 2.73$).

Bezüglich der durchschnittlichen Schul- und Ausbildungszeit ergaben sich keine Unterschiede zwischen den jüngeren ($M = 14.69$ Jahre; $SD = 2.48$) und den älteren Erwachsenen ($M = 14.10$ Jahre; $SD = 4.17$). Auch die subjektive Einschätzung des eigenen Gesundheitszustandes unterschied sich nicht zwischen den beiden Altersgruppen (jüngere Erwachsene: $M = 1.88$; $SD = 0.65$; ältere Erwachsene: $M = 2.18$; $SD = 0.71$). Lediglich für die Zahl der durchschnittlich eingenommenen Medikamente zeigte sich ein bedeutsamer Unterschied zwischen den jüngeren und den älteren Erwachsenen (jüngere Erwachsene: $M = 0.39$; $SD = 0.70$; ältere Erwachsene: $M = 2.55$; $SD = 3.27$; $t(78) = 4.09$; $p < .001$).

6.3.2.3 Versuchsmaterial

Wortpaare. Das Lernmaterial bestand aus zwei Übungswortpaaren sowie 22 fußballbezogenen, konkreten Substantiven (s. Anhang A 3), die zur Hälfte semantisch zusammenhängend (z.B. Feld – Stadion), zur anderen Hälfte unzusammenhängend waren (z.B. Sieger – Pass). Die Wortpaare sowie die Präsentationszeiten wurden auf Grundlage einer Pilotierungsstudie mit $N = 43$ Probanden ausgewählt. Die Erfassung der EOLs, JOLs und der SUs erfolgte wie in der ersten Studie anhand einer siebenstufigen Skala. Die Instruktion für die

EOLs lautete „Wie leicht wäre dieses Wortpaar für dich/Sie zu lernen?“ mit Antwortmöglichkeiten von „sehr schwer“ bis „sehr leicht“.

Fußball-Wissenstest. Die Einteilung der Probanden in Experten und Novizen erfolgte auf der Grundlage von Ergebnissen in einem Fußball-Wissenstest. In der Altersgruppe der Grundschüler kam hierbei ein bereits verwendetes Verfahren von Schneider et al. (1989) zum Einsatz, bei dem die Probanden neun Wissens- sowie drei Interessensfragen, welche bei der Zuteilung der Teilnehmer unberücksichtigt blieben, durch Ankreuzen einer der vorgegebenen Antwortalternativen beantworten sollten (s. Anhang C 1). Bei den Wissensfragen war jeweils nur eine der Möglichkeiten korrekt. Für die drei anderen Altersgruppen musste ein neues Verfahren entwickelt werden. Dafür wurden auf der Grundlage einer Pilotierungsstudie mit $N = 68$ erwachsenen Teilnehmern aus einer Reihe von Fragen zehn ausgewählt, die besonders gut zwischen Experten und Novizen differenzierten sowie unterschiedliche Schwierigkeitsstufen beinhalteten (s. Anhang C 1). Es handelte sich hierbei sowohl um offene Fragen, die kurze, eindeutige Antworten verlangten, als auch um Einfachwahlaufgaben. Auch hier kamen drei weitere Fragen zum Fußballinteresse zum Einsatz, die in der Zuweisung zu Experten und Novizen keine Rolle spielten.

Zusatzvariablen. Zusätzlich wurden wie in Studie 1 die Untertests Zahlennachsprechen und Wortschatz aus HAWIK bzw. WIE (Petermann & Petermann, 2010; von Aster et al., 2006) sowie bei den jüngeren und älteren Erwachsenen der Fragebogen zu demografischen Angaben und Gesundheitszustand eingesetzt. Außerdem wurde in allen Altersgruppen mittels Fragebögen die Leistung im deklarativen Metagedächtnis erhoben. Da die Ergebnisse aus Letzteren nicht zu Zielsetzung dieser Arbeit gehören und deren Umfang überschreiten würden, wird darauf verzichtet, diese zu berichten.

6.3.2.4 Studienablauf

Sämtliche Erhebungen fanden als Einzeltestungen statt. Bei den minderjährigen Teilnehmern lag die schriftliche Einverständniserklärung der Erziehungsberechtigten vor.

Den Probanden wurden nacheinander die Wortpaare am Laptop präsentiert. Jedem Teil des Versuchs gingen zwei nicht fußballbezogene Wortpaare als Übungsaufgaben voraus. Bei der

Erhebung der EOLs gingen die Teilnehmer in ihrem eigenen Tempo vor. In der nachfolgenden Lernphase wurde den Grundschulern und den älteren Erwachsenen jedes Wortpaar acht Sekunden lang präsentiert, den Jugendlichen sechs und den jüngeren Erwachsenen drei Sekunden (vgl. Studie 1). Anschließend erfolgte die Erhebung der JOLs analog zu Studie 1. Im Intervall zwischen Lernen und Erinnern bearbeiteten die Teilnehmer den Fußballtest, die Tests zum Zahlennachsprechen und zum Wortschatz sowie die jüngeren und die älteren Erwachsenen den demografischen Fragebogen. Anschließend wurde das linke Wort jedes Paares präsentiert mit der Aufgabe, das rechte Wort zu nennen und danach die Sicherheit in die eigene Antwort einzuschätzen. Die Teilnehmer hatten hier nicht die Möglichkeit mit „weiß nicht“ zu antworten, sondern wurden aufgefordert, gegebenenfalls zu raten. Die Präsentation der Wortpaare erfolgte in randomisierter Reihenfolge. Jeder Proband erhielt jedoch in jeder Phase des Experiments dieselbe Reihenfolge.

Nach der Hauptphase der Studie wurden die Fragebögen zum deklarativen Metagedächtnis bearbeitet. Als Dank für die Teilnahme erhielten die minderjährigen Teilnehmer kleine Geschenke, die Psychologiestudenten Versuchspersonenstunden und die anderen erwachsenen Teilnehmer zehn Euro.

6.3.2.5 Überblick über die statistischen Auswertungen

Die Berechnungen wurden nach dem gleichen Prinzip wie in Studie 1 durchgeführt. Als weiterer Zwischensubjektfaktor neben der Altersgruppe (vierstufig) wurde die Expertise (zweistufig: Experten vs. Novizen) eingeführt. Wie erwartet ergaben sich unterschiedliche Verteilungen von männlichen und weiblichen Teilnehmern auf die Expertisebedingungen: Insgesamt wurden 66 männliche und 14 weibliche Experten identifiziert. Da die Literatur zu metakognitiven Überwachungsleistungen keinen Einfluss des Geschlechts diskutiert (z.B. Schneider, 2015), ist nicht davon auszugehen, dass dieser Faktor hier – über die Zuweisung zur jeweiligen Expertisebedingung hinaus – Einfluss auf die Ergebnisse nimmt. Aus diesem Grund wird im Weiteren nicht zwischen männlichen und weiblichen Probanden differenziert. Der Innersubjektfaktor Itemschwierigkeit (assoziativ vs. nicht assoziativ) wird in der Auswertung nicht berücksichtigt, da die Novizen z.T. nur sehr wenige unzusammenhängende Wortpaare erinnerten und sich dadurch in einigen Auswertungen die Freiheitsgrade zu stark reduzierte.

6.3.3 Ergebnisse

Die Ergebnisse werden getrennt nach Erinnerungsleistung, EOLs, JOLs und SUs berichtet.

6.3.3.1 Erinnerungsleistung.

Einen Überblick über den jeweiligen Prozentsatz richtiger Antworten bei Experten und Novizen gibt die Tabelle 5. Eine ANOVA mit den Zwischensubjektfaktoren Altersgruppe und Expertise ergab einen signifikanten Haupteffekt der Altersgruppe ($F(3,152) = 8.32; p < .001; \eta^2 = .14$). Post-hoc-Tests zeigten, dass die Grundschüler (35.34%) und die älteren Erwachsenen (40.27%) weniger Items korrekt erinnerten als die Jugendlichen (51.02%) und die jüngeren Erwachsenen (50.11%). Weiterhin wurden der Haupteffekt des Faktors Expertise ($F(1,152) = 25.45; p < .001; \eta^2 = .14$) sowie die Interaktion zwischen Altersgruppe und Expertise ($F(3,152) = 3.35; p < .05; \eta^2 = .06$) signifikant. *T*-Tests für jede Altersgruppe zeigten, dass bei den Grundschulern (27.73% vs. 42.95%; $p < .001$), bei den Jugendlichen (43.64% vs. 58.41%; $p < .05$) und bei den jüngeren Erwachsenen (38.41% vs. 61.82%; $p < .01$) die Experten den Novizen in der Erinnerungsleistung überlegen waren; bei den älteren Erwachsenen unterschieden sich die beiden Gruppen hingegen nicht.

Tabelle 5: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe und Expertise

	Experten	Novizen
Grundschüler	42.95 (14.11)	27.73 (10.31)
Jugendliche	58.41 (16.37)	43.64 (18.91)
Jüngere Erwachsene	61.82 (20.56)	38.41 (20.88)
Ältere Erwachsene	40.31 (14.60)	40.23 (15.70)

Standardabweichungen in Klammern.

6.3.3.2 Ease-of-Learning-Urteile

Wie in Studie 1 wurde die Überwachungsleistung sowohl hinsichtlich der Differenzierungsfähigkeit als auch hinsichtlich der Genauigkeit analysiert. Tabelle 6 zeigt die mittleren EOLs vor richtigen bzw. falschen Antworten. Eine ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren und Korrektheit der Antwort als Innersubjektfaktor

ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,152) = 9.51; p < .001; \eta^2 = .16$). In Post-hoc-Tests zeigte sich, dass die Grundschüler (5.51) höhere mittlere EOLs abgaben als die anderen drei Altersgruppen (Jugendliche: 4.56; jüngere Erwachsene: 4.53; ältere Erwachsene: 4.98). Des Weiteren wurde der Faktor Expertise signifikant ($F(1,152) = 9.74; p < .01; \eta^2 = .06$): Experten (5.13) gaben höhere EOLs als Novizen (4.66). Außerdem resultierte ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,152) = 112.65; p < .001; \eta^2 = .43$), der zeigt, dass die EOLs vor richtigen Antworten (5.27) höher ausfielen als vor falschen Antworten (4.53). Schließlich ergab sich eine signifikante Interaktion zwischen den Faktoren Altersgruppe und Korrektheit der Antwort ($F(3,152) = 4.87; p < .01; \eta^2 = .09$). Getrennte ANOVAs für die einzelnen Altersgruppen mit Korrektheit der Antwort als Innersubjektfaktor zeigten, dass die Teilnehmer aller Altersgruppen in der Lage waren, zwischen richtigen und falschen Antworten zu differenzieren (alle p 's $< .001$). Die Werte für η^2 waren jedoch bei den jüngeren Erwachsenen (.64) etwa doppelt so hoch wie bei den anderen Altersgruppen (Grundschüler: .26; Jugendliche: .37; ältere Erwachsene: .34).

Tabelle 6: Mittlere EOLs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Expertise

	Experten		Novizen	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	5.88 (0.77)	5.50 (1.09)	5.62 (0.90)	5.05 (1.05)
Jugendliche	5.08 (0.74)	4.45 (1.10)	4.65 (0.87)	4.05 (0.83)
Jüngere Erwachsene	5.41 (0.91)	4.17 (1.43)	4.83 (1.04)	3.71 (1.13)
Ältere Erwachsene	5.65 (1.01)	4.88 (1.29)	4.99 (1.29)	4.40 (0.90)

Standardabweichungen in Klammern.

Die Überwachungsleistung in den EOLs wurde in einem zweiten Schritt zusätzlich durch die Berechnung der Gammakorrelationen bestimmt (vgl. Tab. 7). In t -Tests zeigte sich, dass alle Korrelationen signifikant von Null verschieden waren (Grundschüler: alle p 's $< .05$; Jugendliche und Novizen unter den älteren Erwachsenen: alle p 's $< .01$; alle übrigen p 's $< .001$). Für die mittleren Korrelationen zwischen EOLs und Erinnerungsleistung wurde eine ANOVA mit den Zwischensubjektfaktoren Altersgruppe und Expertise durchgeführt, die einen signifikanten Haupteffekt des Faktors Altersgruppe ergab ($F(3,152) = 4.53; p < .05; \eta^2 = .08$).

Post-hoc-Tests zeigten, dass die Grundschüler (.25) im Mittel niedrigere Korrelationen aufwiesen als die jüngeren Erwachsenen (.56). Es ergaben sich weder ein Effekt der Expertise noch eine signifikante Interaktion.

Tabelle 7: Mittlere Gammakorrelationen für EOLs, JOLs und SUs in Abhängigkeit von Altersgruppe und Expertise

	EOLs		JOLs		SUs	
	Experten	Novizen	Experten	Novizen	Experten	Novizen
Grundschüler	.25 (.45)	.25 (.44)	.66 (.35)	.66 (.34)	.91 (.09)	.82 (.26)
Jugendliche	.32 (.46)	.30 (.38)	.85 (.17)	.76 (.20)	.98 (.06)	.92 (.11)
Jüngere Erwachsene	.63 (.28)	.49 (.32)	.83 (.20)	.67 (.36)	.96 (.05)	.88 (.17)
Ältere Erwachsene	.51 (.43)	.34 (.40)	.78 (.24)	.77 (.30)	.89 (.15)	.86 (.16)

Standardabweichungen in Klammern.

6.3.3.3 Judgments of Learning

Die JOLs wurden auf die gleiche Weise wie die EOLs analysiert. Die Tabelle 8 zeigt die mittleren JOLs vor richtigen bzw. falschen Antworten. Eine ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren und Korrektheit der Antwort als Innersubjektfaktor ergab einen signifikanten Haupteffekt des Faktors Expertise ($F(1,152) = 7.76; p < .01; \eta^2 = .05$). Experten (4.61) gaben höhere mittlere JOLs ab als Novizen (4.20). Außerdem zeigte sich ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,152) = 653.39; p < .001; \eta^2 = .81$). Vor richtigen Antworten (5.55) wurden höhere JOLs abgegeben als vor falschen (3.26). Weitere signifikante Haupteffekte oder Interaktionen ergaben sich nicht.

Bezüglich der JOL-Genauigkeit konnte mit *t*-Tests nachgewiesen werden, dass alle Gammakorrelationen von Null verschieden waren (alle p 's $< .001$; s. Tab. 7). Eine ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren ergab weder signifikante Haupteffekte noch Interaktionen.

Tabelle 8: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Expertise

	Experten		Novizen	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	5.72 (1.21)	3.85 (0.96)	5.42 (1.31)	3.23 (1.16)
Jugendliche	5.86 (0.93)	3.39 (0.85)	5.13 (1.08)	2.95 (1.03)
Jüngere Erwachsene	5.99 (0.49)	3.22 (0.95)	5.17 (1.22)	3.28 (1.23)
Ältere Erwachsene	5.68 (0.88)	3.15 (1.21)	5.45 (1.48)	2.98 (0.93)

Standardabweichungen in Klammern.

6.3.3.4 Sicherheitsurteile

Die mittleren SUs nach richtigen bzw. falschen Antworten in Abhängigkeit von Altersgruppe und Expertise sind in Tabelle 9 aufgeführt. Die entsprechende ANOVA ergab einen signifikanten Haupteffekt des Faktors Expertise ($F(1,152) = 18.00; p < .001; \eta^2 = .11$): Experten (4.48) gaben höhere mittlere SUs ab als Novizen (3.97). Weiterhin zeigte sich ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,152) = 1404.23; p < .001; \eta^2 = .90$). Nach richtigen Antworten (5.95) wurden höhere SUs abgegeben als nach falschen (2.50). Die Interaktion zwischen den Faktoren Expertise und Korrektheit der Antwort lag an der Grenze zur Signifikanz ($F(1,152) = 3.91; p = .05; \eta^2 = .03$). Nachgeschaltete *t*-Tests zeigten, dass Experten (6.30) nach richtigen Antworten höhere SUs abgaben als Novizen (5.60; $p < .001$). Der Unterschied bei den falschen Antworten war deutlich geringer, aber dennoch signifikant. Auch hier gaben die Experten etwas höhere SUs (2.67) ab als die Novizen (2.33; $p < .05$). Insgesamt differenzierten die Experten also geringfügig besser zwischen richtigen und falschen Antworten als die Novizen.

Im Hinblick auf die Genauigkeit der SUs zeigte sich, dass alle Gammakorrelationen zwischen SUs und Erinnerungsleistung von Null verschieden waren (alle p 's $< .001$; s. Tab. 7). Eine ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren ergab einen signifikanten Haupteffekt der Expertise ($F(1,152) = 8.87; p < .01; \eta^2 = .06$). Die Experten (.94) wiesen demnach höhere Genauigkeiten auf als die Novizen (.87).

Um die Gültigkeit der Ergebnisse zu überprüfen, wurden sämtliche Analysen mit einer strengeren Klassifikation von Expertise wiederholt. Teilnehmer mit Werten im Fußballtest, die im mittleren Drittel lagen, wurden von den Auswertungen ausgeschlossen. Da die Analysen

sich mit den oben berichteten decken, wird darauf verzichtet, sie an dieser Stelle ausführlich darzustellen.

Tabelle 9: Mittlere SUs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Expertise

	Experten		Novizen	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	6.31 (0.83)	3.31 (1.21)	5.76 (0.99)	2.47 (0.93)
Jugendliche	6.54 (0.58)	2.29 (0.84)	5.44 (1.36)	2.28 (1.15)
Jüngere Erwachsene	6.41 (0.45)	2.66 (0.97)	5.47 (1.06)	2.16 (1.00)
Ältere Erwachsene	5.94 (0.89)	2.40 (1.02)	5.74 (1.07)	2.43 (0.62)

Standardabweichungen in Klammern.

6.3.4 Diskussion

Diese Studie stellt den ersten Ansatz dar, Experten und Novizen einer großen Altersspanne hinsichtlich ihrer Überwachungsleistungen zu vergleichen. Da die vorliegende Literatur keine eindeutigen Rückschlüsse zulässt, ob sich ein hohes Vorwissen positiv auf die metakognitive Überwachung auswirkt (z.B. T. D. Griffin et al., 2009) oder eher zu einer Überschätzung der eigenen Leistungen führt (z.B. Son & Kornell, 2010), kann diese Frage bislang als ungeklärt gelten. Es ist weiterhin nicht eindeutig belegt, ob und inwiefern sich Expertise in verschiedenen Altersstufen unterschiedlich auf die Überwachungsleistung auswirkt. Im vorliegenden Experiment wurde zunächst auf recht einfach strukturierte Stimuli (Wortpaare) zurückgegriffen, um eine größere Vergleichbarkeit mit anderen Befunden zum Paar-Assoziationslernen zu gewährleisten.

Was die Erinnerungsleistung betrifft, ist festzuhalten, dass es durch die Anpassung der Präsentationszeiten gelungen ist, Boden- und Deckeneffekte zu vermeiden. Der Anteil korrekter Antworten lag zwischen 30 und 60%. Erwartungsgemäß wiesen die Experten unter den Grundschulern, den Jugendlichen und den jüngeren Erwachsenen bessere Erinnerungsleistungen auf als die Novizen (vgl. Nietfeld & Schraw, 2002; Schneider et al., 1989). Dies traf jedoch nicht für die Gruppe der älteren Erwachsenen zu: Hier zeigten sich keine Unterschiede zwischen den Experten und den Novizen. Eine differenziertere Betrachtung ihrer Erinnerungsleistung verdeutlicht, dass ältere Experten im Vergleich zu den Novizen

zusammenhängende Items besser erinnerten. Bei den nicht assoziierten Wortpaaren zeigte sich das gegenteilige Bild. Diese deskriptiven Ergebnisse können mit der „Hypothese des assoziativen Defizits“ (Naveh-Benjamin, 2000) erklärt werden, wonach ältere Erwachsene mehr Schwierigkeiten haben, neue Verbindungen zwischen ursprünglich nicht zusammenhängenden Items herzustellen. Dies könnte besonders auf die Experten in dieser Altersgruppe zutreffen, da diese deutlich stärkere Verknüpfungen zwischen fußballbezogenen Wortpaaren aufweisen sollten als die Novizen und somit neue Verbindungen ein stärkeres Überlernen darstellen sollten. Insgesamt kann festgehalten werden, dass das ausgewählte Material und die die Präsentationszeiten als angemessen gesehen werden können.

In den Überwachungsurteilen waren die Teilnehmer aller Altersgruppen erwartungsgemäß in der Lage, zwischen richtigen und falschen Antworten zu unterscheiden. Generell neigten die Probanden dazu, die Schwierigkeit der Aufgabe vor dem eigentlichen Lernvorgang zu unterschätzen, was sich an den hohen EOLs vor später falschen Antworten ablesen lässt. Sowohl an der Differenzierung als auch der Genauigkeit der EOLs ist zu erkennen, dass die jüngeren Erwachsenen den bevorstehenden Lernprozess besser einschätzen konnten als die anderen Altersgruppen, insbesondere als die Grundschüler; dies kann damit zusammenhängen, dass jüngere Erwachsene auch im täglichen Leben häufiger als die anderen Teilnehmer mit abstrakten Stimuli konfrontiert sind. Bei den JOLs und den SUs zeigten sich weder in der Differenzierungsfähigkeit noch in der Genauigkeit Alterseffekte. Insgesamt war also die Überwachungsleistung in den EOLs zufriedenstellend ausgeprägt; bei den JOLs und den SUs zeigten sich diesbezüglich sehr gute Fähigkeiten. Dieses Ergebnis ist in Übereinstimmung mit der Studie 1 sowie anderen Befunden, die zeigen, dass JOLs und SUs ab dem Grundschulalter nur geringen Veränderungen unterliegen (Schneider, 2015), und dass Verschlechterungen im höheren Erwachsenenalter selten sind (Hertzog & Dunlosky, 2011).

Der Schwerpunkt im vorliegenden Experiment lag auf der Frage, inwiefern sich große Vorwissensunterschiede in verschiedenen Lebensaltern auf die Überwachungsleistungen auswirken. Die Ergebnisse weisen insgesamt darauf hin, dass Experten dazu neigen, ihre eigenen Fähigkeiten optimistischer einzuschätzen als Novizen. In allen Differenzierungsmaßen wurde ein signifikanter Haupteffekt des Faktors Expertise gefunden, in dem Sinne, dass Experten höhere mittlere Urteile abgaben als Novizen und dies unabhängig davon, ob die spätere Antwort korrekt war oder nicht. Dies galt für alle vier Altersgruppen gleichermaßen; d.h. die Annahme, dass der Optimismus bei den älteren Erwachsenen am stärksten ausgeprägt ist (vgl. Shing et al., 2009; Toth et al., 2011), konnte nicht bestätigt werden. An dieser Stelle muss noch einmal darauf hingewiesen werden, dass es den Teilnehmern bekannt war, dass in

der Studie der Einfluss des Vorwissens überprüft wird. Von daher ist davon auszugehen, dass die Teilnehmer sich in ihren Urteilen nicht nur auf die Vertrautheit mit dem Stimulusmaterial stützen, sondern auch die Assoziation „Vorwissen – Sicherheit in eigene Leistungen“ getriggert wurde, was zu optimistischeren Selbsteinschätzungen führt (vgl. Son & Kornell, 2010). Dieser Effekt scheint die Auswirkungen von Vertrautheit im höheren Erwachsenenalter zu überlagern und für ähnliche Tendenzen in den Differenzierungsleistungen in allen Altersgruppen verantwortlich zu sein. Dennoch konnte in den SUs ein Vorteil zugunsten der Experten nachgewiesen werden, der jedoch nicht alterskorreliert war: Teilnehmer mit einem höheren Vorwissen differenzierten etwas besser zwischen richtigen und falschen Antworten; auch ihre Genauigkeit war höher ausgeprägt als bei den Novizen (vgl. Nietfeld & Schraw, 2002). Es ist demnach zu vermuten, dass Experten im Vergleich zu Novizen in den SUs besser dazu in der Lage sind, Informationen aus allen Stadien des Lernens und Erinnerns zu nutzen und dadurch zu einem zutreffenderen Urteil gelangen.

Insgesamt gesehen konnte der positive Einfluss von Vorwissen in verschiedenen Lebensaltern auf die Gedächtnisleistung bestätigt werden; für die Überwachungsleistung traf dies jedoch nur eingeschränkt zu. Die Ergebnisse zeigen lediglich für die SUs einen Vorteil zugunsten der Experten. Eine mögliche Einschränkung in Bezug auf die Gültigkeit dieses Befundes könnte die Tatsache sein, dass das Lernen von Wortpaaren nur ansatzweise die Organisation des Vorwissens im semantischen Netzwerk repräsentiert; d.h. eher abstrakte Paar-Assoziationen bilden das Wissen von Fußballexperten unter Umständen nur sehr ungenau ab. Dieses Problem soll in einem zweiten Experiment zum Vorwissen angegangen werden: Hier soll ein Text über ein konkretes Fußballspiel zum Einsatz kommen, also ein Format, welches die Domäne besser repräsentieren und damit mögliche Vorteile in der Überwachung bei Experten deutlicher zeigen sollte.

6.4 Studie 3: Einfluss des Vorwissens auf die Überwachungsleistung beim Lernen aus einem Text

6.4.1 Herleitung der Fragestellung

Der Einfluss einer breiten Wissensbasis auf die Qualität der metakognitiven Überwachung kann weder durch die vorhandene Literatur noch durch Ergebnisse der Studie 2 als abschließend geklärt gelten. Die Vorteile von Experten gegenüber Novizen, die sich in Gedächtnismaßen sehr eindeutig zeigen, ließen sich in der eigenen Studie nicht auf alle Überwachungsmaße gleichermaßen übertragen. Eine mögliche Einschränkung der Studie 2, die oben bereits erwähnt wurde, nämlich die Tatsache, dass das Vorwissen von Fußballexperten beim Lernen von Wortpaaren nicht angemessen abgebildet wird, soll nun in Studie 3 überwunden werden.

Das folgende Experiment ist in der Forschungslinie des Metacomprehension-Ansatzes einzuordnen, d.h. es geht um Überwachungsprozesse beim Lernen und Verstehen von Textmaterial. Bisherige Befunde zeigen in recht großer Übereinstimmung, dass Überwachungsurteile bei komplexeren Materialien, insbesondere bei Texten, häufig ungenauer ausfallen als beim Lernen von Wortpaaren (s. Glenberg & Epstein, 1985; Maki, 1998a; vgl. auch Studie 1). Thiede et al. (2009) führen mehrere Erklärungen für dieses Phänomen an: Insgesamt stellen Texte deutlich ungenauere Stimuli dar als konkrete Wortpaare. Folglich ist es für Probanden schwieriger, eine valide und eindeutige Aussage über das Ausmaß ihres Verstehens und Behaltens zu treffen. Wiley, Griffin und Thiede (2005) schlugen aus diesem Grund vor, anstelle eines einzigen globalen Verständnisurteiles – wie häufig üblich – immer mehrere Urteile für jeden Text zu erheben. Außerdem kann laut Thiede et al. (2009) davon ausgegangen werden, dass die Hinweisreize, die beim Textlernen in Überwachungsurteile einfließen, sich von denen bei Paar-Assoziationsdesigns unterscheiden. Insbesondere dann, wenn das Metagedächtnisurteil deutlich allgemeiner abgefragt wird als die Erinnerungsleistung, ist davon auszugehen, dass sich die Teilnehmer jeweils auf andere Hinweisreize stützen und die Qualität der Überwachung dadurch geringer ausfällt. Weiterhin ist die Überwachung von Texten dadurch anspruchsvoller, da die Aufmerksamkeit des Lesers zwischen der Verarbeitung der Information aus dem Text und der Überwachung aufgeteilt werden muss.

Es ist also wichtig, die Ergebnisse aus Studie 2 in komplexeren und alltagsnäheren Stimuli, wie Texte sie darstellen, zu validieren. Im folgenden Experiment bestand demnach die Aufgabe darin, die eigenen Gedächtnisleistungen bei der Beantwortung von Fragen zu einem Fußballtext und beim Wiedererkennen von Sätzen aus diesem Text zu beurteilen. Es sollen verschiedene

Überwachungsmaße verglichen werden – zum einen, um die Ergebnisse aus Studie 2 in einem anderen Kontext zu überprüfen, zum anderen, um zu untersuchen, wie die Effekte von Vorwissen sich in den verschiedenen Maßen auswirken. Generell kann angenommen werden, dass sich die Befunde aus Studie 2 replizieren lassen: Experten sollten ihre eigenen Leistungen sowohl prospektiv als auch retrospektiv optimistischer einschätzen als Novizen. Außerdem sollte sich ein Vorteil bezüglich der Qualität der Überwachung insbesondere in den SUs zeigen. Es wird insgesamt ein stärkerer Vorteil der Experten bezüglich der Überwachungsleistung angenommen als in Studie 2. Beim Lernen von Wortpaaren sollten Novizen eher auf vergleichbare Strategien zurückgreifen wie Experten, da die grundlegende Struktur der Stimuli recht reduziert und einfach ist. Beim Lernen aus ein- oder zweimalig dargebotenen Texten hingegen sollte es insgesamt schwieriger sein, spontan wirkungsvolle Strategien einzusetzen, sodass Experten hier aufgrund einer erleichterten Verarbeitung Vorteile in der Überwachung haben sollten. Es ist auch zu erwarten, dass Experten ein besseres Situationsmodell ausbilden als Novizen (vgl. Kintsch, 1998), da sie über eine breitere und besser vernetzte Wissensbasis verfügen und aus diesem Grund ihre Überwachungsurteile auf diagnostisch wertvollere Hinweisreize stützen sollten.

Es wird also insgesamt davon ausgegangen, dass die Experten den Novizen in der Überwachung ihrer Gedächtnisleistung überlegen sind, wenn es darum geht, Fragen zum Text zu beantworten. Dieser Vorteil sollte bis zum jüngeren Erwachsenenalter noch zunehmen, da das Vorwissen bis dahin noch ausgebaut wird und sich die Unterschiede zwischen Experten und Novizen vergrößern sollten. Beim Wiedererkennen von Sätzen – eine Aufgabe, die an der Textoberfläche ansetzt – sollten die Unterschiede zwischen beiden Versuchsgruppen kleiner ausfallen, da erwartet wird, dass die Expertise primär den Aufbau eines adäquaten Situationsmodells fördert. Davon abgesehen werden vergleichbare Ergebnisse zu Studie 2 erwartet, in dem Sinne, dass sich das Vorwissen unabhängig von der Korrektheit der Antwort in höheren Überwachungsurteilen niederschlagen sollte und dass die Experten den Novizen besonders in den SUs, aber auch in geringerem Maße in den JOLs überlegen sein sollten.

6.4.2 Methodik

6.4.2.1 Design

Die Studie lehnt sich im Design direkt an Studie 2 an. Es wurde ein 4 (Altersgruppe: Kinder, Jugendliche, jüngere und ältere Erwachsene) x 2 (Expertise: Experten vs. Novizen) faktorielles

Design verwendet. Die Altersgruppe und die Expertise wurden als Zwischensubjektfaktoren erfasst. Als abhängige Variablen wurden die Erinnerungsleistung sowie beim Beantworten von Fragen zum Text die Güte der globalen Einschätzung, der JOLs und der SUs erhoben; beim Wiedererkennen von Sätzen aus dem Text wurden SUs erfasst.

6.4.2.2 Stichprobe

An der Studie nahmen insgesamt $N = 160$ Personen teil. Es handelte sich um 40 Drittklässler (23 männlich, 17 weiblich; mittleres Alter: 8.20 Jahre; $SD = 0.41$; Altersspanne: 8 bis 9 Jahre), 40 Jugendliche (21 männlich, 19 weiblich; mittleres Alter: 13.30 Jahre; $SD = 0.95$; Altersspanne: 12 bis 15), 40 jüngere Erwachsene (20 männlich, 20 weiblich; mittleres Alter: 21.99 Jahre; $SD = 2.82$; Altersspanne: 17 bis 32) sowie 40 ältere Erwachsene (18 männlich, 22 weiblich; mittleres Alter: 69.08 Jahre; $SD = 4.85$; Altersspanne: 59 bis 77). Die Rekrutierung erfolgte analog zu den beiden Vorgängerstudien.

Es ergaben sich keine Unterschiede in den Leistungen im Zahlennachsprechen (Petermann & Petermann, 2010; von Aster et al., 2006) zwischen den Experten und den Novizen. Auch die einzelnen Altersgruppen wiesen vergleichbare Normwerte auf: Grundschüler: Experten: $M = 10.75$ ($SD = 2.05$); Novizen: $M = 11.45$ ($SD = 2.93$); Jugendliche: Experten: $M = 10.79$ ($SD = 3.21$); Novizen: $M = 9.20$ ($SD = 2.26$); jüngere Erwachsene: Experten: $M = 10.30$ ($SD = 2.23$); Novizen: $M = 10.65$ ($SD = 3.87$); ältere Erwachsene: Experten: $M = 11.15$ ($SD = 2.16$); Novizen: $M = 10.60$ ($SD = 2.95$).

Beim Wortschatztest (Petermann & Petermann, 2010; von Aster et al., 2006) zeigten sich ebenfalls keine Mittelwertsunterschiede zwischen den Experten und den Novizen. Allerdings unterschieden sich die vier Altersgruppen voneinander ($F(3,151) = 2.93$; $p < .05$; $\eta^2 = .06$). Der Post-hoc-Test ergab, dass die Jugendlichen (8.77) etwas niedrigere Leistungen erreichten als die Grundschüler (10.13). Im Einzelnen lauteten die Normwerte für die jeweiligen Versuchsgruppen wie folgt: Grundschüler: Experten: $M = 10.90$ ($SD = 2.51$); Novizen: $M = 9.35$ ($SD = 2.35$); Jugendliche: Experten: $M = 8.79$ ($SD = 1.67$); Novizen: $M = 8.75$ ($SD = 2.17$); jüngere Erwachsene: Experten: $M = 9.15$ ($SD = 2.01$); Novizen: $M = 9.50$ ($SD = 2.33$); ältere Erwachsene: Experten: $M = 10.45$ ($SD = 2.26$); Novizen: $M = 9.50$ ($SD = 2.84$).

In dieser Stichprobe unterschieden sich die jüngeren ($M = 14.45$ Jahre; $SD = 2.55$) von den älteren Erwachsenen ($M = 12.08$ Jahre; $SD = 3.14$) bezüglich der durchschnittlichen Schul- und Ausbildungszeit ($t(77) = 3.69$; $p < .001$). Weiterhin schätzen die älteren Erwachsenen ($M =$

2.09; $SD = 0.71$) ihren eigenen Gesundheitszustand etwas schlechter ein als die jüngeren ($M = 1.70$; $SD = 0.65$; $t(77) = 2.65$; $p < .05$). Auch die Zahl der durchschnittlich eingenommenen Medikation unterschied sich zwischen den beiden Altersgruppen (jüngere Erwachsene: $M = 0.44$; $SD = 0.74$; ältere Erwachsene: $M = 1.85$; $SD = 2.31$; $t(75) = 3.67$; $p < .001$).

6.4.2.3 Versuchsmaterial

Fußballtext. Der Fußballtext und die Fragen zum Text bzw. die Sätze zum Wiedererkennen wurden auf der Grundlage von ausführlichen Pilotierungen mit insgesamt $N = 50$ Teilnehmern erstellt. Angelehnt an die Fußballgeschichte von Schneider et al. (1989) wurden drei Texte verschiedener Länge und Schwierigkeit geschrieben: Der Text für die Grundschüler umfasste 626 Wörter, der für die Jugendlichen und die älteren Erwachsenen 795 Wörter und der für die jüngeren Erwachsenen 991 Wörter (s. Anhang A 4). Diese Studie stellt die einzige aus dieser Reihe von sechs Experimenten dar, bei denen den älteren Erwachsenen dieselben Aufgaben wie den Jugendlichen und nicht wie den Grundschulern dargeboten wurden, da nicht vorrangig fluide, sondern zu weiten Teilen kristalline Verarbeitungsprozesse involviert sind, die weniger von altersbedingten Verschlechterungen betroffen sein sollten (vgl. Baltes & Baltes, 1990). Vor der Pilotierung wurden die Texte zunächst von mehreren Fußballexperten sowie Grundschullehrkräften gegengelesen. Inhaltlich wurde im Text erzählt, wie ein Junge namens Max mit seinem Vater ein Fußballspiel besucht. Die Geschichte enthielt schwerpunktmäßig Informationen zum Verlauf des Spiels, aber auch Informationen zum Kontext, z.B. dass das Spiel an Max' Geburtstag stattfindet. Einige Informationen über das Spiel, so z.B. die Anzahl der Tore, wurden nicht explizit erwähnt, sondern mussten erschlossen werden. Obwohl davon ausgegangen wird, dass Vorwissen im Bereich Fußball die Bildung dieser Inferenzen erleichtert, war dies keine Voraussetzung für das Verstehen des Textes. Der Text wurde sowohl in schriftlicher Form als auch auditiv über Kopfhörer präsentiert.

Wie in Studie 1 beim Film wurden offene Fragen zur Messung der Erinnerungsleistung verwendet. Durch die Pilotierungsstudie wurde sichergestellt, dass die einzelnen Aufgaben in ihrer Schwierigkeit variierten und für jede Altersgruppe einen vergleichbaren Anspruch darstellten. Die jüngeren Erwachsenen erhielten 20 Fragen, alle anderen Altersgruppen 18 (s. Anhang A 5). Außerdem erhielt jede Altersgruppe 20 Sätze, die zur Hälfte wörtlich aus dem Text übernommen worden waren und zur Hälfte leichte Änderungen enthielten (s. Anhang A 6).

Die Überwachungsleistung wurde durch zum einen durch ein globales Verständnisurteil erfasst. Die Teilnehmer sollten auf einer siebenstufigen Smiley-Skala von 1 („sehr schlecht“) bis 7 („sehr gut“) angeben, wie gut sie die Geschichte insgesamt verstanden hatten.

Zum anderen wurden JOLs und SUs erfasst. Bei den Fragen zur Geschichte wurden die JOLs und SUs analog zu Studie 1 erhoben. Das heißt, es wurden bei den JOLs die gleichen Fragen wie die zur Erfassung der Erinnerungsleistung gestellt, mit dem Unterschied, dass der Einschub „Wie sicher bist du, dich später daran zu erinnern, dass...“ vorangestellt wurde. Die SUs wurden sowohl für die Beantwortung der Fragen als auch für das Wiedererkennen der Sätze erhoben. Hier sollten die Teilnehmer jeweils angeben, wie sicher sie sind, dass ihre Antwort korrekt ist. Die JOLs und die SUs wurden wieder auf siebenstufigen Smiley-Skalen erfasst.

Zusatzvariablen. Das fußballbezogene Vorwissen wurde wie in Studie 2 anhand des Fußballtests gemessen. Analog zu den bereits vorgestellten Studien kamen außerdem die Untertests Zahlennachsprechen und Wortschatz aus HAWIK bzw. WIE (Petermann & Petermann, 2010; von Aster et al., 2006) sowie bei den jüngeren und älteren Erwachsenen der Fragebogen zu demografischen Angaben und Gesundheitszustand zum Einsatz. Es wurde ebenfalls in allen Altersgruppen die deklarative Metagedächtnisleistung erhoben, welche an dieser Stelle aus genannten Gründen nicht näher berichtet werden soll.

6.4.2.4 Studienablauf

Die Erhebungen fanden wiederum als Einzeltestungen statt. Bei den minderjährigen Teilnehmern lag die schriftliche Einverständniserklärung der Erziehungsberechtigten vor.

Die Teilnehmer wurden als erstes dazu aufgefordert, sich die Fußballgeschichte über die Kopfhörer anzuhören. Es wurde ihnen mitgeteilt, dass ihnen später Fragen zur Geschichte gestellt werden würden. Nach dieser ersten Präsentation sollten die Probanden die Geschichte noch einmal anhören und gleichzeitig den geschriebenen Text mitlesen. Diese doppelte Darbietung sollte die Komplexität der Aufgabe verringern und eventuelle Nachteile von schwächeren Lesern ausgleichen. Anschließend sollten die Teilnehmer angeben, wie gut sie die Geschichte verstanden hatten (globales Verständnisurteil) und ihre JOLs zu den einzelnen Fragen zum Text angeben. Zwischen Lernen und Erinnern wurden die Untertests zum Wortschatz und zum Zahlennachsprechen durchgeführt. Die jüngeren und älteren Erwachsenen füllten zusätzlich den demografischen Fragebogen aus. Anschließend beantworteten die

Teilnehmer die Fragen zum Text und gaben für jedes Item ein SU ab. Es bestand nicht die Option, mit „weiß nicht“ zu antworten. Danach wurden den Teilnehmern die Sätze aus dem Text bzw. die Distraktoren dargeboten; sie sollten mit Ja oder Nein entscheiden, ob diese genauso im Originaltext vorkamen. Jede Antwort sollte mit einem SU versehen werden.

Nach der Hauptphase der Studie wurden die Fragebögen zum deklarativen Metagedächtnis bearbeitet. Als Dank für die Teilnahme erhielten die minderjährigen Teilnehmer kleine Geschenke, die Psychologiestudenten Versuchspersonenstunden und die anderen erwachsenen Teilnehmer zehn Euro.

6.4.2.5 Überblick über die statistischen Auswertungen

Die Berechnungen erfolgten analog zu Studie 2. Für die Variable globales Verständnisurteil wurden separate Korrelationen für jede Alters- und Versuchsgruppe durchgeführt und nach Cohen (1988) interpretiert. Wie in Studie 2 waren erwartungsgemäß deutlich mehr männliche als weibliche Teilnehmer unter den Experten (64 vs. 16). Aus den gleichen Gründen wie in Studie 2 wurde dennoch darauf verzichtet, den Faktor Geschlecht in die Auswertungen einzuschließen.

6.4.3 Ergebnisse

Die Ergebnisse werden getrennt nach Erinnerungsleistung für Fragen zum Text und für die Wiedererkennungsaufgabe, globalem Verständnisurteil, JOLs sowie SUs (ebenfalls für Fragen zum Text und für das Wiedererkennen) berichtet.

6.4.3.1 Erinnerungsleistung

Die Tabelle 10 zeigt die Prozentsätze richtiger Antworten auf Fragen zum Text. Eine ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,152) = 8.46; p < .001; \eta^2 = .14$). Der Post-hoc-Test zeigte, dass die Jugendlichen (.49) sowie die älteren Erwachsenen (.45) weniger Fragen korrekt beantworteten als die jüngeren Erwachsenen (.61). Weiterhin ergab sich ein signifikanter Haupteffekt des Faktors Expertise ($F(1,152) = 25.97; p < .001; \eta^2 = .15$): Experten (.59) wiesen mehr korrekte Antworten auf als Novizen (.46).

Für die Aufgabe zum Wiedererkennen von Sätzen wurde ebenfalls eine ANOVA mit den Zwischensubjektfaktoren Altersgruppe und Expertise durchgeführt (s. auch Tab. 10). Hier wurde ebenfalls der Haupteffekt des Faktors Altersgruppe signifikant ($F(3,152) = 6.07$; $p < .01$; $\eta^2 = .11$). Post-hoc-Tests zeigten, dass die Grundschüler (.64) sowie die jüngeren Erwachsenen (.65) mehr Sätze korrekt zuordneten als die älteren Erwachsenen (.54). Weiterhin wurde auch hier der Haupteffekt des Faktors Expertise signifikant ($F(1,152) = 6.85$; $p < .05$; $\eta^2 = .04$): Experten (.64) erzielten bessere Leistungen als Novizen (.56).

Tabelle 10: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe, Expertise und Modalität der Erinnerungsabfrage

	Experten		Novizen	
	Fragen zum Text	Klassifikation von Sätzen	Fragen zum Text	Klassifikation von Sätzen
Grundschüler	60.28 (17.22)	66.00 (14.74)	47.22 (16.57)	61.25 (9.85)
Jugendliche	54.58 (18.98)	63.25 (16.80)	42.78 (14.43)	60.50 (9.58)
Jüngere Erwachsene	67.50 (12.83)	67.75 (13.62)	55.50 (13.95)	62.00 (9.23)
Ältere Erwachsene	51.58 (15.19)	57.75 (12.08)	38.61 (13.55)	51.00 (7.88)

Standardabweichungen in Klammern

6.4.3.2 Globales Verständnisurteil

Um die Vorhersagekraft des globales Verständnisurteils zu überprüfen, wurden bivariate Korrelationen zwischen diesem Urteil und der späteren Leistung bei der Beantwortung der Fragen einerseits sowie beim Wiedererkennen von Sätzen andererseits berechnet (s. Tab. 11). Bis auf eine Ausnahme war keine Korrelation von Null verschieden. Lediglich bei den Novizen unter den Grundschulern war die Korrelation zwischen dem Globalurteil und der Leistung im Satz wiedererkennen signifikant ($p < .05$).

Tabelle 11: Korrelationen von globalem Verständnisurteil und Erinnerungsleistung in Abhängigkeit von Altersgruppe, Expertise und Modalität der Erinnerungsabfrage

	Experten		Novizen	
	Fragen zum Text	Klassifikation von Sätzen	Fragen zum Text	Klassifikation von Sätzen
Grundschüler	.02	-.09	-.26	-.48
Jugendliche	.05	.06	.21	.18
Jüngere Erwachsene	-.21	.29	.13	.01
Ältere Erwachsene	.26	.03	.14	-.22

6.4.3.3 Judgments of Learning

Zur Analyse der Differenzierungsleistungen wurden wieder die mittleren JOLs vor richtigen vs. falschen Antworten verglichen (s. Tab. 12). Neben der Korrektheit der Antworten als Innersubjektfaktor gingen die Altersgruppe und die Expertise als Zwischensubjektfaktoren in die ANOVA ein. Es ergab sich ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,152) = 336.90$; $p < .001$; $\eta^2 = .69$): Die mittleren JOLs vor den richtigen Antworten (5.53) lagen höher als die vor den falschen (4.40). Außerdem zeigten sich ein signifikanter Haupteffekt des Faktors Expertise ($F(1,152) = 30.34$; $p < .001$; $\eta^2 = .17$) und ein signifikanter Haupteffekt des Faktors Altersgruppe ($F(3,152) = 3.66$; $p < .05$; $\eta^2 = .07$). Weiterhin wurden sowohl die Wechselwirkung zwischen diesen beiden Faktoren ($F(3,152) = 5.08$; $p < .01$; $\eta^2 = .10$) als auch die Dreifachinteraktion zwischen Korrektheit der Antwort, Altersgruppe und Expertise signifikant ($F(3,152) = 3.10$; $p < .05$; $\eta^2 = .06$). Für eine erleichterte Interpretation dieser Befunde wurden zunächst nur die JOLs vor korrekten Antworten mit den Faktoren Altersgruppe und Expertise ausgewertet. Hier ergab die entsprechende ANOVA lediglich einen signifikanten Haupteffekt des Faktors Expertise ($F(1,152) = 18.15$; $p < .001$; $\eta^2 = .11$). Experten (5.48) gaben höhere JOLs vor richtigen Antworten als Novizen (5.21). Bei den falschen Antworten zeigte sich ein signifikanter Haupteffekt des Faktors Altersgruppe ($F(3,152) = 6.44$; $p < .001$; $\eta^2 = .11$), ein signifikanter Haupteffekt des Faktors Expertise ($F(1,152) = 27.02$; $p < .001$; $\eta^2 = .15$) und eine signifikante Interaktion zwischen beiden Faktoren ($F(3,152) = 3.34$; $p < .05$; $\eta^2 = .06$). Bei einer weiteren Untergliederung der Analyse bei falschen Antworten in die einzelnen Altersgruppen ergaben sich lediglich bei den Kindern ($F(1,38) = 25.08$; $p < .001$; $\eta^2 = .40$; Experten: 4.69; Novizen: 2.91) und den älteren Erwachsenen ($F(1,38) = 7.67$; $p < .01$; $\eta^2 = .17$; Experten: 5.06; Novizen: 4.28) signifikante Einflüsse des Faktors Expertise. Die Experten schätzen ihre eigenen Leistungen vor richtigen

Antworten demnach insgesamt optimistischer ein als die Novizen, bei den falschen Antworten war dies nur bei den Grundschulern und den älteren Erwachsenen der Fall.

Tabelle 12: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Expertise

	Experten		Novizen	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	5.97 (0.92)	4.69 (1.18)	5.27 (0.95)	2.91 (1.06)
Jugendliche	5.74 (0.81)	4.30 (1.10)	5.11 (1.05)	3.70 (1.06)
Jüngere Erwachsene	5.51 (0.94)	3.90 (1.13)	5.29 (0.79)	3.53 (1.22)
Ältere Erwachsene	6.15 (0.85)	4.27 (0.64)	5.18 (1.13)	5.06 (1.10)

Standardabweichungen in Klammern.

Die weiterführende Analyse der JOLs mit Gammakorrelationen zeigte, dass alle Korrelationen von Null verschieden waren (bei Novizen unter den älteren Erwachsenen: $p < .01$; alle anderen p 's $< .001$; s. Tab. 13). Eine ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,152) = 3.45$; $p < .05$; $\eta^2 = .06$). In den Post-hoc-Tests zeigte sich, dass die älteren Erwachsenen (.42) niedrigere Korrelationen aufwiesen als die Grundschüler (.65) und die jüngeren Erwachsenen (.59).

Tabelle 13: Mittlere Gammakorrelationen für EOLs, JOLs und SUs in Abhängigkeit von Altersgruppe, Expertise und Modalität der Erinnerungsabfrage

	JOLs		SUs (Fragen zum Text)		SUs (Klassifikation von Sätzen)	
	Experten	Novizen	Experten	Novizen	Experten	Novizen
Grundschüler	.55 (.29)	.75 (.23)	.61 (.44)	.74 (.21)	.17 (.63)	.12 (.47)
Jugendliche	.57 (.37)	.51 (.24)	.76 (.19)	.51 (.29)	.48 (.38)	.13 (.53)
Jüngere Erwachsene	.58 (.26)	.60 (.26)	.65 (.29)	.62 (.25)	.24 (.40)	.24 (.32)
Ältere Erwachsene	.51 (.47)	.33 (.47)	.68 (.45)	.52 (.31)	.37 (.17)	.38 (.14)

Standardabweichungen in Klammern.

6.4.3.4 Sicherheitsurteile

Fragen zum Text. Die mittleren Sicherheitsurteile nach richtigen bzw. falschen Antworten bei den Fragen zum Text flossen ebenfalls zur Bestimmung der Differenzierungsleistung als Innersubjektfaktor in eine ANOVA ein. Die entsprechenden Mittelwerte nach Altersgruppe und Expertise getrennt sind in Tabelle 14 einzusehen. In der ANOVA, die außerdem die Zwischensubjektfaktoren Altersgruppe und Expertise beinhaltete, wurde ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort gefunden ($F(1,152) = 398.13; p < .001; \eta^2 = .72$). Nach richtigen Antworten (5.86) wurden im Mittel höhere SUs abgegeben als nach falschen (4.26). Es ergaben sich außerdem ein signifikanter Haupteffekt des Faktors Expertise ($F(3,152) = 33.63; p < .001; \eta^2 = .18$) sowie eine signifikante Dreifachinteraktion zwischen allen eingeschlossenen Faktoren ($F(3,152) = 2.76; p < .05; \eta^2 = .05$). Für eine erleichterte Interpretation dieser Effekte wurde zunächst nur für die SUs nach richtigen Antworten eine ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren berechnet. Hier zeigte sich ein signifikanter Haupteffekt des Faktors Expertise ($F(1,152) = 28.03; p < .001; \eta^2 = .16$). Die SUs der Experten (6.22) lagen höher als die der Novizen (5.50). Für die SUs nach falschen Antworten zeigte sich ebenfalls ein signifikanter Haupteffekt des Faktors Expertise ($F(1,152) = 25.17; p < .001; \eta^2 = .14$). Auch hier gaben die Experten (4.74) höhere SUs ab als die Novizen (3.79). Die Interaktion zwischen Altersgruppe und Expertise verfehlte das Signifikanzniveau.

Tabelle 14: Mittlere SUs bei Fragen zum Text in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Expertise

	Experten		Novizen	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	6.26 (0.73)	4.84 (1.38)	5.51 (1.22)	3.13 (0.99)
Jugendliche	5.54 (1.05)	4.53 (1.28)	6.12 (0.75)	3.93 (1.13)
Jüngere Erwachsene	5.98 (0.76)	4.56 (1.22)	5.61 (0.81)	3.93 (1.13)
Ältere Erwachsene	6.53 (0.52)	4.74 (1.27)	5.33 (0.89)	3.92 (1.01)

Standardabweichungen in Klammern.

Die Gammakorrelationen zwischen SUs nach Fragen zum Text und Leistung sind in Tabelle 13 abgebildet. *T*-Tests zeigten, dass alle Korrelationen von Null verschieden waren (alle p 's $< .001$). Eine ANOVA mit den Zwischensubjektfaktoren Altersgruppe und Expertise ergab keine

signifikanten Haupteffekte, aber eine Wechselwirkung zwischen beiden Faktoren ($F(3,151) = 2.24$; $p < .05$; $\eta^2 = .05$). Nachfolgende t -Tests, die durchgeführt wurden, um Experten und Novizen in jeder Altersgruppe zu vergleichen, zeigten, dass die Experten (.76) nur bei den Jugendlichen höhere Korrelationen aufwiesen als die Novizen (.51). In allen anderen Altersgruppen unterschieden sich die Korrelationen nicht zwischen den beiden Versuchsbedingungen.

Klassifikation von Sätzen. Die mittleren SUs nach richtigen vs. falsch klassifizierten Sätzen sind in Tabelle 15 einzusehen. Eine ANOVA mit dem Innersubjektfaktor Korrektheit der Antwort und den Zwischensubjektfaktoren Altersgruppe und Expertise ergab einen signifikanten Haupteffekt des Faktors Expertise ($F(1,152) = 11.26$; $p < .05$; $\eta^2 = .03$). Experten (5.82) gaben demnach im Mittel höhere SUs ab als Novizen (5.44). Weiterhin ergaben sich ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,152) = 4.32$; $p < .05$; $\eta^2 = .03$), ein signifikanter Haupteffekt der Altersgruppe ($F(3,152) = 9.51$; $p < .05$; $\eta^2 = .07$) sowie eine Interaktion zwischen Altersgruppe und Korrektheit der Antwort ($F(3,152) = 4.53$; $p < .01$; $\eta^2 = .08$). T -Tests für die einzelnen Altersgruppen zeigten, dass die Grundschüler und die älteren Erwachsenen nicht in der Lage waren, zwischen korrekt und falsch klassifizierten Sätzen zu unterscheiden; bei den Jugendlichen und den jüngeren Erwachsenen waren die Unterschiede in den SUs bei richtigen vs. falschen Klassifikationen signifikant (Jugendliche: $t(39) = 3.50$; $p < .01$; 5.90 vs. 5.42; jüngere Erwachsene: $t(39) = 5.18$; $p < .001$; 5.56 vs. 4.98).

Tabelle 15: Mittlere SUs bei der Klassifikation von Sätzen in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Expertise

	Experten		Novizen	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	6.41 (0.55)	6.20 (0.72)	5.45 (1.73)	6.30 (3.61)
Jugendliche	6.06 (0.79)	5.49 (1.07)	5.73 (0.60)	5.34 (1.18)
Jüngere Erwachsene	5.76 (0.47)	5.20 (0.90)	5.36 (0.67)	4.76 (1.11)
Ältere Erwachsene	5.78 (0.83)	5.64 (0.91)	5.30 (0.94)	5.31 (1.10)

Standardabweichungen in Klammern.

Die SU-Genauigkeit bei der Klassifikation von Sätzen wurde ebenfalls durch die Berechnung von Gammakorrelationen bestimmt (s. Tab. 13). *T*-Tests für die einzelnen Versuchsgruppen ergaben, dass die Korrelationen bei den Experten unter den Jugendlichen ($p < .01$), bei den jüngeren Erwachsenen ($p < .05$ für die Experten; $p < .01$ für die Novizen) sowie den älteren Erwachsenen (p 's $< .001$) signifikant von Null verschieden waren, in den anderen Gruppen hingegen nicht. Die ANOVA mit Altersgruppe und Expertise als Zwischensubjektfaktoren ergab weder signifikante Haupteffekte noch Wechselwirkungen.

6.4.4 Diskussion

Die Studie 3 diente dem Ziel, die Ergebnisse der Studie 2 in einem Metacomprehension-Design auf ihre Generalisierbarkeit zu überprüfen. Da Texte komplexere und alltagsnähere Materialien darstellen, können die Ergebnisse aus Paar-Assoziationsstudien nicht unmittelbar übertragen werden (de Bruin et al., 2007). Außerdem sollte mit dieser Studie die bislang unzureichende Datenlage zum Einfluss der Expertise auf die Überwachungsleistung in verschiedenen Altersstufen erweitert werden.

Die Ergebnisse zeigen, dass die Schwierigkeit der Items wieder in zufriedenstellendem Maße zwischen den Altersgruppen angepasst werden konnte und weder Boden- noch Deckeneffekte auftraten: Die Erinnerungsleistung lag zwischen 40% und 70% für die Fragen zum Text. 50% bis 70% der Sätze wurden richtig klassifiziert. Es wurden nur geringfügige Alterseffekte in der Erinnerungsleistung gefunden. Es zeigte sich lediglich, dass die jüngeren Erwachsenen etwas besser abschnitten als die Jugendlichen. Erwartungsgemäß erbrachten die Experten in beiden Aufgabenbereichen bessere Ergebnisse als die Novizen. Das Wiedererkennen der Sätze schien den Teilnehmern insgesamt schwer zu fallen, da die Leistung kaum über dem Zufallsniveau lag.

Das Globalurteil stellte sich sowohl in Bezug auf die Beantwortung der Fragen als auch in Bezug auf die Klassifikation von Sätzen als sehr ungenaues Rating heraus. Dieses Ergebnis deckt sich mit Vorbefunden, die besagen, dass eine große Vergleichbarkeit zwischen der Abfrage des Überwachungsurteils und der Erinnerungsleistung wichtig ist, damit ein Selbsttest von diagnostischem Wert ist (z.B. Begg et al., 1989). Selbst wenn also ein allgemeines Urteil einen „Standard“ (Dunlosky, Rawson & McDonald, 2002, S. 79) in der Metacomprehension-Literatur darstellt, zeigen die Ergebnisse, dass die Überwachung deutlich detaillierter erfasst werden muss.

Diese Anforderung wurde bei den JOLs umgesetzt. Diese konnten nur bei den Fragen zum Text erfasst werden. Sowohl die Differenzierungsfähigkeit als auch die Genauigkeit fielen insgesamt zufriedenstellend aus. Die Teilnehmer aller Alters- und beider Vorwissensgruppen gaben höhere mittlere Urteile vor richtigen als vor falschen Antworten. Auch zwischen den Altersgruppen waren die Leistungen vergleichbar; lediglich die älteren Erwachsenen waren insgesamt etwas optimistischer als die anderen drei Altersgruppen bezüglich der späteren Erinnerungsleistung. Die mittlere Gammakorrelation zwischen Vorhersage und Erinnerungsleistung von .55 ist vergleichbar mit Befunden aus anderen Studien zur Metacomprehension (z.B. Dunlosky, Rawson & Hacker, 2002). Alterseffekte zeigten sich hier nur in der Hinsicht, dass die Grundschüler etwas höhere Werte aufwiesen als die älteren Erwachsenen. Dieses Ergebnis könnte daran liegen, dass sich die Senioren insgesamt mehr auf ihr größeres Weltwissen verließen und sich deshalb etwas optimistischer einschätzten.

Bei den SUs zu den Fragen zum Text waren die Teilnehmer ebenfalls in der Lage, zwischen richtigen und falschen Antworten zu unterscheiden. Die Genauigkeit in den SUs lag auf einem moderaten Niveau und war vergleichbar mit anderen Studien (z.B. von der Linden & Roebbers, 2006). Bei der Klassifikation der Sätze gaben die Probanden höhere SUs nach richtigen als nach falschen Antworten; dieser Unterschied war jedoch wesentlich kleiner als bei den Fragen zum Text und nur bei den Jugendlichen und den jüngeren Erwachsenen statistisch bedeutsam. Die eher niedrigen Gammakorrelationen bei dieser Aufgabe zeigen ebenfalls, dass es den Teilnehmern schwerfiel, die Korrektheit der Antworten einzuschätzen. Auch aus anderen Studien ist bekannt, dass die Überwachungsleistung bei Aufgaben zum Wiedererkennen deutlich niedriger ist als bei Aufgaben zur freien Erinnerung, da bei ersteren deutlich mehr Raten auftritt und dadurch die Unsicherheit größer ist (z.B. Buratti & Allwood, 2012).

Wie in Studie 2 lag der Schwerpunkt auf dem Einfluss des Vorwissens auf die verschiedenen Überwachungsmaße in den jeweiligen Altersgruppen. Beim Globalurteil konnten keine konsistenten Expertiseeffekte identifiziert werden. Bei den spezifischeren Maßen (JOLs und SUs) zeigten sich dagegen mehr Auswirkungen des Vorwissens. Bei der Differenzierungsleistung in den JOLs schnitten die Experten unter den Grundschulern und den älteren Erwachsenen schlechter ab als die Novizen, was auf höhere JOLs vor falschen Antworten zurückzuführen ist; in den beiden anderen Altersgruppen zeigten sich hier keine Unterschiede. Dieses Muster ergab sich auch für die SUs bei der Beantwortung der Fragen zum Text, verfehlte hier aber das Signifikanzniveau. Die Ergebnisse konnten nicht direkt auf die Genauigkeit übertragen werden. Bei den JOLs ergab sich kein Einfluss der Expertise. Bei den SUs zu den Fragen zum Text waren die mittleren Gammakorrelationen in den drei älteren

Altersgruppen bei den Experten höher als bei den Novizen. Der Unterschied erreichte jedoch nur bei den Jugendlichen ein signifikantes Ausmaß. In beiden Maßen kann die Tatsache, dass sich die Expertise nur so wenig auf die Überwachungsleistung auswirkt, auch mit den insgesamt hohen Standardabweichungen (bis zu .47) erklärt werden. Es gab also sowohl bei den Experten als auch bei den Novizen Teilnehmer, die ihre Leistung sehr gut überwachten, wie auch Teilnehmer, denen dies weniger gut gelang.

Bei der Klassifikation von Sätzen war der Einfluss des Vorwissens noch weniger ausgeprägt. In der Auswertung der Differenzierungsleistung zeigte sich, dass die Experten insgesamt höhere SUs gaben, unabhängig davon, ob die Antwort richtig oder falsch war. Bei den Gammakorrelationen ergaben sich keine Unterschiede zwischen den beiden Gruppen. Das heißt, die Experten waren insgesamt etwas sicherer bezüglich ihrer Antworten, profitierten jedoch nicht von ihrem Vorwissen in Bezug auf die Qualität der Überwachung. Wörtliches und sinngemäßes Erinnern scheinen demnach zwar nicht von völlig identischen Überwachungsvorgängen begleitet zu sein, der Einfluss der Expertise scheint jedoch vergleichbar zu sein.

Die Ergebnisse der beiden hier vorgestellten Studien legen nahe, dass die kognitiven und die metakognitiven Effekte von Vorwissen getrennt betrachtet werden müssen. Die oft berichteten Vorteile von Experten gegenüber Novizen bezüglich der Gedächtnisleistung (vgl. Feltovich, Prietula & Ericsson, 2006; Simon & Chase, 1973) konnten in beiden Studien gefunden werden. Dieser positive Effekt ließ sich jedoch nicht vollständig auf die Überwachungsleistung übertragen – ein Ergebnis, das im Gegensatz zu anderen Befunden steht (z.B. T. D. Griffin et al., 2009; Nietfeld & Schraw, 2002).

Die beiden Studien zeigen, dass verschieden komplexe Lernmaterialien und Formen der Erinnerungsabfrage nur zum Teil vergleichbar sind. Die Hypothese, dass Fußballexperten beim Lernen aus Textmaterial mehr von ihrem Vorwissen profitieren als beim Lernen von Wortpaaren konnte nicht bestätigt werden. Es zeigte sich vielmehr, dass Experten aller Altersgruppen beim Paar-Assoziationslernen bezüglich der Überwachungsgenauigkeit in den SUs von ihrem Vorwissen profitieren konnten. Bei den Fragen zum Fußballtext zeigen zwar die deskriptiven Daten, dass die SU-Gammakorrelationen der Experten im Mittel in allen Altersgruppen größer ausfielen als bei den Novizen; dieser Unterschied erreichte jedoch nur für die Gruppe der Jugendlichen das Signifikanzniveau. Dies könnte einerseits daran liegen, dass die Korrelationen in den SUs bei den Fragen zum Text insgesamt ein sehr hohes Niveau aufwiesen, sodass die Unterschiede deutlich größer ausfallen müssten, um signifikant zu werden. Andererseits ist es möglich, dass das Lernen von Paar-Assoziationen eher dem

semantischen Netzwerk von Experten entspricht und von daher die Überwachungsleistung – zumindest bei den SUs – genauer ausfällt als bei den Fragen zum Text.

Bezüglich der Differenzierungsleistung waren in Studie 3 bei den Fragen zum Text die Novizen den Experten sogar bei den Grundschulern und den älteren Erwachsenen in den JOLs überlegen. Hier scheint das Vorwissen demnach einen negativen Effekt auf die Überwachungsleistung zu haben. Dieses Ergebnis steht bei den Kindern im Kontrast zu den Vorbefunden von Schneider et al. (1989), die in dieser Altersgruppe eine bessere Überwachungsleistung bei den Experten gefunden hatten. Eventuell fiel für die Grundschüler in der eigenen Untersuchung die Überschneidung zwischen Vorwissen und präsentiertem Stimulusmaterial eher gering aus und es musste eine große Menge vollständig neuer Information verarbeitet werden (vgl. van Loon et al., 2013). Für die älteren Erwachsenen entsprechen die Ergebnisse in etwa den Befunden von Toth et al. (2011) in dem Sinne, dass diese Altersgruppe im Vergleich zu den jüngeren Erwachsenen bei vorwissensbezogenem Material etwas schlechtere Überwachungsleistungen aufweist, da hier vermehrt auf Vertrautheit mit den Items statt auf die tatsächliche Verarbeitung der neuen Information zurückgegriffen wird (vgl. auch Shing et al., 2009).

Das für alle Altersgruppen durchgängigste Ergebnis besteht jedoch darin, dass die Experten ihre eigenen Leistungen im Vergleich zu Novizen positiver einschätzten und zwar unabhängig von der Korrektheit der Antwort. Auch wenn dies nicht direkt in der Instruktion erwähnt wurde, war es für die Probanden offensichtlich, dass die Studien sich mit dem Themenbereich Fußball beschäftigten, wodurch vermutlich bei den Experten „Sicherheit“ getriggert wurde, bei den Novizen eher „Unsicherheit“ (vgl. Son & Kornell, 2010). Die Einstellung zur eigenen Expertise könnte demnach wie ein Anker gewirkt haben, der im Verlauf des Lernprozesses mehr oder weniger erfolgreich gemäß der Lernerfahrungen angepasst wird (Zhao & Linderholm, 2008). Sowohl dieser Anker als auch das bei Experten wahrscheinlich automatisiertere Verarbeiten der Information (de Bruin et al., 2007) könnten als Hinweisreize für die Überwachungsurteile herangezogen worden sein und zu einer Überschätzung der eigenen Leistung geführt haben. In diesem Kontext muss angemerkt werden, dass es bei der Mehrzahl der vorliegenden Studien zum Einfluss des Vorwissens auf die metakognitive Überwachung nicht möglich ist, das Ausmaß eines Überoptimismus der Experten abzuleiten, da in der Regel Gammakorrelationen verwendet werden. So gaben die Experten bei de Bruin et al. (2007) positivere Leistungsvorhersagen als die Novizen ab, erreichten jedoch auch insgesamt bessere Ergebnisse. In der Arbeit von T. D. Griffin et al. (2009) zeigten sich die Probanden mit Vorwissen weniger pessimistisch bezüglich der eigenen Leistung, was der Überschätzung in den eigenen Studien entsprechen könnte.

Zusammenfassend lässt sich festhalten, dass die Experten in unseren Studien sowohl pro- als auch retrospektiv und unabhängig von der Korrektheit ihrer Antworten zu höheren Einschätzungen als die Novizen tendierten. Diese insgesamt höheren Urteile waren nur bei den SUs in der Paar-Assoziationsaufgabe genauer als die von Novizen. Bessere Gedächtnisleistungen und ein größeres Vorwissen scheinen also unabhängig vom Alter der Probanden nicht zwangsläufig zu besseren Überwachungsleistungen zu führen. Es ist anzunehmen, dass sich bei der Mehrzahl der Urteile andere Hinweisreize, wie z.B. Itemcharakteristika oder der Zeitpunkt der Abfrage, stärker auf die Qualität der Überwachung auswirken als die Expertise. Eine genauere Untersuchung der zugrunde liegenden Mechanismen könnte einen Erklärungsansatz für die divergierenden Ergebnisse der vorliegenden Studien in diesem Bereich sein und bietet einen vielversprechenden Ausgangspunkt für weitere Forschungsarbeiten.

Da bislang als weitgehend ungeklärt gelten kann, welche Hinweisreize bei besseren im Vergleich zu schlechteren Gedächtnisleistungen für die metakognitive Überwachung genutzt werden, soll in der folgenden Studie ein anderer Ansatz gewählt werden, um dieser Frage nachzugehen.

6.5 Studie 4: Einfluss eines Strategietrainings auf die Überwachungsleistung

6.5.1 Herleitung der Fragestellung

Während sich die beiden vorherigen Studien damit befassten, wie sich Vorwissen, das der Person bereits zur Verfügung steht, in den verschiedenen Altersgruppen auf die Überwachungsleistung auswirkt, soll nun in der vierten Studie untersucht werden, inwiefern Überwachungsvorgänge „von außen“, d.h. durch eine Instruktion verbessert werden können. Dazu soll die Hälfte der Teilnehmer in der Anwendung einer Strategie unterwiesen werden.

Um nachzuvollziehen, wie ein Strategietraining die Überwachungsleistung verbessern kann, muss zunächst auf die Grundlagen von JOLs und SUs – die beiden Maße, die in der vorliegenden Studie zum Einsatz kommen – eingegangen werden. Wie in Kapitel 3.1.3 und 3.1.5 beschrieben, ist die Güte der pro- und retrospektiven Urteile davon abhängig, welche Hinweisreize für diese Einschätzungen genutzt werden. Je valider (d.h. „diagnostischer“) diese für die Gedächtnisleistung sind, desto genauer und differenzierter fallen die Urteile aus (Dunlosky & Metcalfe, 2009; Koriat, 1997, 2012). Hinweisreize, die Erinnerungsleistung und Überwachung in ähnlicher Weise beeinflussen, sind demnach am aussagekräftigsten. Bislang wurden sowohl für unmittelbar und verzögert erfasste JOLs als auch für SUs jeweils eine Reihe verschiedener Hinweisreize diskutiert, wie z.B. die Leichtigkeit der Verarbeitung für unmittelbare JOLs (Begg et al., 1989) oder die Lebhaftigkeit (vividness) des Abrufs für SUs (M. D. Robinson et al., 2000). Zur Klassifikation der verschiedenen Hinweisreize wird in neuerer Zeit vermehrt zwischen „Recollection“- und vertrauthetsbasierten Hinweisreizen unterschieden (Daniels et al., 2009; McCabe & Soderstrom, 2011; Metcalfe & Finn, 2008; Toth et al., 2011). Erstere beziehen sich auf einen bewussten, kontrollierten und intentionalen Zugang zu Gedächtnisinhalten, wodurch auch deren Details abrufbar werden. Vertrauthetsbasierte Hinweisreize meinen dagegen eher unscharfe Aktivierungen vorangegangener Erfahrungen, die auf vorhandenen semantischen Repräsentationen beruhen.

Verschiedene Vorbefunde stützen die Annahme, dass „Recollection“-Prozesse die Qualität von Überwachungsurteilen verbessern. So sind beispielsweise verzögerte JOLs deutlich genauer als unmittelbar erfasste JOLs, was damit erklärt werden kann, dass bei einer Verzögerung vermehrt bewusst auf Informationen im Langzeitgedächtnis zurückgegriffen wird und die Probanden somit mehr auf idiosynkratische Hinweisreize des Einspeicher- und Abrufvorgangs rekurren (Koriat, 1997). Idiosynkratische Hinweisreize sind solche, die persönliche, auf die Items

bezogene Details, wie z.B. Bilder oder Assoziationen enthalten. Es konnte weiterhin auch für unmittelbar erfasste JOLs gezeigt werden, dass für „Recollection“-basierte Gedächtnisinhalte höhere (Daniels et al. 2009; Toth et al. 2011) und genauere Urteile abgegeben wurden (Toth et al., 2011).

Auch für SUs gibt es Befunde, die zeigen, dass „Recollection“-Prozesse zu besseren – d.h. genaueren – Urteilen führen als das Stützen auf Vertrautheit (z.B. Kelley & Sahakyan, 2003). Auch für Defizite in der Qualität der SUs im höheren Erwachsenenalter wurden Verschlechterungen in den „Recollection“-Prozessen verantwortlich gemacht (Kelley & Sahakyan, 2003; Shing et al., 2009; Wong, Cramer & Gallo, 2012). Für Kinder existieren bislang keine Studien, die die Rolle von Recollection bei SUs untersuchen.

Es liegt also die Annahme nahe, dass ein Training, welches durch die vermehrte Nutzung idiosynkratischer Hinweisreize „Recollection“-Prozesse fördert, positive Auswirkungen auf die Qualität von JOLs und SUs hat. Im Entwicklungsverlauf sollten vor allem ältere Erwachsene, aber auch Kinder am meisten von einem solchen Training bezüglich ihrer Überwachungsleistung profitieren, denn in diesen beiden Altersgruppen sind Produktionsdefizite im Strategiegebrauch am stärksten ausgeprägt (Naveh-Benjamin, Brav & Levy, 2007; Pressley & Levin, 1977). Obwohl insgesamt recht wenige entwicklungsbedingte Veränderungen in den JOLs gefunden wurden, zeigen neuere Befunde, dass „Recollection“-Prozesse in bestimmten Situationen für eine geringere Genauigkeit bei älteren Erwachsenen verantwortlich sind (Daniels et al., 2009; Toth et al., 2011). Die Qualität der SUs unterliegt stärkeren Anstiegen im Grundschulalter (Schneider, 2015), die von „Recollection“-Prozessen mitbestimmt werden (Roderer & Roebbers, 2010). Auch Leistungseinbußen in den SUs im höheren Erwachsenenalter hängen vermutlich damit zusammen, dass in dieser Altersgruppe Vertrautheitsprozesse eine größere Rolle spielen als Recollection (Kelley & Sahakyan, 2003).

In der folgenden Studie bestanden die Lernaufgaben aus Wortpaaren. Je die Hälfte der Teilnehmer jeder Altersgruppe bekam ein Training in der „visual-imagery“-Strategie, um die Nutzung von idiosynkratischen Hinweisreizen zu fördern. Die Probanden wurden also angeleitet, zu zwei Wörtern ein interaktives Bild zu entwickeln, um somit sowohl beim Enkodier- als auch beim Abrufprozess mehr Details zu dem jeweiligen Wortpaar zur Verfügung zu haben. Diese Strategie stellt dem heutigen Stand nach eine der wirkungsvollsten Methoden dar, Paarassoziationen zu memorieren (Richardson, 1998). Die Wirkung auf die Überwachungsleistung wurde dagegen bislang nur bedingt untersucht. Hertzog et al. (2010) fanden, dass bei jüngeren Erwachsenen der spontane Gebrauch einer „Interactive-imagery“-

Strategie die JOL-Genauigkeit bedeutsam verbesserte. Bei Robinson, Hertzog und Dunlosky (2006) erhielten jüngere und ältere Erwachsene eine Instruktion, jedoch kein Training in „Interactive imagery“. Die Autoren konnten zeigen, dass der Gebrauch der Strategie sowohl mit der Erinnerungsleistung als auch mit der Höhe der JOLs zusammenhing. Weiterhin konnte nachgewiesen werden, dass die Strategie einen diagnostischen Hinweis für die JOLs darstellte. Insgesamt fehlt es also bei den JOLs vor allem an Studien mit Kindern und Jugendlichen, die ein solches Training anwenden.

Bezüglich der SUs zeigten sich in einer Studie von Shing et al. (2009), dass Teilnehmer im Alter zwischen 10 und 75 Jahren von einem Training in „Visual imagery“ profitierten: Die Differenz in den SUs zwischen Treffern und falschen Alarmen war nach dem Training deutlich größer.

Es wird also erwartet, dass das Strategietraining sich in allen Altersgruppen positiv auf die Erinnerungs- und die Überwachungsleistung (JOLs und SUs) auswirkt. Kinder und ältere Erwachsene sollten dabei am meisten profitieren.

Neben der pro- und retrospektiven Leistungseinschätzung, wie sie auch in den drei bisherigen Studien erfasst wurde, wurden die Teilnehmer in diesem Experiment zusätzlich gebeten, die Grundlage ihrer Antwort einzuschätzen, d.h. ob „Recollection“- oder Vertrautheitsprozesse bzw. keine Erinnerung vorlagen (RFN-Urteil; vgl. Daniels et al., 2009; Toth et al., 2011).

6.5.2 Methodik

6.5.2.1 Design

Es kam ein 4 (Altersgruppe: Kinder, Jugendliche, jüngere und ältere Erwachsene) x 2 (Versuchsbedingung: mit vs. ohne Strategietraining) faktorielles Design zum Einsatz. Die Altersgruppe und die Expertise wurden als Zwischensubjektfaktoren erfasst. Als abhängige Variablen wurden die Erinnerungsleistung sowie die Güte der JOLs und der SUs erhoben. Die Teilnehmer wurden nach dem Zufallsprinzip der Experimentalgruppe (mit Training) und der Kontrollgruppe (ohne Training) zugewiesen.

6.5.2.2 Stichprobe

An der Studie nahmen insgesamt $N = 160$ Personen teil. Darunter befanden sich 40 Drittklässler (23 männlich, 17 weiblich; mittleres Alter: 8.38 Jahre; $SD = 0.49$; Altersspanne: 8 bis 9 Jahre), 40 Jugendliche (28 männlich, 12 weiblich; mittleres Alter: 12.73 Jahre; $SD = 0.72$; Altersspanne: 12 bis 14), 40 jüngere Erwachsene (11 männlich, 29 weiblich; mittleres Alter: 22.75 Jahre; $SD = 2.02$; Altersspanne: 19 bis 26) sowie 40 ältere Erwachsene (23 männlich, 17 weiblich; mittleres Alter: 68.40 Jahre; $SD = 4.08$; Altersspanne: 60 bis 75). Die Teilnehmer stammten aus der gleichen Population wie in den anderen Studien.

Im Zahlennachsprechen (Petermann & Petermann, 2010; von Aster et al., 2006) waren die Leistungen der beiden Versuchsgruppen vergleichbar; es zeigten sich keine Effekte der Versuchsbedingung. Die ANOVA mit Altersgruppe und Versuchsbedingung als Zwischensubjektfaktoren ergab einen signifikanten Haupteffekt der Altersgruppe ($F(3,145) = 3.31$; $p < .05$; $\eta^2 = .06$). In Post-hoc-Tests konnten jedoch keine signifikanten Gruppenunterschiede identifiziert werden. Im Einzelnen ergaben sich folgende gemittelte Normwerte: Grundschüler: Experimentalgruppe (EG): $M = 9.30$ ($SD = 2.03$); Kontrollgruppe (KG): $M = 10.43$ ($SD = 2.65$); Jugendliche: EG: $M = 9.00$ ($SD = 2.31$); KG: $M = 10.00$ ($SD = 2.85$); jüngere Erwachsene: EG: $M = 10.90$ ($SD = 2.88$); KG: $M = 11.20$ ($SD = 2.33$); ältere Erwachsene: EG: $M = 11.00$ ($SD = 2.62$); KG: $M = 10.65$ ($SD = 2.62$).

Auch im Wortschatztest (Petermann & Petermann, 2010; von Aster et al., 2006) ergab die ANOVA mit Altersgruppe und Versuchsbedingung als Zwischensubjektfaktoren lediglich einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,144) = 14.00$; $p < .001$; $\eta^2 = .23$). Im Post-hoc-Test zeigte sich, dass die Jugendlichen (8.13) insgesamt niedrigere Werte aufwiesen als die Teilnehmer der drei anderen Altersgruppen (Grundschüler: 10.01, jüngere Erwachsene: 10.73; ältere Erwachsene: 11.30). Auch hier hatten demnach beide Versuchsgruppen vergleichbare Leistungen.

Die Gruppe der jüngeren und der älteren Erwachsenen unterschied sich nicht hinsichtlich der durchschnittlichen Schul- und Ausbildungszeit (jüngere Erwachsene: $M = 15.14$ Jahre; $SD = 2.18$; ältere Erwachsene: $M = 14.10$ Jahre; $SD = 3.84$). Ihren Gesundheitszustand schätzten die jüngeren Erwachsenen ($M = 1.63$; $SD = 0.67$) positiver ein als die älteren ($M = 2.25$; $SD = 0.71$; $t(78) = 4.07$; $p < .001$). Weiterhin berichteten die älteren Erwachsenen ($M = 2.34$; $SD = 2.31$), durchschnittlich mehr Medikamente pro Tag einzunehmen als die jüngeren ($M = 0.64$; $SD = 0.75$; $t(77) = 4.36$; $p < .001$).

6.5.2.3 Versuchsmaterial

Wortpaare. Das Lernmaterial umfasste 45 Wortpaare bei den Grundschulern und den älteren Erwachsenen, 54 bei den Jugendlichen und 60 bei den jüngeren Erwachsenen (s. Anhang A 7). Die Wortpaare gehörten verschiedenen semantischen Kategorien (z.B. Möbelstücke, Tiere, Fahrzeuge) an; bei jeweils der Hälfte der Wortpaare stammten in der Lernphase beide Teile aus derselben Kategorie, bei der anderen Hälfte waren beide Wörter aus verschiedenen Kategorien. Die Erinnerungsphase beinhaltete zu 50% Items, die identisch, d.h. in derselben Kombination, in der Lernphase präsentiert worden waren. 15 Wortpaare bei den Grundschulern und den älteren Erwachsenen, 18 bei den Jugendlichen und 20 bei den älteren Erwachsenen wurden neu kombiniert dargeboten, d.h. beide Wörter kamen in der Lernphase vor, jedoch nicht in dieser Zusammensetzung. Zusätzlich kam die gleiche Anzahl von vollständig neuen Item-Paaren in jeder Altersgruppe zum Einsatz. Sowohl die identischen, als auch die neu kombinierten und die neuen Paare enthielten jeweils zur Hälfte Wörter aus derselben und aus verschiedenen Kategorien. Die Auswahl der Items erfolgte auf Grundlage einer Pilotstudie mit $N = 35$ Teilnehmern.

Im Gegensatz zu den bisherigen Experimenten fanden hier die JOL- wie auch die spätere SU-Abgabe am Computer statt. Auf der visuellen Analog-Skala, die in Anlehnung an ein Thermometer mit einer Einstellung von blau („sehr kalt“) bis rot („sehr heiß“) gestaltet war, konnte von den Teilnehmern eingestellt werden, wie sicher sie sind, das Itempaar korrekt wiederzuerkennen, bzw. wie sicher sie sich sind, dass die gegebene Antwort korrekt ist. Dabei entsprachen die Einstellungen der Probanden in der Auswertung Werten von 0 (sehr kalt = sehr unsicher) bis 100 (sehr heiß = sehr sicher). Weiterhin wurden RFN-Urteile für jede gegebene Antwort erfasst mit der Frage „Worauf beruht deine/Ihre Antwort?“. Die Antwortoptionen umfassten „Ich kann mich genau daran erinnern“ (Recollection), „Das Wortpaar kommt mir vertraut vor“ (Familiarity/Vertrautheit) und „Ich kann mich nicht an das Wortpaar erinnern“ (No Memory/Keine Erinnerung).

Strategietraining. Die Probanden in der Experimentalgruppe erhielten ein ca. zehnminütiges Training zur „Interactive-imagery“-Strategie. Dabei wurde zunächst theoretisch und anhand von zwei Beispielbildern erklärt, inwiefern die Bildung eines interaktiven mentalen Bildes den Lernvorgang unterstützt. Anschließend wurden die Teilnehmer aufgefordert, die Strategie bei zehn Wortpaaren einzuüben und dem Versuchsleiter jeweils ein interaktives Bild zu nennen (z.B. zu „Wiese“ und „Bleistift“ das Bild „ein Bleistift zeichnet eine Wiese“). Die Probanden

erhielten ein Feedback über die Qualität ihres Bildes und wurden gegebenenfalls aufgefordert, ein anderes Bild zu generieren. Die Grundschüler erhielten eine sprachlich leicht vereinfachte Version des Trainings (s. Anhang C 2). Auch das Training wurde in der Pilotierungsstudie erfolgreich getestet.

Fragebogen zum Strategietraining. Zur Überprüfung einer erfolgreichen Manipulation wurden die Teilnehmer in der Experimentalgruppe am Ende der Untersuchung befragt, ob und wie oft die instruierte Strategie verwendet wurde. Die Probanden in der Kontrollgruppe wurden aufgefordert anzugeben, welche Strategien sie verwendet hatten (s. Anhang C 3).

Zusatzvariablen. Als Zusatzvariablen wurden wie in den bereits vorgestellten Studien die Leistungen im Zahlennachsprechen und im Wortschatz (Petermann & Petermann, 2010; von Aster et al., 2006) erfasst. Wieder kam bei den jüngeren und älteren Erwachsenen der Fragebogen zu demografischen Angaben und zum Gesundheitszustand zum Einsatz. Die Ergebnisse aus den zusätzlich verwendeten Fragebögen zum deklarativen Metagedächtnis werden hier ebenfalls nicht berichtet.

6.5.2.4 Studienablauf

Wie in den anderen Experimenten handelte es sich um Einzeltestungen in den Schulen oder im Labor, die bei den minderjährigen Teilnehmern mit dem schriftlichen Einverständnis der Eltern durchgeführt wurden.

Die Experimentalgruppe erhielt zunächst das Strategietraining. Vor der Lernphase wurden die Teilnehmer darüber unterrichtet, dass sie die Items anschließend in der korrekten Zuordnung wiedererkennen müssen. Ihnen wurden vier Übungspaare sowie die zu lernenden Items präsentiert. Die Darbietungsdauer betrug bei den Grundschulern und den älteren Erwachsenen 8 Sekunden, bei den Jugendlichen 6 Sekunden und bei den jüngeren Erwachsenen 2,5 Sekunden. Auch diese Zeiten ergaben sich aus den Ergebnissen der Pilotstudie. Anschließend wurden die JOLs abgefragt; hier wurde jeweils nur das linke Wort dargeboten, um ein erneutes Lernen zu verhindern. In der Phase zwischen Lernen und Wiedererkennen wurden die Untertests aus HAWIK bzw. WIE sowie bei den jüngeren und älteren Erwachsenen der demografische Fragebogen bearbeitet. In der Erinnerungsabfrage wurden den Probanden wie

oben beschrieben identische, neu kombinierte und vollständig neue Wortpaare dargeboten. Es sollte jeweils die Frage „Kam dieses Wortpaar genau in dieser Form in der Lernphase vor?“ mit „Ja“ oder „Nein“ beantwortet werden. Anschließend wurde für jedes Itempaar das SU eingeschätzt sowie ein RFN-Urteil abgegeben. Danach wurden der Fragebogen zum Strategietraining sowie die Fragebögen zum deklarativen Metagedächtnis bearbeitet.

Die Entschädigungen für die Probanden entsprachen denen der vorherigen Studien.

6.5.2.5 Überblick über die statistischen Auswertungen

Die Auswertungen erfolgten vergleichbar zu den bisherigen Studien. In die ANOVAs gingen nun die Zwischensubjektfaktoren Altersgruppe und Versuchsbedingung ein. Abgesehen von einer zusätzlichen Auswertung bei der Erinnerungsleistung wurde aus Gründen der Interpretierbarkeit auf eine Aufteilung nach Art des Wortpaares (identisch, neu kombiniert, neu) verzichtet. Aus denselben Gründen wurde der Faktor Assoziativität nicht berücksichtigt. Für die Auswertung der RFN-Urteile wurde – da sich die Anzahl der Items zwischen den Altersgruppen unterschied – der Prozentsatz ausgerechnet, zu dem von jedem Teilnehmer jede Option ausgewählt wurde. Der Innersubjekt-Faktor RFN-Urteil wurde dann zusätzlich in die ANOVA aufgenommen. Um den spontanen und den instruierten Strategie-Gebrauch zu vergleichen, wurde ebenfalls anhand von ANOVAs verglichen, wie viele Teilnehmer jeder Altersgruppe und Versuchsbedingung berichteten, eine „Interactive-imagery“-Strategie anzuwenden. Die offenen Antworten der Teilnehmer wurden von zwei unabhängigen Beobachtern kategorisiert ($\kappa = .95$).

6.5.3 Ergebnisse

6.5.3.1 Erinnerungsleistung

Die Gedächtnisleistung wurde als Prozentsatz korrekt wiedererkannter Items definiert; dies schloss sowohl identische Wortpaare ein, die als „alt“ klassifiziert wurden, als auch neu kombinierte oder vollständig neue Paare, die als „neu“ eingestuft wurden. Die erste Spalte der Tabelle 16 zeigt den mittleren Anteil richtig klassifizierter Wortpaare. Eine ANOVA mit Altersgruppe und Versuchsbedingung als Zwischensubjektfaktoren ergab einen signifikanten Haupteffekt der Altersgruppe ($F(3,152) = 5.34; p < .01; \eta^2 = .10$). Die Post-hoc-Tests zeigten, dass die jüngeren Erwachsenen etwas besser abschnitten als die Grundschüler (80.05% vs.

70.99% korrekt). Außerdem wurde ein signifikanter Haupteffekt der Versuchsbedingung gefunden ($F(1,152) = 21.82$; $p < .001$; $\eta^2 = .13$): Die Teilnehmer in der Trainingsgruppe schnitten demnach besser ab als die in der Kontrollgruppe (89.75% vs. 72.04%).

Tabelle 16: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe, Versuchsbedingung und Art des Wortpaares

	Gesamt	Identische WP	Neu kombinierte WP	Neue WP
Grundschüler				
Experimentalgruppe	76.62 (11.12)	71.16 (18.25)	72.14 (19.65)	91.43 (10.26)
Kontrollgruppe	65.36 (6.20)	59.56 (16.34)	53.93 (19.32)	88.21 (12.75)
Jugendliche				
Experimentalgruppe	80.65 (9.44)	76.23 (11.97)	75.73 (14.85)	94.17 (9.45)
Kontrollgruppe	69.83 (11.48)	63.89 (16.70)	62.67 (20.30)	88.61 (11.18)
Jüngere Erwachsene				
Experimentalgruppe	81.27 (12.12)	78.56 (14.57)	74.61 (22.56)	93.22 (10.05)
Kontrollgruppe	78.82 (9.15)	75.00 (10.38)	72.99 (18.74)	61.50 (9.69)
Ältere Erwachsene				
Experimentalgruppe	79.75 (10.87)	73.79 (18.12)	77.14 (21.81)	95.00 (14.34)
Kontrollgruppe	72.04 (10.99)	79.11 (14.32)	52.86 (28.33)	85.00 (21.23)

Standardabweichungen in Klammern. WP = Wortpaare.

Die übrigen Spalten der Tabelle 16 unterteilen die Wiedererkennensrate noch etwas genauer nach Art des Wortpaares. Es wurden getrennte ANOVAs mit Altersgruppe und Versuchsbedingung als Zwischensubjektfaktoren für jeden Typ durchgeführt, um differenziertere Aussagen über die Wirkung des Strategietrainings treffen zu können. Für die identischen Wortpaare ergab sich ein signifikanter Haupteffekt der Altersgruppe ($F(3,152) = 5.14$; $p < .01$; $\eta^2 = .09$). Die Post-hoc-Tests verdeutlichten, dass die Grundschüler (65.37%) weniger identische Items richtig identifizierten als die jüngeren (76.78%) und die älteren Erwachsenen (76.45%). Der Haupteffekt der Versuchsbedingung erreichte ebenfalls das Signifikanzniveau ($F(1,152) = 5.24$; $p < .05$; $\eta^2 = .03$). Auch hier war dies auf bessere Leistungen der Experimentalgruppe (74.93%) im Vergleich zur Kontrollgruppe (69.39%) zurückzuführen. Bei den neu kombinierten Wortpaaren wurde lediglich der Faktor Versuchsbedingung signifikant ($F(1,152) = 18.52$; $p < .001$; $\eta^2 = .11$). Die Teilnehmer in der Trainingsgruppe (74.91%) klassifizierten mehr Items richtig als die in der Kontrollgruppe (60.61%). Auch für die neuen Wortpaare wies der Haupteffekt des Faktors Versuchsbedingung

die gleiche Richtung auf ($F(3,152) = 7.04; p < .01; \eta^2 = .04$). Die Teilnehmer in der Experimentalgruppe (93.45%) waren denen in der Kontrollgruppe (88.33%) überlegen.

6.5.3.2 Judgments of Learning

Tabelle 17 zeigt die mittleren JOLs vor richtigen vs. falschen Antworten in Abhängigkeit von Altersgruppe und Versuchsbedingung. Die dazugehörige ANOVA ergab einen signifikanten Haupteffekt des Faktors Korrektheit der Antwort ($F(1,151) = 135.38; p < .001; \eta^2 = .74$). Vor später richtigen Antworten (57.98) wurden demnach im Mittel höhere JOLs abgegeben als vor falschen (46.78). Außerdem wurden die Interaktion der Faktoren Korrektheit der Antwort und Versuchsbedingung ($F(3,151) = 6.22; p < .05; \eta^2 = .04$) sowie die Dreifachinteraktion zwischen Korrektheit der Antwort, Versuchsbedingung und Altersgruppe signifikant ($F(3,151) = 3.43; p < .05; \eta^2 = .06$). Um die Richtung dieser Wechselwirkungen zu bestimmen, wurden getrennte ANOVAs für Experimental- und Kontrollgruppe mit dem Innersubjektfaktor Korrektheit der Antwort und dem Zwischensubjektfaktor Altersgruppe berechnet. In der Experimentalgruppe wurde hierbei ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort gefunden ($F(1,75) = 89.07; p < .001; \eta^2 = .54$). Die JOLs vor den richtigen Antworten (59.58) fielen demnach höher aus als die vor den falschen (45.98). Auch in der Kontrollgruppe wurde der Faktor Korrektheit der Antwort signifikant ($F(1,76) = 4.36; p < .001; \eta^2 = .38$). Hier ergab sich zusätzlich eine signifikante Interaktion zwischen den Faktoren Korrektheit der Antwort und Altersgruppe ($F(3,76) = 5.93; p < .01; \eta^2 = .19$). Nachfolgende *t*-Tests für die einzelnen Altersgruppen zeigten, dass in der Kontrollbedingung nur die Jugendlichen und die jüngeren Erwachsenen zwischen richtigen und falschen Antworten differenzierten (Jugendliche: 58.47 vs. 48.92; $t(19) = 4.45; p < .001$; jüngere Erwachsene: 59.62 vs. 42.60; $t(19) = 8.13; p < .01$).

Die mittleren Gammakorrelationen für die einzelnen Altersgruppen und Versuchsbedingungen sind in Tabelle 18 zu sehen. Mithilfe von *t*-Tests wurde festgestellt, dass – mit Ausnahme der Grundschüler in der Kontrollgruppe – alle Werte signifikant von Null verschieden waren (ältere Erwachsene in der Kontrollgruppe: $p < .05$; alle anderen p 's $< .001$). Eine ANOVA mit Altersgruppe und Versuchsbedingung als Zwischensubjektfaktoren ergab einen signifikanten Haupteffekt der Versuchsbedingung ($F(1,151) = 11.36; p < .001; \eta^2 = .07$) und eine signifikante Interaktion zwischen beiden Faktoren ($F(3,151) = 3.58; p < .05; \eta^2 = .05$). Zur Interpretation dieser Effekte wurden *t*-Tests für die einzelnen Altersgruppen getrennt berechnet. Bei den Grundschulern zeigte sich ein signifikanter Unterschied zugunsten der Experimentalgruppe

(.32 vs. .03; $t(38) = 3.50$; $p < .01$). Bei den älteren Erwachsenen wiesen die Ergebnisse dieselbe Richtung auf, verfehlten jedoch das Signifikanzniveau (.29 vs. .14; $t(38) = 1.82$; $p = .077$).

Tabelle 17: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Versuchsbedingung

	Experimentalgruppe		Kontrollgruppe	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	58.40 (16.06)	44.33 (16.88)	50.39 (13.72)	48.06 (12.12)
Jugendliche	53.96 (14.84)	37.16 (12.52)	58.47 (13.50)	48.92 (15.21)
Jüngere Erwachsene	60.60 (12.14)	48.76 (13.37)	59.62 (7.06)	42.60 (10.83)
Ältere Erwachsene	65.36 (17.83)	53.68 (21.98)	57.01 (14.04)	50.71 (18.05)

Standardabweichungen in Klammern.

Tabelle 18: Mittlere Gammakorrelationen für JOLs und SUs in Abhängigkeit von Altersgruppe und Versuchsbedingung

	JOLs		SUs	
	Experimental- gruppe	Kontroll- gruppe	Experimental- gruppe	Kontroll- gruppe
Grundschüler	.32 (.28)	.03 (.24)	.38 (.29)	.18 (.20)
Jugendliche	.30 (.19)	.20 (.19)	.42 (.21)	.33 (.21)
Jüngere Erwachsene	.27 (.23)	.32 (.15)	.35 (.23)	.42 (.20)
Ältere Erwachsene	.29 (.25)	.14 (.26)	.49 (.26)	.38 (.26)

Standardabweichungen in Klammern.

6.5.3.3 Sicherheitsurteile

Auch für die mittleren SUs nach richtigen bzw. falschen Antworten (s. Tab. 19) wurde eine ANOVA mit dem Innersubjektfaktor Korrektheit der Antwort und den Zwischensubjektfaktoren Altersgruppe und Versuchsbedingung berechnet. Hier zeigte sich ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,151) = 188.75$; $p < .001$; $\eta^2 = .56$); nach richtigen Antworten (79.02) wurden höhere SUs abgegeben als nach falschen (66.60). Weiterhin wurden der Haupteffekt des Faktors Altersgruppe ($F(3,151) = 6.26$; $p < .001$;

$\eta^2 = .11$) sowie die Interaktion zwischen Korrektheit der Antwort und Altersgruppe signifikant ($F(3,151) = 5.61; p < .001; \eta^2 = .10$). Zur leichteren Interpretation wurden separate ANOVAs mit dem Faktor Altersgruppe für die richtigen und die falschen Antworten berechnet. Für die SUs nach korrekten Antworten ergab die ANOVA einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,156) = 3.42; p < .05; \eta^2 = .06$). In den Post-hoc-Tests zeigte sich, dass die älteren Erwachsenen (82.57) höhere SUs abgaben als die jüngeren (74.25). Auch bei den SUs nach falschen Antworten wurde ein signifikanter Haupteffekt des Faktors Altersgruppe gefunden ($F(3,155) = 7.89; p < .001; \eta^2 = .13$). Hier ergaben die Post-hoc-Tests, dass die jüngeren Erwachsenen (57.33) niedrigere SUs abgaben als die drei anderen Altersgruppen (Grundschüler: 72.63; Jugendliche: 69.00; ältere Erwachsene: 67.42).

Tabelle 19: Mittlere SUs in Abhängigkeit von Korrektheit der Antwort, Altersgruppe und Versuchsbedingung

	Experimentalgruppe		Kontrollgruppe	
	Richtige Antworten	Falsche Antworten	Richtige Antworten	Falsche Antworten
Grundschüler	81.27 (16.91)	72.18 (16.63)	77.86 (13.53)	73.08 (14.65)
Jugendliche	81.82 (10.33)	69.52 (14.55)	78.54 (11.35)	68.48 (11.95)
Jüngere Erwachsene	73.24 (8.82)	58.42 (14.54)	74.28 (11.01)	56.29 (14.95)
Ältere Erwachsene	84.18 (10.22)	65.76 (15.86)	80.96 (11.90)	69.08 (14.70)

Standardabweichungen in Klammern.

Die mittleren Gammakorrelationen für die SUs waren in allen Altersgruppen und Versuchsbedingungen von Null verschieden (Grundschüler in der Kontrollbedingung: $p < .01$; alle anderen p 's $< .001$; vgl. Tab. 18). In der ANOVA mit diesen beiden Zwischenssubjektfaktoren zeigte sich ein signifikanter Haupteffekt des Faktors Altersgruppe ($F(3,151) = 2.93; p < .05; \eta^2 = .06$). In den Post-hoc-Tests waren die mittleren Korrelationen der Grundschüler (.28) signifikant niedriger als die der älteren Erwachsenen (.44). Weiterhin wurde der Haupteffekt des Faktors Versuchsbedingung signifikant ($F(1,151) = 4.75; p < .05; \eta^2 = .03$). Teilnehmer mit Strategietraining (.41) erreichten höhere Gammakorrelationen als die in der Kontrollgruppe (.33).

6.5.3.4 Urteile zu „Recollection“ oder Vertrautheit (RFN-Urteile)

Zusätzlich wurden die RFN-Urteile der Teilnehmer analysiert. Dazu wurde der Prozentsatz, zu dem die jeweilige Antwortoption ausgewählt wurde, verglichen (s. Tab. 20). Eine ANOVA mit RFN-Urteil als Innersubjektfaktor sowie Altersgruppe und Versuchsbedingung als Zwischensubjektfaktoren fand einen signifikanten Haupteffekt des Faktors RFN-Urteil ($F(1,152) = 29.99; p < .001; \eta^2 = .17$). Paarweise Kontraste zeigten, dass die Teilnehmer sich deutlich seltener für die Option „Vertrautheit“ (22.40%) als für „Recollection“ (39.82%) und „Keine Erinnerung“ (37.78%) entschieden. Außerdem ergab sich eine signifikante Interaktion zwischen den Faktoren RFN-Urteil und Versuchsbedingung ($F(1,152) = 4.12; p < .05; \eta^2 = .03$). Getrennte ANOVAs mit Versuchsbedingung als Innersubjektfaktor für jedes der drei Urteile zeigten nur für „Vertrautheit“ ($F(1,158) = 12.63; p < .01; \eta^2 = .07$) und „Keine Erinnerung“ ($F(1,158) = 5.54; p < .05; \eta^2 = .03$) signifikante Effekte: Die Teilnehmer in der Experimentalbedingung wählten die Option „Vertrautheit“ (18.99%) weniger oft und die Option „Keine Erinnerung“ (41.42%) öfter als die Teilnehmer der Kontrollgruppe (25.81 bzw. 34.15%).

Tabelle 20: Gewähltes RFN-Urteil in Prozent in Abhängigkeit von Altersgruppe und Versuchsbedingung

	Recollection	Vertrautheit	Keine Erinnerung
Grundschüler			
Experimentalgruppe	35.36 (22.73)	18.93 (11.37)	45.71 (21.04)
Kontrollgruppe	41.61 (24.40)	26.70 (15.36)	31.70 (25.07)
Jugendliche			
Experimentalgruppe	41.93 (18.14)	20.34 (11.17)	37.73 (20.07)
Kontrollgruppe	40.94 (19.38)	30.43 (15.70)	28.64 (17.76)
Jüngere Erwachsene			
Experimentalgruppe	40.75 (22.49)	21.58 (10.72)	37.67 (17.16)
Kontrollgruppe	35.20 (16.23)	24.43 (10.36)	40.37 (16.71)
Ältere Erwachsene			
Experimentalgruppe	40.35 (21.97)	15.10 (8.62)	44.55 (21.69)
Kontrollgruppe	42.41 (20.63)	21.70 (11.24)	35.89 (15.97)

Standardabweichungen in Klammern.

6.5.3.5 Spontaner versus instruierter Strategiegebrauch

Zur Überprüfung der Wirksamkeit der Intervention wurde verglichen, inwiefern sich der Gebrauch einer „visual-imagery“-Strategie zwischen den beiden Versuchsgruppen unterschied (s. Tab. 21). Eine ANOVA mit Altersgruppe und Versuchsbedingung als Zwischensubjektfaktoren ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,154) = 10.97$; $p < .001$; $\eta^2 = .18$). Hier zeigten die Post-hoc-Tests, dass die jüngeren Erwachsenen (82.50%) deutlich häufiger die Strategie anwendeten als die übrigen drei Altersgruppen (Grundschüler: 47.50%; Jugendliche: 52.63%; ältere Erwachsene: 55.00%). Weiterhin wurde ein signifikanter Haupteffekt des Faktors Versuchsbedingung gefunden ($F(1,154) = 221.46$; $p < .001$; $\eta^2 = .63$): Die Teilnehmer in der Experimentalbedingung (95.00%) berichteten häufiger, die „visual-imagery“-Strategie angewendet zu haben, als diejenigen in der Kontrollgruppe (23.82%). Außerdem ergab sich eine signifikante Interaktion zwischen beiden Faktoren ($F(3,154) = 9.38$; $p < .001$; $\eta^2 = .16$). Separate ANOVAs für beide Versuchsbedingungen ergaben für die Probanden in der Experimentalgruppe keinen signifikanten Effekt des Faktors Altersgruppe. Die Teilnehmer waren demnach altersunabhängig in der Lage, die Strategie zu nutzen. Für die Kontrollgruppe wurde der Haupteffekt des Faktors Altersgruppe signifikant ($F(3,74) = 12.30$; $p < .001$; $\eta^2 = .35$). Hier zeigte sich in den Post-hoc-Tests, dass die jüngeren Erwachsenen (65.00%) deutlich häufiger spontan die „visual-imagery“-Strategie nutzten als die anderen Altersgruppen (Grundschüler: 0%; Jugendliche: 5.00%; ältere Erwachsene: 25.00%).

Tabelle 21: Anteil der Teilnehmer in Prozent, die angaben, eine „visual-imagery“-Strategie genutzt zu haben

	Experimentalgruppe	Kontrollgruppe
Grundschüler	95.00 (22.36)	0.00 (0.00)
Jugendliche	100.00 (0.00)	5.26 (22.94)
Jüngere Erwachsene	100.00 (0.00)	65.00 (58.94)
Ältere Erwachsene	85.00 (36.64)	25.00 (44.43)

Standardabweichungen in Klammern

6.5.4 Diskussion

Das Ziel dieser Studie lag darin, die Überwachungsleistungen in den eingeschlossenen Altersgruppen durch ein kurzes Strategietraining zu verbessern. Es wurde erwartet, dass die Instruktion, idiosynkratische Inhalte mit den zu lernenden Items zu verknüpfen, dazu führt, dass sich die Probanden bei der Überwachung weniger auf Aspekte der Vertrautheit mit den Reizen als vielmehr auf „Recollection“-Prozesse beziehen. Da bislang besonders bei Kindern wenig Studien zu den einem Überwachungsurteil zugrunde liegenden Hinweisreizen existieren und besonders bei älteren Erwachsenen Defizite in der Überwachungsleistung auf deren Fokus auf die Vertrautheit mit dem Lernmaterial zurückzuführen sind (Shing et al., 2009; Toth et al., 2011), ist die Untersuchung einer breiten Altersspanne von besonderer Bedeutung.

Die Ergebnisse zeigen, dass die Anpassung der Aufgabenschwierigkeit zwischen den Altersgruppen weitgehend gelungen ist. Es traten weder Decken- noch Bodeneffekte in der Wiedererkennensleistung auf. Außerdem klassifizierten die Teilnehmer der Experimentalgruppe deutlich mehr Items korrekt als die der Kontrollgruppe, was darauf hinweist, dass die Manipulation durch das Strategietraining erfolgreich war. Der Effekt war für die neu kombinierten Wortpaare am deutlichsten ausgeprägt, jedoch auch für die gesamte Wiedererkennensleistung bedeutsam und in Einklang mit vorhandenen Befunden zur positiven Wirkung einer „Interactive-imagery“-Strategie beim Paar-Assoziationslernen (Richardson, 1998; Verhaeghen, Marcoen & Goossens, 1992).

Wie erwartet zeigte sich, dass sowohl die Differenzierungsleistung als auch die Genauigkeit der JOLs durch das Strategietraining grundsätzlich, jedoch nicht in allen Altersgruppen verbessert werden konnte. Während alle Teilnehmer der Experimentalgruppe in ihren JOLs zwischen richtigen und falschen Antworten unterschieden, gelang dies in der Kontrollgruppe nur den Jugendlichen und den jüngeren Erwachsenen. Bei den Jugendlichen wies jedoch die Experimentalgruppe eine deutlich bessere Differenzierungsleistung auf als die Kontrollgruppe. Bezüglich der Genauigkeit ergab sich nur bei den Grundschulern ein bedeutsamer positiver Effekt des Trainings, bei den älteren Erwachsenen wurde eine Verbesserung in den Gammakorrelationen lediglich deskriptiv gefunden. Auch bei den SUs profitierten nur die Grundschüler und die älteren Erwachsenen hinsichtlich ihrer Differenzierungsfähigkeit von dem Training: Hier war der Unterschied in den mittleren SUs zwischen richtigen und falschen Antworten bei der Experimentalgruppe jeweils etwa doppelt so hoch wie bei der Kontrollgruppe. Für die SU-Genauigkeit zeigten sich für alle Altersgruppen signifikante Vorteile der Experimental- gegenüber der Kontrollgruppe.

Das Training wies also die erwarteten positiven Effekte auf die Überwachungsleistung auf. Diese waren bei den Grundschulern und den älteren Erwachsenen am deutlichsten ausgeprägt. Dies entspricht Vorbefunden zu Produktionsdefiziten bei Grundschulern (Schneider, 2015) und älteren Erwachsenen (Naveh-Benjamin, Brav & Levy, 2007). Bei den jüngeren Erwachsenen hatte das Training am wenigsten Auswirkungen sowohl auf die Erinnerungs- als auch auf die Überwachungsleistung. Dies ist wahrscheinlich darauf zurückzuführen, dass die jüngeren Erwachsenen am häufigsten spontan die Strategie anwendeten. Hier ist zu vermuten, dass eine Strategieinstruktion eher bei komplexeren Aufgaben Einfluss nehmen würde (Nietfeld & Schraw, 2002). Außerdem trafen die positiven Effekte des Trainings für die SUs auf eine breitere Altersspanne zu als für die JOLs. Dies deutet darauf hin, dass sowohl JOLs als auch SUs auf Hinweisreizen beruhen, auf die das Training Einfluss nahm. Allerdings scheinen die beiden Maße zum Teil auch auf unterschiedlichen Grundlagen zu beruhen. Da in die SUs mehr Informationen aus dem Abrufprozess eingehen als in die JOLs, ist anzunehmen, dass hier „Recollection“-Prozesse noch mehr Bedeutung haben als bei JOLs.

Als für das Training zugrunde liegender Wirkmechanismus wurden Änderungen in den genutzten Hinweisreizen vermutet, in dem Sinne, dass die Urteile in der Experimentalgruppe weniger auf vertrauthheits- als vielmehr auf „Recollection“-basierten Prozessen beruhen sollten. Diese Annahme konnte in Teilen durch die RFN-Urteile bestätigt werden: In der Experimentalgruppe zeigte sich in allen Altersgruppen die erwartete Abnahme an Vertrauthheitsurteilen. Diese erfolgte jedoch nicht zugunsten der „Recollection“-Urteile, sondern zugunsten der No-Memory-Urteile. Das heißt, die Teilnehmer konnten anscheinend besser zwischen verschiedenen „Recollection“-Zuständen und sicherer zwischen nicht erinnerten und erinnerten Items unterscheiden. Die Zahl der Klassifikationsentscheidungen, bei denen Raten eine Rolle spielte und die den Vertrauthheits-Urteilen zuzuordnen sind, schien also durch das Strategie-Training zurückzugehen. Die Tatsache, dass bezüglich der RFN-Urteile keine Alterseffekte gefunden wurden, macht deutlich, dass es sich wohl lebensspannenübergreifend um ähnliche Wirkmechanismen zu handeln scheint.

Insgesamt gesehen fielen die positiven Effekte des Strategietrainings auf die Überwachungsleistung nicht so hoch aus wie erwartet. Ein Grund dafür könnte das gewählte Paradigma sein: Um den Einfluss von Vertrautheit und Recollection im Design klar zu trennen, wurde eine Wiedererkennenaufgabe gestellt, bei der die Genauigkeit der Überwachung generell geringer ausfällt als bei Aufgaben zum freien Erinnerungsabruf (Buratti & Allwood, 2012), sodass die Vermutung nahe liegt, dass hier auch Recollection und Vertrautheit für die Probanden schwerer zu unterscheiden sind. Weiterhin ist hinreichend bekannt, dass verzögert

erfasste JOLs an sich schon recht genau ausfallen, da hier auf das Langzeitgedächtnis, also vermehrt auf „Recollection“-Prozesse rekurriert wird (Nelson & Dunlosky, 1991). Dennoch fiel die Entscheidung in dieser Studie für verzögerte JOLs, da das Strategietraining zusätzlich zu bekannten Einflussfaktoren einen positiven Effekt auf die Qualität von Überwachungsurteilen belegen sollte. Es ist zu erwarten, dass noch deutlichere Auswirkungen eines Strategietrainings auf die Überwachungsleistung gezeigt werden können, wenn die Intervention über einen längeren Zeitraum oder – besonders bei jüngeren Erwachsenen – auch bei komplexeren Materialien stattfindet. So steht eine Überprüfung in alltagsnäheren Kontexten noch aus. Dennoch konnte diese Studie erfolgreich zeigen, dass schon kurze und ökonomische Maßnahmen zur Verbesserung der Qualität der verwendeten Hinweisreize sich positiv auf die metakognitive Überwachung auswirken.

6.6 Studie 5: Sequenzielle Verknüpfung von Überwachungs- und Kontrollprozessen

6.6.1 Herleitung der Fragestellung

Die beiden letzten Studien dieser Arbeit erweitern den Fokus dahingehend, dass hier nicht verschiedene Einflussfaktoren auf Überwachungsprozesse – wie die Komplexität des Lernmaterials, das Vorwissen oder ein Strategietraining – untersucht werden, sondern Kontrollprozesse einbezogen werden. Ziel der Studien 5 und 6 ist es, zu untersuchen, wie Überwachungs- und Kontrollprozesse sich in verschiedenen Lebensaltern wechselseitig beeinflussen.

Im Design der Studie 5 werden dabei mögliche Wirkrichtungen in zwei Lerndurchgängen getrennt untersucht. Der erste Durchgang befasst sich dabei mit der Frage, inwiefern Personen verschiedenen Alters in der Lage sind, Informationen aus Überwachungsvorgängen für die Kontrolle des Lernens, hier gemessen durch die Lernzeit, zu nutzen („Monitoring-affects-control“-Hypothese, auch MC-Modell; Nelson & Leonesio, 1988). Diese intuitiv eingängige Annahme, dass also beispielsweise die in hohen JOLs zum Ausdruck kommende Überzeugung, dass ein Item zu einem späteren Zeitpunkt noch erinnert werden kann, beim erneuten Lernen zu kurzen Lernzeiten führt, konnte mehrheitlich experimentell bestätigt werden (vgl. Son & Metcalfe, 2000). Auch in entwicklungspsychologischen Studien zeigte sich, dass bereits Kinder zu Beginn der Grundschulzeit Items mit niedrigeren JOLs mehr Lernzeit widmen; der Zusammenhang steigert sich dann noch im Verlauf der Grundschulzeit (Destan et al., 2014; Lockl & Schneider, 2003). Bis ins jüngere Erwachsenenalter scheint dieser Zusammenhang stabil zu sein und tendenziell noch enger zu werden (Koriat et al., 2006; Kurtz & Schneider, 1988; Son & Metcalfe, 2000). Für das höhere Erwachsenenalter existieren bislang z.T. widersprüchliche Befunde in Bezug auf das MC-Modell. Zwar weist diese Altersgruppe bekanntlich gute Überwachungsleistungen auf (vgl. Dunlosky & Metcalfe, 2009), doch wurde in einigen Studien ein geringerer Zusammenhang zwischen Überwachung und Kontrolle als bei jüngeren Erwachsenen gefunden (Dunlosky & Connor, 1997; Froger, Sacher, Gaudouen, Isingrini & Taconnat, 2011). Dies könnte daran liegen, dass ältere Erwachsene effektive Strategien weniger häufig spontan einsetzen als jüngere (Dunlosky & Connor, 1997; vgl. auch Studie 4). Unter günstigen Bedingungen, wie z.B. nach einer ausführlichen Strategieinstruktion (Froger, Bouazzaoui, Isingrini & Taconnat, 2012) oder nach einer dem Experiment vorgeschalteten Phase, in der sämtliche Wortpaare schon einmal gelernt werden konnten

(Hines, Hertzog & Touron, 2012), erreichten die älteren Erwachsenen dagegen vergleichbare Ergebnisse wie die jüngeren.

Auch wenn für eine Reihe von Altersgruppen erste Belege für das MC-Modell gefunden werden konnten, steht auch hier ein direkter Vergleich verschiedener Altersgruppen unter identischen Versuchsbedingungen aus. Deshalb soll in Lernphase 1 des folgenden Experiments mit dem Design von Lockl und Schneider (2003) für eine breite Altersgruppe (Grundschüler bis ältere Erwachsene) untersucht werden, inwiefern sich das Ausmaß des Zusammenhangs zwischen Überwachung und Kontrolle in den verschiedenen Altersstufen unterscheidet. Es wird erwartet, dass bereits bei Grundschulern Evidenz für das MC-Modell gefunden werden kann, der Zusammenhang jedoch im Verlauf des Jugendalters bis zum jüngeren Erwachsenenalter noch etwas enger wird. Da das Paradigma – abgesehen von der Anpassung der grundsätzlichen Aufgabenschwierigkeit – keine Maßnahmen zur Erleichterung der kognitiven Verarbeitung (z.B. eine Strategieinstruktion) vorsieht, wird erwartet, dass der Zusammenhang zwischen Überwachung und Kontrolle bei den älteren Erwachsenen wieder etwas geringer ausfällt als bei den jüngeren.

Noch weniger empirische Belege existieren bislang für einen umgekehrten Zusammenhang zwischen Kontroll- und Überwachungsprozessen (CM-Modell), der im zweiten Lerndurchgang dieses Experiments untersucht werden soll. Seit der Arbeit von Koriat et al. (2006) wird erstmals diskutiert, dass bei datenorientierten Lernsituationen (im Gegensatz zum zielorientierten Lernen) Informationen aus Kontrollprozessen auch als Hinweisreize für die Höhe der Überwachungsurteile genutzt werden können. Je kürzer ein Item demnach gelernt wurde, desto höher sollte auch das entsprechende Urteil ausfallen (Easily-Learned-Easily-Remembered-Heuristik; Koriat et al., 2009a). In Studien mit jüngeren Erwachsenen fielen diese Zusammenhänge zwischen Lernzeit und JOLs statistisch bedeutsam, jedoch etwas niedriger aus als die Korrelationen im MC-Modell (Koriat et al., 2006). Im Entwicklungsverlauf wurden Belege für das CM-Modell etwa ab der dritten Klasse gefunden (Koriat et al., 2009b). Die Stärke des Zusammenhangs scheint sich bis zum Jugendalter weiter zu steigern (Hoffmann-Biencourt et al., 2010). Diese ersten Studien sprechen dafür, dass das Zustandekommen von Überwachungsurteilen im Verlauf der Entwicklung unterschiedlichen Hinweisreizen unterliegt, und dass wahrscheinlich in der Zeit zwischen dem mittleren Grundschul- und dem jüngeren Erwachsenenalter zunehmend variable und damit auch effektivere Grundlagen für die Urteile herangezogen werden. Die bisherige empirische Datenlage für das CM-Modell, besonders was altersbedingte Unterschiede betrifft, ist jedoch noch als unzureichend anzusehen. So gibt es bislang noch keine entsprechenden Befunde für das höhere Erwachsenenalter. Da Studien, die

Kontrollmaße erfassen, mehrheitlich zum Schluss kommen, dass ältere Erwachsene aufgrund von Defiziten in Exekutivfunktionen in der Steuerung ihres Lernvorgangs schlechter abschneiden als jüngere Erwachsene (z.B. Murphy, Schmitt, Caruso & Sanders, 1987; Pansky, Goldsmith, Koriat & Pearlman-Avni, 2009; Souchay & Isingrini, 2004), ist davon auszugehen, dass auch der Zusammenhang zwischen Kontrolle und Überwachung in dieser Altersgruppe geringer ausfällt.

Insgesamt werden also für das CM-Modell ähnliche, wenngleich etwas deutlicher ausgeprägte Altersverläufe wie beim MC-Modell erwartet, im Sinne einer Zunahme des Zusammenhangs zwischen Kontrolle und Überwachung bis zum jüngeren Erwachsenenalter und eine Verringerung vom jüngeren zum höheren Erwachsenenalter.

Zusammenfassend soll also im folgenden Experiment im ersten Lerndurchgang erstmals ein lebensspannenübergreifender Vergleich für das MC-Modell erfolgen; es wird daher eine zielorientierte Lernsituation geschaffen. Im zweiten Lerndurchgang soll datenorientiertes Lernen stattfinden. Damit sollen Belege für das CM-Modell in einer breiten Altersspanne erbracht werden. Der Fokus des Experiments lag also auf dem Zusammenhang zwischen JOLs und Lernzeit. Die SUs wurden zusätzlich erhoben, um die Leistungen mit denen in den anderen Studien zu vergleichen.

6.6.2 Methodik

6.6.2.1 Design

Beim Versuchsplan handelte es sich um ein 4 (Altersgruppen: Kinder, Jugendliche, jüngere und ältere Erwachsene) x 2 (Item-Schwierigkeit: Zusammenhängend vs. unzusammenhängend) faktorielles Design. Die Altersgruppe wurde als Zwischensubjektfaktor, die Item-Schwierigkeit als Innersubjektfaktor erfasst. Die Erinnerungsleistung, die JOLs, die SUs sowie die selbst gewählte Lernzeit bildeten die abhängigen Variablen.

6.6.2.2 Stichprobe

Es nahmen insgesamt $N = 112$ Personen, darunter 28 Drittklässler (12 männlich, 16 weiblich), 28 Jugendliche der siebten und achten Jahrgangsstufe (21 männlich, 7 weiblich), 28 jüngere Erwachsene (11 männlich, 17 weiblich) sowie 28 ältere Erwachsene (11 männlich, 17 weiblich) teil. Die Grundschüler waren im Mittel 9.00 Jahre alt ($SD = 0.38$; Altersspanne: 8 bis 10 Jahre),

die Jugendlichen 13.25 Jahre ($SD = 1.04$; Altersspanne: 11 bis 15 Jahre), die jüngeren Erwachsenen 22.00 Jahre ($SD = 2.29$; Altersspanne: 18 bis 27 Jahre) und die älteren Erwachsenen 68.39 Jahre ($SD = 5.63$; Altersspanne: 59 bis 76 Jahre). Die Probanden wurden gleichermaßen wie bei den bisher vorgestellten Studien rekrutiert.

Die Leistungen im Zahlennachsprechen (Petermann & Petermann, 2010; von Aster et al., 2006) unterschieden sich leicht zwischen den Altersgruppen. Eine ANOVA mit Altersgruppe als Zwischensubjektfaktor ergab einen signifikanten Haupteffekt dieses Faktors ($F(3,108) = 3.04$; $p < .05$; $\eta^2 = .08$). Post-hoc konnten jedoch keine bedeutsamen Gruppenunterschiede identifiziert werden. Die Normwerte der einzelnen Altersgruppen lauteten wie folgt: Grundschüler: $M = 9.64$ ($SD = 2.21$); Jugendliche: $M = 10.96$ ($SD = 2.67$); jüngere Erwachsene: $M = 11.78$ ($SD = 2.91$); ältere Erwachsene: $M = 11.43$ ($SD = 3.43$).

Im Wortschatztest (Petermann & Petermann, 2010; von Aster et al., 2006) wurde bei der ANOVA mit Altersgruppe als Zwischensubjektfaktor ebenfalls ein signifikanter Haupteffekt gefunden ($F(3,107) = 3.13$; $p < .05$; $\eta^2 = .08$); im Post-hoc-Test wurden keine signifikanten Unterschiede mehr zwischen den Altersgruppen sichtbar. Für den Wortschatz lauteten die Normwerte für die einzelnen Altersgruppen folgendermaßen: Grundschüler: $M = 8.39$ ($SD = 2.31$); Jugendliche: $M = 9.96$ ($SD = 3.40$); jüngere Erwachsene: $M = 10.07$ ($SD = 2.14$); ältere Erwachsene: $M = 10.37$ ($SD = 2.54$).

Die mittlere Schul- und Ausbildungszeit unterschied sich nicht zwischen den jüngeren und den älteren Erwachsenen (jüngere Erwachsene: $M = 14.65$ Jahre; $SD = 2.10$; ältere Erwachsene: $M = 13.22$ Jahre; $SD = 3.48$). Die jüngeren Erwachsenen schätzten ihren Gesundheitszustand positiver ein als die älteren ($M = 1.71$; $SD = 0.71$ vs. $M = 2.61$; $SD = 0.76$) und gaben an, weniger Medikamente pro Tag einzunehmen ($M = 0.45$; $SD = 0.92$ vs. $M = 2.41$; $SD = 2.16$).

6.6.2.3 Versuchsmaterial

Wortpaare. Das Lernmaterial bestand aus zweisilbigen konkreten Substantiven. Um die Schwierigkeit zu variieren, war die Hälfte der Wortpaare semantisch assoziiert (z.B. Käfer – Spinne), die andere Hälfte nicht assoziiert (z.B. Socke – Bagger). Die Itemliste für Kinder und ältere Erwachsene umfasste 60 Wortpaare, von denen 36 randomisiert dem ersten und 24 dem zweiten Lerndurchgang zugewiesen wurden. Bei den Jugendlichen handelte es sich um 90 Paare (54 im ersten und 36 im zweiten Lerndurchgang) und bei den jüngeren Erwachsenen um 120 Paare (72 im ersten und 48 im zweiten Lerndurchgang; s. Anhang A 8). Jede Altersgruppe

bekam zusätzlich vier Übungspaare dargeboten. Die Auswahl der Items wurde basierend auf eine Pilotierungsstudie mit $N = 20$ Teilnehmern getroffen.

Wie im vorherigen Experiment fanden die JOL- und die SU-Abfrage am Computer mithilfe der Thermometerskala statt.

Zusatzvariablen. Als Zusatzvariablen wurden wie in den bereits vorgestellten Studien die Leistungen im Zahlennachsprechen und im Wortschatz (Petermann & Petermann, 2010; von Aster et al., 2006) erfasst. Wieder kam bei den jüngeren und älteren Erwachsenen der Fragebogen zu demographischen Angaben und zum Gesundheitszustand zum Einsatz.

6.6.2.4 Studienablauf

Es handelte sich um Einzeltestungen in den Schulen oder im Labor, die bei den minderjährigen Teilnehmern mit dem schriftlichen Einverständnis der Eltern durchgeführt wurden.

Die Hälfte der Probanden bearbeitete zunächst den Lerndurchgang 1, die andere Hälfte zunächst den Lerndurchgang 2, um Reihenfolgeeffekte auszugleichen.

Der erste Durchgang wurde analog zur Studie von Lockl und Schneider (2003) konzipiert. Zunächst wurden die Wortpaare in der Lernphase mit festgelegten Lernzeiten (Grundschüler und ältere Erwachsene: 6 Sekunden; Jugendliche: 4 Sekunden; jüngere Erwachsene: 3 Sekunden) präsentiert. Die Darbietungsdauer wurde auf Grundlage der Ergebnisse der Pilotierungsstudie festgelegt. Danach wurden die JOLs abgefragt. Anschließend hatten die Teilnehmer die Möglichkeit, sich die Wortpaare noch einmal in ihrem eigenen Lerntempo einzuprägen. Sie erhielten die Instruktion, die Items so lange zu lernen, wie sie benötigten, um sie in etwa 15 Minuten zu erinnern. Im Intervall zwischen Lernen und Erinnern wurde der Test zum Zahlennachsprechen bearbeitet. In der Erinnerungsphase wurde wiederum das linke Wort jedes Paares präsentiert mit der Aufgabe, das rechte zu ergänzen und die Sicherheit der Antwort einzuschätzen. Die Teilnehmer sollten raten, falls ihnen keine Antwort einfiel.

Der Ablauf des zweiten Lerndurchgangs richtete sich im Wesentlichen nach Koriat et al. (2009b). Die Teilnehmer sollten nun als erstes die Wortpaare im selbst gewählten Tempo lernen, wieder mit der Vorgabe, so lange zu lernen, dass sie in etwa 15 Minuten das rechte Wort ergänzen konnten, wenn das linke dargeboten wurde. Danach erfolgten die JOL-Abfrage sowie

anschließend der Wortschatztest. In der Erinnerungsphase sollte dann das rechte Wort ergänzt bzw. gegebenenfalls erraten werden. Außerdem sollten die SUs abgegeben werden.

Die Belohnungen für die Teilnehmer waren identisch zu denen der bisher vorgestellten Studien.

6.6.2.5 Überblick über die statistischen Auswertungen

Um die Übersichtlichkeit und die Interpretierbarkeit zu gewährleisten, wurden die beiden Lerndurchgänge getrennt ausgewertet. Die Analysen zur Erinnerungsleistung sowie zur Güte der JOLs und SUs entsprechen denen der vorherigen Studien; d.h. es wurden ANOVAs mit dem Zwischensubjektfaktor Altersgruppe berechnet. Bei der Auswertung der Erinnerungsleistung floss zusätzlich der Innersubjektfaktor Itemschwierigkeit ein. Um den Einfluss der JOLs auf die selbst gewählte Lernzeit (Lerndurchgang 1; MC-Modell) bzw. den Einfluss der Lernzeit auf die Höhe der JOLs (Lerndurchgang 2; CM-Modell) zu bestimmen, wurden jeweils Pearson-Korrelationen zwischen den JOLs und der Lernzeit berechnet. In Lerndurchgang 2 wurde zusätzlich der Mediansplit der Lernzeit gebildet und verglichen, inwiefern sich die JOLs für lange vs. kurz gelernte Items unterschieden (vgl. Koriat et al., 2006, 2009a). Eine vergleichbare Auswertung auf die JOLs bezogen konnte für den ersten Lerndurchgang nicht erfolgen, da sich hier eine u-förmige Verteilung ergab und somit kein Mediansplit möglich war. Schließlich wurde für den zweiten Lerndurchgang analysiert, inwiefern die Lernzeit tatsächlich einen validen Hinweisreiz für die objektive Erinnerungsleistung darstellt. Dazu wurde mithilfe von Gammakorrelationen der Zusammenhang zwischen beiden Variablen für jeden Teilnehmer berechnet (vgl. Hoffmann et al., 2010; Koriat et al., 2009b). Auch diese Ergebnisse wurden durch eine weitere Analyse ergänzt: Hier wurde ebenfalls die Erinnerungsleistung für Items mit Lernzeiten über bzw. unter dem Median verglichen (vgl. Koriat et al., 2009b). Um signifikante Unterschiede in *t*-Tests miteinander zu vergleichen, wurde Cohens *d* herangezogen. Effektstärken zwischen .2 und .4 stehen dabei für kleine, zwischen .5 und .7 für mittlere und über .8 für große Effekte (Cohen, 1988).

6.6.3 Ergebnisse

6.6.3.1 Lerndurchgang 1

6.6.3.1.1 Erinnerungsleistung

Die Tabelle 22 zeigt den prozentualen Anteil richtiger Antworten im ersten Lerndurchgang. Die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Itemschwierigkeit als Innersubjektfaktor ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,108) = 8.85; p < .001; \eta^2 = .20$). Im Post-hoc-Test zeigte sich, dass die jüngeren Erwachsenen (85.02%) besser abschnitten als die Teilnehmer der drei anderen Altersgruppen (Grundschüler: 62.40%; Jugendliche: 63.62%; ältere Erwachsene: 63.49%). Weiterhin wurde der Haupteffekt des Faktors Itemschwierigkeit ($F(1,108) = 258.89; p < .001; \eta^2 = .71$) sowie die Interaktion zwischen den beiden Faktoren signifikant ($F(3,108) = 10.72; p < .001; \eta^2 = .22$). Insgesamt lag die Erinnerungsleistung bei assoziativen Wortpaaren (82.42%) deutlich höher als bei nicht assoziativen (54.85%). *T*-Tests zeigten, dass alle Altersgruppen bessere Leistungen bei leichten als bei schwierigen Wortpaaren aufwiesen (alle *p*'s < .001). Die Wechselwirkung verdeutlicht, dass der Einfluss der Itemschwierigkeit bei den jüngeren Erwachsenen weniger stark ausgeprägt war als bei den anderen Altersgruppen (Grundschüler: Cohens $d = 2.07$; Jugendliche: $d = 1.31$; jüngere Erwachsene: $d = 0.64$; ältere Erwachsene: $d = 1.25$).

Tabelle 22: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe und Itemschwierigkeit im Lerndurchgang 1

	Gesamt	Einfache Items	Schwierige Items
Grundschüler	62.40 (15.24)	80.36 (13.05)	44.44 (20.73)
Jugendliche	63.62 (20.89)	78.57 (17.32)	48.68 (27.12)
Jüngere Erwachsene	85.02 (16.44)	90.77 (12.58)	79.27 (22.24)
Ältere Erwachsene	63.49 (23.98)	79.96 (18.48)	47.02 (32.48)

Standardabweichungen in Klammern.

6.6.3.1.2 Judgments of Learning

Die Güte der prospektiven Überwachungsleistung wurde wieder durch die Kontrastierung der mittleren JOLs vor richtigen bzw. falschen Antworten bestimmt (vgl. Tab. 23). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Korrektheit der Antwort als Innersubjektfaktor

ergab einen signifikanten Haupteffekt der Korrektheit der Antwort ($F(1,105) = 323.64$; $p < .001$; $\eta^2 = .75$); d.h. die Probanden gaben im Mittel höhere JOLs vor richtigen (63.65) als vor falschen Antworten (32.24). Außerdem wurde die Interaktion zwischen beiden Faktoren signifikant ($F(3,105) = 3.32$; $p < .05$; $\eta^2 = .08$). Bei univariaten ANOVAs, nach richtigen und falschen Antworten getrennt, wurde für die JOLs vor richtigen Antworten ein signifikanter Haupteffekt des Faktors Altersgruppe gefunden (beide p 's $< .001$). Im Post-hoc-Test zeigte sich, dass die älteren Erwachsenen (67.60) etwas höhere Werte aufwiesen als die jüngeren Erwachsenen (57.47; $p < .05$). Für die JOLs vor falschen Antworten ergab sich kein signifikanter Haupteffekt. Bezüglich der Gammakorrelationen zwischen JOLs und Erinnerungsleistung zeigte sich in der ANOVA mit Altersgruppe als Zwischensubjektfaktor kein signifikanter Effekt. Alle Korrelationen waren signifikant von Null verschieden (alle p 's $< .001$; s. Tab. 23).

Tabelle 23: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe; mittlere Gammakorrelationen für JOLs in Abhängigkeit der Altersgruppe im Lerndurchgang 1

	Mittlere JOLs		Gamma- korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	65.74 (14.16)	33.40 (16.66)	.55 (.17)
Jugendliche	64.88 (11.39)	34.73 (15.70)	.49 (.25)
Jüngere Erwachsene	57.35 (14.01)	33.47 (23.10)	.41 (.39)
Ältere Erwachsene	66.64 (13.17)	27.35 (16.04)	.60 (.24)

Standardabweichungen in Klammern.

6.6.3.1.3 Sicherheitsurteile

Bezüglich der Differenzierungsleistung in den SUs ergab die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Korrektheit der Antwort als Innersubjektfaktor ebenfalls einen signifikanten Haupteffekt des Faktors Korrektheit der Antwort ($F(1,105) = 1128.33$; $p < .001$; $\eta^2 = .92$) mit deutlich höheren SUs nach richtigen (92.73) als nach falschen Antworten (28.79; s. Tab. 24). Auch die Interaktion zwischen beiden Faktoren wurde signifikant ($F(3,105) = 3.14$; $p < .05$; $\eta^2 = .08$). Die getrennte Auswertung nach richtigen und falschen Antworten ergab für die richtigen Antworten keinen Effekt des Faktors Altersgruppe. Für die falschen Antworten verfehlte der Effekt knapp das Signifikanzniveau ($F(3,105) = 2.52$; $p < .10$; $\eta^2 = .07$). Der größte

Unterschied zeigte sich zwischen den jüngeren Erwachsenen (35.86) und den Grundschulern (22.49; $p = .101$). Alle mittleren Gammakorrelationen für die SUs waren bedeutsam von Null verschieden (alle p 's $< .001$; s. Tab. 24). In der dazugehörigen ANOVA wurde der Effekt des Zwischensubjekt-Faktors Altersgruppe signifikant ($F(3,105) = 3.15$; $p < .05$; $\eta^2 = .08$). Post-hoc zeigte sich, dass die Grundschüler (.96) marginal höhere Werte aufwiesen als die jüngeren Erwachsenen (.75; $p < .1$).

Tabelle 24: Mittlere SUs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe; mittlere Gammakorrelationen für SUs in Abhängigkeit der Altersgruppe im Lerndurchgang 1

	Mittlere SUs		Gamma- korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	94.30 (9.08)	22.49 (15.88)	.96 (.05)
Jugendliche	92.71 (7.45)	25.60 (15.21)	.94 (.10)
Jüngere Erwachsene	92.86 (8.15)	35.86 (26.84)	.75 (.54)
Ältere Erwachsene	91.03 (8.41)	31.19 (18.18)	.87 (.13)

Standardabweichungen in Klammern.

6.6.3.1.4 Judgments of Learning als Hinweisreize für die Lernanstrengung

Der Zusammenhang zwischen JOLs und Lernzeit wurde mit Pearson-Korrelationen zwischen beiden Variablen für jeden Teilnehmer bestimmt (s. Tab. 25). Sämtliche Korrelationen waren von Null verschieden (alle p 's $< .001$). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor fand einen signifikanten Haupteffekt ($F(3,108) = 5.05$; $p < .01$; $\eta^2 = .12$). In den Post-hoc-Tests zeigte sich, dass die jüngeren Erwachsenen (-.53) im Mittel höhere Korrelationen aufwiesen als die Jugendlichen (-.27).

Tabelle 25: Mittlere Pearson-Korrelationen zwischen JOLs und Lernzeiten in Abhängigkeit der Altersgruppe und Lerndurchgang

	Lerndurchgang 1	Lerndurchgang 2
Grundschüler	-.34 (.26)	-.05 (.29)
Jugendliche	-.27 (.24)	-.04 (.18)
Jüngere Erwachsene	-.53 (.19)	-.15 (.22)
Ältere Erwachsene	-.39 (.32)	-.11 (.24)

Standardabweichungen in Klammern.

6.6.3.2 Lerndurchgang 2

6.6.3.2.1 Erinnerungsleistung

Für den Lerndurchgang 2 (s. Tab. 26) ergab sich in der ANOVA mit dem Zwischensubjektfaktor Altersgruppe und dem Innersubjektfaktor Itemschwierigkeit ein signifikanter Haupteffekt des Faktors Altersgruppe ($F(3,108) = 13.94; p < .001; \eta^2 = .28$). Die Post-hoc-Tests zeigten, dass die jüngeren Erwachsenen (79.54%) im Mittel eine höhere Erinnerungsleistung aufwiesen als die drei anderen Altersgruppen (Grundschüler: 54.83%, $p < .001$; Jugendliche: 61.11%, $p < .001$; ältere Erwachsene: 52.42%, $p < .001$). Weiterhin wurden ein signifikanter Haupteffekt des Faktors Itemschwierigkeit ($F(1,108) = 278.82; p < .001; \eta^2 = .72$) sowie eine Interaktion zwischen beiden Faktoren gefunden ($F(3,108) = 13.67; p < .001; \eta^2 = .28$). *T*-Tests für die einzelnen Altersgruppen ergaben, dass jeweils deutlich mehr einfache als schwierige Items erinnert wurden (alle p 's $< .001$). Dieser Unterschied war jedoch für die Grundschüler ($d = 1.93$) und die älteren Erwachsenen ($d = 1.44$) stärker ausgeprägt als für die Jugendlichen ($d = 1.00$) und die jüngeren Erwachsenen ($d = 0.67$).

Tabelle 26: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe und Itemschwierigkeit im Lerndurchgang 2

	Gesamt	Einfache Items	Schwierige Items
Grundschüler	.46 (.19)	.66 (.21)	.26 (.20)
Jugendliche	.61 (.21)	.72 (.16)	.50 (.28)
Jüngere Erwachsene	.80 (.19)	.86 (.16)	.73 (.24)
Ältere Erwachsene	.53 (.24)	.71 (.24)	.33 (.28)

Standardabweichungen in Klammern.

6.6.3.2.2 Judgments of Learning

Die ANOVA zur Analyse der Differenzierungsleistung in den JOLs (s. Tab. 27) mit dem Zwischensubjektfaktor Altersgruppe und dem Innersubjektfaktor Korrektheit der Antwort fand lediglich einen signifikanten Effekt der Korrektheit der Antwort ($F(1,105) = 417.27$; $p < .001$; $\eta^2 = .80$): Vor richtigen Antworten (85.22) wurden höhere JOLs abgegeben als vor falschen (37.29). Die mittleren Gammakorrelationen (s. Tab. 27) waren für die JOLs im zweiten Lerndurchgang ebenfalls alle deutlich von Null verschieden (alle p 's $< .001$). Auch hier ergab sich in der dazugehörigen ANOVA kein Haupteffekt des Faktors Altersgruppe.

Tabelle 27: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe; mittlere Gammakorrelationen für JOLs in Abhängigkeit der Altersgruppe im Lerndurchgang 2

	Mittlere JOLs		Gamma- korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	85.00 (18.75)	35.29 (19.77)	.81 (.29)
Jugendliche	84.63 (12.16)	38.17 (15.76)	.81 (.14)
Jüngere Erwachsene	56.78 (8.71)	39.87 (25.48)	.66 (.47)
Ältere Erwachsene	84.48 (15.21)	35.82 (18.97)	.76 (.27)

Standardabweichungen in Klammern.

6.6.3.2.3 Sicherheitsurteile

Auch für die Differenzierungsleistung in den SUs (s. Tab. 28) wurde in der ANOVA mit dem Zwischensubjektfaktor Altersgruppe und dem Innersubjektfaktor Korrektheit der Antwort lediglich der Haupteffekt der Korrektheit der Antwort signifikant ($F(1,105) = 895.15$; $p < .001$; $\eta^2 = .90$). Die Teilnehmer gaben im Mittel deutlich höhere SUs nach richtigen (90.38) als nach falschen Antworten (28.08) ab. Bezüglich der Gammakorrelationen (s. Tab. 28) ergab die ANOVA mit Zwischensubjektfaktor Altersgruppe ebenfalls keinen signifikanten Haupteffekt. In jeder Altersgruppe waren die gemittelten Korrelationen statistisch bedeutsam von Null verschieden (alle p 's $< .001$).

Tabelle 28: Mittlere SUs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe; mittlere Gammakorrelationen für SUs in Abhängigkeit der Altersgruppe im Lerndurchgang 2

	Mittlere SUs		Gamma- korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	88.98 (13.57)	23.80 (16.01)	.90 (.13)
Jugendliche	91.15 (8.75)	24.09 (16.20)	.93 (.09)
Jüngere Erwachsene	93.21 (5.59)	33.39 (26.55)	.80 (.43)
Ältere Erwachsene	88.19 (13.90)	31.04 (19.19)	.87 (.16)

Standardabweichungen in Klammern.

6.6.3.2.4 Lernzeit als Hinweisreiz für die Judgments of Learning

Um beurteilen zu können, inwiefern die Lernzeit als Hinweisreiz für die JOLs genutzt wurde, wurden auch für den zweiten Lerndurchgang zunächst Pearson-Korrelationen zwischen Lernzeit und JOLs für jeden Teilnehmer berechnet (s. Tab. 25). Bei den jüngeren ($p < .01$) und den älteren Erwachsenen ($p < .05$) waren die mittleren Korrelationen signifikant von Null verschieden. Eine ANOVA mit Altersgruppe als Zwischensubjektfaktor ergab keinen signifikanten Haupteffekt. Um diese Ergebnisse weiter zu untermauern, wurden für jeden Probanden zusätzlich die mittleren JOLs über und unter dem individuellen Median der Lernzeit berechnet (s. Tab. 29). Bei der ANOVA mit Altersgruppe als Zwischensubjektfaktor und Lernzeit (über vs. unter dem Median) wurde ein signifikanter Haupteffekt des Faktors Altersgruppe ($F(3,108) = 5.54; p < .01; \eta^2 = .13$) gefunden. Im Post-hoc-Test zeigte sich, dass die jüngeren Erwachsenen (75.44) höhere JOLs abgaben als die Grundschüler (60.34; $p < .01$) und die älteren Erwachsenen (62.15; $p < .05$). Außerdem wurde der Haupteffekt des Faktors Lernzeit signifikant ($F(3,108) = 21.24; p < .001; \eta^2 = .16$). Im Mittel wurden höhere JOLs für kürzer gelernte (69.41) als für länger gelernte Items (63.11) abgegeben.

Tabelle 29: Mittlere JOLs und Erinnerungsleistung in Prozent für Items mit einer Lernzeit über bzw. unter dem Median in Abhängigkeit der Altersgruppe im Lerndurchgang 2

	Mittlere JOLs		Mittlere Erinnerungsleistung	
	Unter dem Median	Über dem Median	Unter dem Median	Über dem Median
Grundschüler	62.24 (18.13)	58.43 (18.02)	49.40 (23.01)	42.26 (21.63)
Jugendliche	68.94 (14.02)	65.27 (16.71)	64.91 (19.98)	56.76 (27.26)
Jüngere Erwachsene	78.65 (12.89)	72.23 (12.30)	82.39 (17.74)	76.57 (20.77)
Ältere Erwachsene	67.81 (19.54)	56.49 (20.90)	63.10 (25.10)	41.96 (26.88)

Standardabweichungen in Klammern.

6.6.3.2.5 Validität der Lernzeit als Prädiktor für die Erinnerungsleistung

Um zu beurteilen, ob die Lernzeit einen validen Hinweisreiz für die spätere Erinnerungsleistung darstellt, wurden Gammakorrelationen zwischen Lernzeit und Erinnerungsleistung für jeden Teilnehmer berechnet (vgl. Tab. 30). Die gemittelten Korrelationen waren für jede Altersgruppe signifikant von Null verschieden (Grundschüler und Jugendliche: $p < .05$; jüngere Erwachsene: $p < .01$; ältere Erwachsene: $p < .001$): Kürzere Lernzeiten gingen also mit einer besseren Erinnerungsleistung einher. Die ANOVA mit Altersgruppe als Zwischensubjektfaktor ergab einen signifikanten Haupteffekt ($F(3,108) = 3.50$; $p < .05$; $\eta^2 = .09$). Die Post-hoc-Tests zeigten, dass die älteren Erwachsenen eine höhere Korrelation (-.34) aufwiesen als die Jugendlichen (-.11; $p < .05$). Auch hier wurde in einer zweiten Analyse die Erinnerungsleistung nach dem Median der Lernzeit für jeden Teilnehmer getrennt berechnet (s. Tab. 29). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Lernzeit (über vs. unter dem Median) ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,108) = 13.77$; $p < .001$; $\eta^2 = .28$), einen signifikanten Haupteffekt der Lernzeit ($F(1,108) = 36.93$; $p < .001$; $\eta^2 = .25$) sowie eine signifikante Interaktion zwischen beiden Faktoren ($F(3,108) = 4.19$; $p < .01$; $\eta^2 = .10$). Der Haupteffekt der Altersgruppe entspricht dem oben schon post-hoc geprüften Effekt (s. Erinnerungsleistung). Weiterhin zeigte sich, dass eine höhere Erinnerungsleistung für kürzer gelernte Items nur bei den drei älteren Altersgruppen auftrat, nicht aber bei den Grundschulern (Jugendliche: 64.91 vs. 56.76% richtige Antworten; $p < .05$; jüngere Erwachsene: 82.39 vs. 76.57%; $p < .001$; ältere Erwachsene: 63.10 vs. 41.96%; $p < .001$).

Tabelle 30: Mittlere Gammakorrelationen zwischen Lernzeit und Erinnerungsleistung in Abhängigkeit der Altersgruppe im Lerndurchgang 2

Grundschüler	Jugendliche	Jüngere Erwachsene	Ältere Erwachsene
-0.13 (.29)	-0.11 (.27)	-0.21 (.27)	-0.34 (.32)

Standardabweichungen in Klammern.

6.6.4 Diskussion

Das Ziel dieser Studie war es, innerhalb eines Experiments, aber getrennt in zwei Lerndurchgängen zu untersuchen, inwiefern bei Personen verschiedenen Alters Nachweise sowohl für das MC- als auch für das CM-Modell erbracht werden können. Bislang stand für beide Modelle ein lebensspannenübergreifender Vergleich aus, für das CM-Modell existierten bis jetzt keine Studien, die ältere Erwachsene einschließen.

Um zu untersuchen, wie Informationen aus den JOLs für die nachfolgende Anpassung der Lernzeit genutzt werden (MC-Modell), mussten den Probanden die Wortpaare im ersten Lerndurchgang zweimal dargeboten werden, einmal mit vorgegebener und einmal mit freier Lernzeit. Um dennoch die Schwierigkeit zwischen den beiden Lerndurchgängen anzupassen, mussten demnach im zweiten Lerndurchgang weniger Wortpaare präsentiert werden. Auch hier wurden Anstrengungen unternommen, die Ausgangsleistung zwischen den Altersgruppen anzupassen, sodass die jüngeren Erwachsenen wieder die größte Anzahl an Items erhielten, gefolgt von den Jugendlichen und den Grundschulern und älteren Erwachsenen. Die Angleichung der Schwierigkeit zwischen den beiden Lerndurchgängen kann als weitgehend gelungen bezeichnet werden: Im Lerndurchgang 1 wurden etwa 10% mehr Wortpaare korrekt erinnert als im zweiten. Dieser Unterschied scheint jedoch dem Aufbau der Aufgabe geschuldet: So ist effizientes Lernen grundsätzlich einfacher, wenn vorher schon bekannt ist, welche Inhalte zu lernen sind, was in Lerndurchgang 2 nicht der Fall war. Die Anpassung der Leistung zwischen den Altersgruppen ist ebenfalls größtenteils geglückt. Lediglich die jüngeren Erwachsenen wiesen in beiden Lerndurchgängen höhere Leistungen auf als die anderen Altersgruppen.

Sowohl für den ersten als auch für den zweiten Teil des Experiments wurden gute Überwachungsleistungen bei den JOLs und nahezu perfekte bei den SUs gezeigt. Unterschiede zwischen den Altersgruppen sind hier zu vernachlässigen. Dies entspricht Befunden der vorliegenden Literatur, dass bereits Grundschüler vor allem bei einfach strukturierten Aufgaben mit offener Erinnerungsabfrage ihren Lernvorgang adäquat überwachen können (für einen

Überblick s. Schneider, 2015) und diese Fähigkeit über die Lebensspanne hinweg weitgehend stabil bleibt (Dunlosky & Metcalfe, 2009).

Bezüglich der zentralen Frage dieses Experiments, nämlich dem Nachweis des MC- und des CM-Modells, zeigten sich differenziertere Ergebnisse. Im Lerndurchgang 1, welcher das MC-Modell untersucht, ergaben sich signifikante negative Zusammenhänge zwischen JOLs und Lernzeit für alle Altersgruppen. Das heißt, im Mittel wurden Items, die höhere JOLs erhielten, in der Folge kürzer gelernt. Dies spricht dafür, dass die Teilnehmer sich bei der Anpassung ihrer Lernzeit nach den JOLs richteten und – in Übereinstimmung mit Vorbefunden – für alle Altersgruppen das MC-Modell als bestätigt gelten kann (vgl. Dunlosky & Connor, 1997; Koriat et al., 2006; Lockl & Schneider, 2003). Die erwarteten Altersverläufe in dem Sinne, dass sich der Zusammenhang zwischen JOLs und Lernzeit von der Grundschule bis zum jüngeren Erwachsenenalter steigert und danach bis zum höheren Erwachsenenalter wieder sinkt, konnten zum Teil bestätigt werden. Deskriptiv zeigen die Daten, dass die jüngeren Erwachsenen die höchsten Korrelationen aufwiesen, die Grundschüler und die älteren Erwachsenen im Mittelfeld lagen und bei den Jugendlichen die niedrigsten Korrelationen resultierten. Der Unterschied wurde jedoch lediglich im Vergleich zwischen Jugendlichen und jüngeren Erwachsenen signifikant. Grundsätzlich gehen also die jüngeren Erwachsenen am strategischsten und ökonomischsten vor: Sie lernen vor allem diejenigen Items, die sie nach einer ersten Lernphase als noch unzureichend gelernt einschätzen. Bei den anderen Altersgruppen zeigt sich hier noch Potenzial zur Verbesserung der Selbstregulation (vgl. auch Schneider, 2015; Souchay & Isingrini, 2004).

Auch das CM-Modell konnte prinzipiell nachgewiesen werden, jedoch nicht in allen Altersgruppen in gleicher Ausprägung. Die Auswertungsmethode nach dem Mediansplit erbrachte über alle untersuchten Gruppen hinweg einen signifikanten Haupteffekt in dem Sinn, dass kurz gelernte Items im Mittel höhere JOLs erhielten als lang gelernte. Die Korrelationen zwischen Lernzeit und JOLs waren jedoch insgesamt recht niedrig und erreichten nur bei den jüngeren und den älteren Erwachsenen das Signifikanzniveau. Die Ergebnisse sind insofern in Übereinstimmung mit der bisherigen Befundlage, als ab etwa der dritten Jahrgangsstufe ein Nachweis für das CM-Modell zu erwarten ist; die Stärke des Zusammenhangs sollte dann bis zum jüngeren Erwachsenenalter ansteigen (Hoffmann-Biencourt et al., 2010; Koriat et al., 2006; Koriat et al., 2009b). Die erwarteten Unterschiede zwischen Grundschulern und Jugendlichen konnten in dieser Studie jedoch nicht gefunden werden. Dies könnte daran liegen, dass – im Gegensatz zu den Arbeiten von Hoffmann-Biencourt et al. (2010) sowie Koriat et al. (2009b) – die Aufgaben für die Siebt- und Achtklässler schwieriger waren als die für die

Drittklässler und somit bei den Jugendlichen die Fähigkeit zur Selbstregulation durch die erhöhte Anforderung der Aufgabe im Vergleich zu anderen Studien näher am Niveau der Grundschüler lag (vgl. Schunk & Zimmerman, 1998). Die hohen Standardabweichungen bei den gemittelten Korrelationen zwischen Lernzeit und JOLs weisen – und das sowohl für das MC- als auch für das CM-Modell – darauf hin, dass es in jeder Altersgruppe sowohl Personen gab, die Lernzeiten und Überwachung recht stark anpassten, als auch Teilnehmer, bei denen die Lernzeit kein Hinweisreiz für die JOLs darstellte. Genauer zu untersuchen, was diese beiden Gruppen z.B. bezüglich ihrer Erinnerungsleistung unterscheidet, stellt eine sinnvolle Fragestellung für weitere Untersuchungen in größeren Stichproben dar. Bei den älteren Erwachsenen konnte für die Pearson-Korrelationen die erwartete Verringerung des Zusammenhangs zwischen Lernzeit und JOLs im Vergleich zu den jüngeren Erwachsenen gezeigt werden (im Sinne eines niedrigeren p-Werts bei den jüngeren Erwachsenen im Vergleich zu den älteren). In der inferenzstatistischen Analyse wurde dieser Alterseffekt jedoch nicht gefunden. Es scheint demnach für beide Modelle zu gelten, dass es älteren Erwachsenen, wahrscheinlich durch eine Verschlechterung der Leistung in den Exekutivfunktionen bedingt, etwas schwerer fällt, Hinweise aus Kontrolle bzw. Überwachung für die Anpassung des jeweils anderen Prozesses zu nutzen (vgl. Pansky et al., 2009; Souchay & Isingrini, 2004).

Es wurde zusätzlich für den zweiten Lerndurchgang untersucht, inwiefern die Lernzeit einen validen Prädiktor für die Erinnerungsleistung darstellt, d.h. inwiefern die Lernzeit tatsächlich einen diagnostischen Wert hat und als wertvoller Hinweisreiz für die JOLs dienen kann. Hier zeigten die Korrelationen, dass dies für jede Altersgruppe der Fall war. Durch Analyse mit dem Mediansplit wurde jedoch deutlich, dass der Unterschied in der Erinnerungsleistung zwischen lang und kurz gelernten Items erst für Teilnehmer ab dem Jugendalter signifikant wurde. Das heißt, die Lernzeit scheint im Verlauf der Entwicklung eine etwas stärkere Bedeutung für die tatsächliche Erinnerungsleistung zu bekommen. Grundsätzlich zeigen jedoch die insgesamt höheren Korrelationen zwischen Lernzeit und Erinnerungsleistung als zwischen Lernzeit und JOLs, dass die Lernzeit noch stärker als Hinweisreiz für die Höhe der Überwachungsurteile genutzt werden sollte.

Zusammenfassend lässt sich festhalten, dass sowohl für das MC-Modell als auch für das CM-Modell in allen Altersgruppen Nachweise gefunden werden konnten; dabei war letzterer Wirkzusammenhang etwas schwächer ausgeprägt als ersterer. Entwicklungsverläufe zeigten sich dahingehend, dass sich der erwartete Anstieg zum jüngeren Erwachsenenalter und die Abnahme des Zusammenhangs im höheren Erwachsenenalter für das MC-Modell vor allem deskriptiv zeigten, für das CM-Modell jedoch für den Verlauf vom Grundschul- zum jüngeren

Erwachsenenalter auch statistisch bedeutsam wurden. In der vorliegenden Studie wurden die Modelle getrennt in zwei Lerndurchgängen untersucht. Es wird jedoch angenommen und konnte für Personen ab dem Jugendalter bereits nachgewiesen werden, dass beide Modelle sich nicht gegenseitig ausschließen, sondern innerhalb derselben Aufgabe sowohl daten- als auch zielorientiertes Lernen auftreten kann (Koriat et al., 2006). Dies erstmals für ältere Erwachsene und in einem lebensspannenübergreifenden Ansatz zu untersuchen, ist das Ziel der nächsten Studie.

6.7 Studie 6: Simultane Verknüpfung von Überwachungs- und Kontrollprozessen

6.7.1 Herleitung der Fragestellung

Nachdem in Studie 5 erstmals für die Lebensspanne sowohl Belege für ziel- als auch für datenorientiertes Lernen erbracht wurden, soll im letzten Experiment dieser Arbeit der Frage nachgegangen werden, inwiefern das MC- und das CM-Modell in verschiedenen Altersstufen innerhalb eines einzigen Lernvorgangs auftreten. Denn schon Koriat et al. (2006) wiesen in ihrem ersten Artikel zu datenorientierten Lernprozessen darauf hin, dass im Alltag häufig beide Formen der Selbstregulation ineinander verschränkt auftreten. Dies prüften die Autoren für jüngere Erwachsene experimentell, indem sie den Items verschiedene hohe Anreizwerte zuordneten und diese in selbst eingeteilter Lernzeit darboten. Wie erwartet zeigte sich, dass innerhalb einer Anreizkategorie Items mit niedrigen Lernzeiten höhere JOLs erhielten (datenorientiertes Lernen) und im Vergleich zwischen den Anreizkategorien höher bewertete Items länger gelernt wurden (zielorientiertes Lernen).

Diese simultane Verknüpfung von MC- und CM-Modell wurde bislang im Entwicklungsverlauf nur für das Jugendalter untersucht (Koriat, Ackerman et al., 2014). Für Grundschüler und ältere Erwachsene existieren bislang keine Befunde. Bei Schülern der fünften und sechsten Klasse konnte mit dem Paradigma von Koriat et al. (2006) gezeigt werden, dass diese nicht in der Lage waren, flexibel zwischen daten- und zielorientiertem Lernen zu wechseln. Wenn innerhalb eines Blocks sowohl Items mit hohen und niedrigen Anreizwerten als auch einfache und schwierige Items dargeboten wurden, zeigten sich Hinweise auf das CM-Modell, nicht aber auf das MC-Modell. Das heißt, die Itemschwierigkeit hatte einen größeren Effekt auf die Regulation des Lernvorgangs als die Punktwerte, wenn beides gleichzeitig variiert wurde (Koriat, Ackerman et al., 2014). Obwohl bekannt ist, dass schon Sechsjährige ihr Lernverhalten nach dem Wert von Items ausrichten können (Castel, Lee, Humphrey & Moore, 2011), scheinen die Behaltenswerte weniger salient zu werden, wenn sich die Itemschwierigkeit stark zwischen den einzelnen Paaren unterscheidet. Erst Teilnehmer der neunten Jahrgangsstufe waren vergleichbar zu den jüngeren Erwachsenen in Koriat et al. (2006) in der Lage, flexibel zwischen MC- und CM-Modell innerhalb eines Lernvorgangs zu wechseln (Koriat, Ackerman et al., 2014). Dies könnte unter anderem mit einer Verbesserung der exekutiven Funktionen im Jugendalter zusammenhängen (Diamond, 2002) bzw. mit der Fähigkeit, mehrere Dimensionen gleichzeitig bei der Bearbeitung einer Aufgabe zu

berücksichtigen, wie dies laut Piaget ab dem konkret-operationalen Stadium, für abstraktere Inhalte womöglich jedoch erst ab dem formal-operationalen Stadium möglich ist (Piaget, 2000).

Für ältere Erwachsene weisen die Ergebnisse der Studie 5 erstmals das CM-Modell nach, d.h. auch Personen dieser Altersgruppe sind in der Lage, Informationen aus der Lernzeit für die JOLs zu nutzen. Allerdings steht eine Überprüfung der Frage, inwiefern im höheren Erwachsenenalter flexibel zwischen ziel- und datenorientiertem Lernen innerhalb einer Aufgabe gewechselt werden kann, noch aus. Studien zur Wirkung von Anreizen im Vergleich zwischen jüngeren und älteren Erwachsenen zeigen, dass auch Senioren ihre Lernzeit in Abhängigkeit des Wertes der Items einteilen. So wiesen Castel et al. (2002) den zu memorierenden Wörtern einer Liste gestaffelte Werte (z.B. von 1 bis 12) zu. Obwohl die jüngeren Erwachsenen insgesamt bessere Erinnerungsleistungen aufwiesen als die älteren, zeigte sich, dass letztere ihre Itemauswahl in der Mehrzahl der Bedingungen deutlich stärker an den Itemwert anpassten als die jüngeren Erwachsenen (vgl. auch Castel, Farb & Craik, 2007). Bei Price et al. (2010) wurde gleichzeitig zum Anreizwert die Schwierigkeit der Items manipuliert. Hier waren die älteren Teilnehmer im Vergleich zu den jüngeren etwas weniger sensitiv für den Behaltenswert einzelner Wörter und richteten ihre Lernzeit mehr nach der Schwierigkeit aus. Die Autoren erklären dieses Selbstregulationsverhalten mithilfe des Agenda-Based-Regulation-Modells von Ariel et al. (2009): Probanden jeden Alters orientieren demnach ihr Lernverhalten an ihren spezifischen Zielen. Da sich die älteren Erwachsenen bei Price et al. (2010) insgesamt niedrigere Lernziele gesetzt hatten als die jüngeren, kontrollierten sie den Lernvorgang entsprechend, indem sie schwierigere Items unabhängig vom Behaltenswert mit niedrigerer Wahrscheinlichkeit auswählten. Auch die Ergebnisse von Castel et al. (2002) lassen sich nach diesem Modell interpretieren: Wenn alle Items gleich schwierig sind, ist es für ältere Probanden adaptiv, selektiver zu sein und die verfügbare kognitive Energie vermehrt auf die hoch bewerteten Items zu richten.

Aufgrund dieser Ergebnisse ist bei einer gleichzeitigen Manipulation von Ziel- und Datenorientierung innerhalb einer Aufgabe für die älteren Erwachsenen zu erwarten, dass diese zwar in der Lage sind, ihr Lernverhalten an den Behaltenswerten von Items auszurichten; wenn dies jedoch mit ihrem Ziel interferiert, sich insgesamt möglichst viele Wortpaare zu merken, ist anzunehmen, dass sich die Itemschwierigkeit in höherem Maße auf die Kontrolle des Lernvorgangs auswirkt als der Anreizwert, da sie vermutlich im Vergleich salienter ist (vgl. Price et al., 2010).

Es ist also insgesamt zu vermuten, dass sowohl die Grundschüler als auch die älteren Erwachsenen im Gegensatz zu den Jugendlichen und den jüngeren Erwachsenen Schwierigkeiten dabei aufweisen, flexibel innerhalb einer Aufgabe zwischen MC- und CM-Modell zu wechseln. Für beide Altersgruppen sollte die Itemschwierigkeit salienter sein als der Behaltenswert, sodass sich eher Nachweise für datenorientiertes Lernen finden lassen sollten. Da die Studie 5 dieser Arbeit die erste lebensspannenübergreifende Prüfung des CM-Modells darstellt, soll im ersten Lerndurchgang eine Replikation des Lerndurchgangs 2 dieser Studie stattfinden.

6.7.2 Methodik

6.7.2.1 Design

Beim Versuchsplan handelte es sich um ein 4 (Altersgruppe: Kinder, Jugendliche, jüngere und ältere Erwachsene) x 2 (Itemschwierigkeit: zusammenhängend vs. unzusammenhängend) faktorielles Design. Im zweiten Lerndurchgang wurde zusätzlich der Anreizwert der Items (hoch vs. niedrig) als Faktor aufgenommen. Die Altersgruppe wurde als Zwischensubjektfaktor, die Itemschwierigkeit und der Anreizwert als Innersubjektfaktoren erfasst. Die Erinnerungsleistung, die JOLs, die SUs sowie die selbst gewählte Lernzeit bildeten die abhängigen Variablen.

6.7.2.2 Stichprobe

An der Studie nahmen insgesamt $N = 112$ Personen teil. Darunter waren 28 Drittklässler (17 männlich, 11 weiblich), 28 Jugendliche der siebten und achten Klasse (15 männlich, 13 weiblich), 28 jüngere Erwachsene (11 männlich, 17 weiblich) sowie 28 ältere Erwachsene (10 männlich, 18 weiblich). Im Mittel waren die Grundschüler 8.14 Jahre alt ($SD = 0.36$; Altersspanne: 8 bis 9 Jahre), die Jugendlichen 13.07 Jahre ($SD = 0.81$; Altersspanne: 12 bis 15), die jüngeren Erwachsenen 21.89 Jahre ($SD = 2.59$; Altersspanne: 18 bis 26) und die älteren Erwachsenen 67.18 Jahre ($SD = 4.78$; Altersspanne: 60 bis 75). Die Probanden wurden gleichermaßen wie bei den bisher vorgestellten Studien rekrutiert.

Bezüglich der Normwerte im Zahlennachsprechen (Petermann & Petermann, 2010; von Aster et al., 2006) ergaben sich keine Unterschiede zwischen den Altersgruppen. Die Normwerte der einzelnen Altersgruppen lauteten wie folgt: Grundschüler: $M = 11.54$ ($SD = 2.76$); Jugendliche:

$M = 10.79$ ($SD = 2.91$); jüngere Erwachsene: $M = 11.79$ ($SD = 3.29$); ältere Erwachsene: $M = 11.04$ ($SD = 2.40$).

Auch im Wortschatztest (Petermann & Petermann, 2010; von Aster et al., 2006) unterschieden sich die beiden Altersgruppen in ihren Normwerten nicht voneinander. Hier ergaben sich folgende Werte: Grundschüler: $M = 10.00$ ($SD = 2.64$); Jugendliche: $M = 9.04$ ($SD = 2.06$); jüngere Erwachsene: $M = 9.82$ ($SD = 2.36$); ältere Erwachsene: $M = 10.25$ ($SD = 3.37$).

Die mittlere Schul- und Ausbildungszeit unterschied sich nicht zwischen den jüngeren und den älteren Erwachsenen (jüngere Erwachsene: $M = 13.93$ Jahre; $SD = 2.57$; ältere Erwachsene: $M = 13.18$ Jahre; $SD = 3.71$). Der Wert für den subjektiv eingeschätzten Gesundheitszustand lag bei den jüngeren Erwachsenen etwas niedriger, also positiver, als bei den älteren ($M = 1.79$; $SD = 0.63$ vs. $M = 2.27$; $SD = 0.75$; $t(54) = 2.60$; $p < .05$). Die durchschnittliche Anzahl täglich eingenommener Medikamente lag ebenfalls bei den jüngeren Erwachsenen niedriger als bei den älteren ($M = 0.18$; $SD = 0.39$ vs. $M = 2.35$; $SD = 2.00$; $t(53) = 5.78$; $p < .001$).

6.7.2.3 Versuchsmaterial

Wortpaare. Das Lernmaterial bestand aus zweisilbigen konkreten Substantiven, die aus den Wortpaaren des Experiments 5 ausgewählt wurden (s. Anhang A 9). Auch hier waren 50% der Items einfach, d.h. semantisch assoziiert, und 50% schwierig, d.h. semantisch nicht zusammenhängend. Den Grundschulern und den älteren Erwachsenen wurden in beiden Lerndurchgängen 24 Wortpaare dargeboten, den Jugendlichen 36 und den jüngeren Erwachsenen 60. Im Vergleich zum Lerndurchgang 2 in Studie 5 ist die Zahl der Items bei den jüngeren Erwachsenen erhöht worden, um bestehende Unterschiede in der Erinnerungsleistung weiter auszugleichen.

Der zweite Lerndurchgang unterschied sich dahingehend vom ersten, dass nun in der Lernphase die Hälfte der einfachen und die Hälfte der schwierigen Wortpaare wieder jeweils zu 50% mit einem großen Stern, der innen mit einer „5“ beschriftet war, bzw. mit einem kleinen Stern mit der Beschriftung „1“ versehen waren.

Die JOLs und die SUs wurden am Computer mithilfe der Thermometerskala erhoben.

Zusatzvariablen. Um die Bedeutsamkeit des Anreizes und der leistungsabhängigen Belohnung zu messen, kam ein kurzer Fragebogen zum Einsatz (s. Anhang C 4). Dieser erfasste, wie

wichtig es den Teilnehmern war, viele Punkte zu sammeln und möglichst viel von der zusätzlichen Belohnung zu erhalten, ob die Teilnehmer der Ansicht waren, dass die Belohnung sie motiviert habe, die hoch bewerteten Items länger und besser zu lernen, und – nur bei den beiden jüngeren Altersgruppen – wie gut ihnen die Gummibänder als Extrageschenke gefallen hätten. Die Antworten wurden jeweils auf einer Smiley-Skala von 1 bis 7 festgehalten, wobei 1 (trauriger Smiley) für die stärkste Ablehnung, 7 (fröhlicher Smiley) für die stärkste Zustimmung stand. Weiterhin wurden wie in den bereits vorgestellten Studien die Leistungen im Zahlennachsprechen und im Wortschatz (Petermann & Petermann, 2010; von Aster et al., 2006) erfasst. Bei den jüngeren und älteren Erwachsenen kam zusätzlich der Fragebogen zu den demografischen Angaben und zum Gesundheitszustand zum Einsatz.

6.7.2.4 Studienablauf

Die Studie wurde als Einzeltestung in den Räumen des Instituts für Psychologie bzw. in den Schulen durchgeführt. Bei den minderjährigen Teilnehmern lag das schriftliche Einverständnis der Eltern vor.

Der grundlegende Ablauf beider Lerndurchgänge erfolgte identisch zum zweiten Lerndurchgang in Studie 5 (vgl. auch Koriatic et al., 2009b). Der zweite Durchgang unterschied sich vom ersten lediglich dahingehend, dass nun die Hälfte der assoziativen und der nicht assoziativen Wortpaare in der Lernphase mit einem Behaltenswert von einem bzw. fünf Punkten versehen war. Die Teilnehmer erhielten die Instruktion, so viele Punkte wie möglich zu sammeln. Um den Anreiz zu steigern, erhielten die Grundschüler und die Jugendlichen je nach Erinnerungsleistung eine Anzahl von Gummibändern, die zu Armreifen gebunden werden können (ein Band je erreichtem Punkt). Die jüngeren und die älteren Erwachsenen erhielten eine monetäre Belohnung (jüngere Erwachsene: 25 Cent je 10 Punkte; ältere Erwachsene: 50 Cent je 10 Punkte). Die Teilnehmer wurden am Ende des zweiten Lerndurchgangs zu ihrer Einstellung zu den Punktwerten der Items und der Extrabelohnung befragt (s. o.). Zusätzlich erhielt jeder Teilnehmer leistungsunabhängig die gleichen Belohnungen wie in den anderen Studien als Dank für die Teilnahme.

6.7.2.5 Überblick über die statistischen Auswertungen

Beide Lerndurchgänge wurden getrennt ausgewertet. Die Auswertung zum Lerndurchgang 1 entspricht der zum Lerndurchgang 2 in Studie 5. Um in Lerndurchgang 2 zusätzlich zu

bestimmen, inwiefern zielorientiertes Lernen auftrat, wurden für die Höhe der JOLs und die Lernzeit ANVOAs mit den Faktoren Itemschwierigkeit und Lernanreiz als Innersubjektfaktoren und Altersgruppe als Zwischensubjektfaktor berechnet (vgl. Koriat, Ackerman et al., 2014). Der Fragebogen zur subjektiven Bedeutsamkeit der Anreize wurde mit ANOVAs bzw. *t*-Tests für die einzelnen Fragen mit Altersgruppe als Zwischensubjektfaktor ausgewertet.

6.7.3 Ergebnisse

6.7.3.1 Lerndurchgang 1

6.7.3.1.1 Erinnerungsleistung

Der Anteil richtiger Antworten im Lerndurchgang 1 ist in Tabelle 31 abgebildet. Eine ANOVA mit Altersgruppe als Zwischensubjektfaktor und Itemschwierigkeit als Innersubjektfaktor ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,108) = 12.68; p < .001; \eta^2 = .26$). In den Post-hoc-Tests zeigte sich, dass die Grundschüler (46.43%) und die älteren Erwachsenen (50.60%) etwas schlechter abschnitten als die Jugendlichen (66.87%) und die jüngeren Erwachsenen (73.51%). Außerdem erreichten der Haupteffekt des Faktors Itemschwierigkeit ($F(1,108) = 303.36; p < .001; \eta^2 = .74$) sowie die Interaktion zwischen den beiden Faktoren ($F(3,108) = 5.59; p < .01; \eta^2 = .13$) das Signifikanzniveau. Insgesamt gesehen wurden mehr einfache (78.01%) als schwierige Items (40.69%) erinnert. *T*-Tests zeigten, dass dieser Trend auch für jede Altersgruppe einzeln galt (alle *p*'s < .001). Allerdings wird anhand der Interaktion deutlich, dass der Unterschied zwischen leichten und schwierigen Items bei den jüngeren Erwachsenen geringer ausfiel als bei den drei anderen Altersgruppen (Grundschüler: $d = 2.15$; Jugendliche: $d = 1.65$; jüngere Erwachsene: $d = 1.51$; ältere Erwachsene: $d = 2.15$).

Tabelle 31: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe und Item-Schwierigkeit im Lerndurchgang 1

	Gesamt	Einfache Items	Schwierige Items
Grundschüler	46.43 (18.28)	68.15 (23.58)	24.70 (21.21)
Jugendliche	66.87 (18.69)	82.14 (14.92)	51.59 (25.52)
Jüngere Erwachsene	73.51 (20.37)	87.02 (15.88)	60.00 (29.06)
Ältere Erwachsene	50.60 (19.40)	74.70 (18.77)	26.49 (25.46)

Standardabweichungen in Klammern.

6.7.3.1.2 Judgments of Learning

Die Differenzierungsleistung in den JOLs wurde wieder durch einen Vergleich der mittleren JOLs vor richtigen bzw. falschen Antworten bestimmt (s. Tab. 32). In der ANOVA mit Altersgruppe als Zwischensubjektfaktor und Korrektheit der Antwort als Innersubjektfaktor wurde lediglich ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort gefunden ($F(1,108) = 914.48$; $p < .001$; $\eta^2 = .89$). Vor richtigen Antworten (87.39) wurden deutlich höhere JOLs abgegeben als vor falschen (32.88). Zur Bestimmung der Genauigkeit wurden die Gammakorrelationen zwischen JOLs und Erinnerungsleistung berechnet (s. Tab. 32). *T*-Tests zeigten, dass die mittleren Korrelationen in allen Altersgruppen signifikant von Null verschieden waren (alle p 's $< .001$). Die dazugehörige ANOVA mit Altersgruppe als Zwischensubjektfaktor ergab keine signifikanten Effekte.

Tabelle 32: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe und Versuchsbedingung; mittlere Gammakorrelationen für JOLs in Abhängigkeit der Altersgruppe im Lerndurchgang 1

	Mittlere JOLs		Gamma- korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	86.44 (14.00)	30.37 (17.33)	.88 (.23)
Jugendliche	87.54 (12.31)	29.35 (17.73)	.90 (.12)
Jüngere Erwachsene	88.86 (10.78)	34.83 (16.59)	.87 (.13)
Ältere Erwachsene	86.72 (12.60)	36.98 (17.76)	.82 (.15)

Standardabweichungen in Klammern.

6.7.3.1.3 Sicherheitsurteile

Bei den SUs ergab sich hinsichtlich der Differenzierungsleistung (s. Tab. 33) in der ANOVA mit Altersgruppe als Zwischensubjektfaktor und Korrektheit der Antwort als Innersubjektfaktor ein signifikanter Haupteffekt des Faktors Korrektheit der Antwort ($F(1,108) = 1333.55$; $p < .001$; $\eta^2 = .93$). Nach richtigen Antworten (90.40) wurden deutlich höhere SUs abgegeben als nach falschen (21.21). Die mittleren Gammakorrelationen zwischen SUs und Erinnerungsleistung als Maß für die Genauigkeit der Überwachung waren für alle Altersgruppen von Null verschieden (alle p 's $< .001$; vgl. Tab. 33). In der entsprechenden

ANOVA mit Altersgruppe als Zwischenssubjektfaktor wurden keine signifikanten Effekte gefunden.

Tabelle 33: Mittlere SUs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe; mittlere Gammakorrelationen für SUs in Abhängigkeit der Altersgruppe in Lerndurchgang 1

	Mittlere SUs		Gamma-korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	91.35 (9.68)	22.60 (20.06)	.97 (.06)
Jugendliche	90.67 (18.85)	19.58 (16.83)	.97 (.05)
Jüngere Erwachsene	93.99 (5.89)	24.35 (19.79)	.87 (.38)
Ältere Erwachsene	85.58 (14.20)	18.30 (14.18)	.91 (.14)

Standardabweichungen in Klammern.

6.7.3.1.4 Lernzeit als Hinweisreiz für die Judgments of Learning

Wie in Studie 5 wurde auch hier für den ersten Lerndurchgang berechnet, inwiefern die Lernzeit als Hinweisreiz für die JOLs genutzt wurde. Dafür wurden Pearson-Korrelationen zwischen Lernzeiten und JOLs für jeden Teilnehmer berechnet und für die Altersgruppen gemittelt (s. Tab. 34). Bei den drei älteren Gruppen waren diese signifikant von Null verschieden (Jugendliche und jüngere Erwachsene: $p < .05$; ältere Erwachsene: $p < .01$), bei den Grundschulern wurde die Signifikanzschwelle leicht verfehlt ($p = .059$). Auch hier wurde eine ergänzende Analyse durchgeführt, bei der die JOLs nach dem Median der Lernzeit (JOLs bei kurz versus lang gelernten Items) verglichen wurden (s. Tab. 35). In der ANOVA mit Altersgruppe als Zwischenssubjektfaktor und Lernzeit (über vs. unter dem Median) als Innersubjektfaktor wurde ein signifikanter Haupteffekt des Faktors Altersgruppe gefunden ($F(3,108) = 6.71$; $p < .001$; $\eta^2 = .16$). Im Post-hoc-Test zeigte sich, dass die jüngeren Erwachsenen (73.12) höhere JOLs abgaben als die Grundschüler (57.75) und die älteren Erwachsenen (60.75). Außerdem ergab sich ein signifikanter Haupteffekt des Faktors Lernzeit ($F(1,108) = 27.82$; $p < .001$; $\eta^2 = .21$). Kürzer gelernte Items erhielten höhere JOLs (69.09) als länger gelernte (60.96).

Tabelle 34: Mittlere Pearson-Korrelationen zwischen JOLs und Lernzeiten in Abhängigkeit der Altersgruppe, Lerndurchgang und Anreizwert

	Lerndurchgang 1	Lerndurchgang 2	
		1 Pkt.	5 Pkt.
Grundschüler	-.09 (.24)	-.02 (.32)	-.12 (.32)
Jugendliche	-.13 (.25)	-.18 (.29)	-.13 (.22)
Jüngere Erwachsene	-.09 (.18)	-.16 (.21)	-.05 (.24)
Ältere Erwachsene	-.16 (.28)	-.15 (.29)	-.21 (.28)

Standardabweichungen in Klammern.

Tabelle 35: Mittlere JOLs und Erinnerungsleistung in Prozent für Items mit einer Lernzeit über bzw. unter dem Median in Abhängigkeit der Altersgruppe im Lerndurchgang 1

	Mittlere JOLs		Mittlere Erinnerungsleistung	
	Unter dem Median	Über dem Median	Unter dem Median	Über dem Median
Grundschüler	62.97 (14.88)	52.53 (13.80)	54.46 (20.72)	38.39 (20.34)
Jugendliche	72.40 (13.48)	64.53 (17.81)	69.87 (17.97)	63.43 (24.29)
Jüngere Erwachsene	76.00 (14.87)	70.24 (16.24)	76.92 (21.19)	70.56 (21.48)
Ältere Erwachsene	64.97 (22.05)	56.54 (17.38)	57.74 (24.42)	43.45 (21.56)

Standardabweichungen in Klammern.

6.7.3.1.5 Validität der Lernzeit als Prädiktor für die Erinnerungsleistung

Auch hier wurden Gammakorrelationen zwischen Lernzeit und Erinnerungsleistung für jeden Teilnehmer berechnet, um beurteilen zu können, inwieweit erstere tatsächlich ein valider Prädiktor für die Lernleistung ist (s. Tab. 36). Die gemittelten Korrelationen waren für alle Altersgruppen bis auf die Jugendlichen signifikant von Null verschieden (Grundschüler: $p < .001$; jüngere und ältere Erwachsene: $p < .01$). Die dazugehörige ANOVA mit Altersgruppe als Zwischensubjekt-Faktor ergab keine signifikanten Ergebnisse. Auch hier wurde eine zweite Analyse, bei der die Erinnerungsleistung nach dem Median der Lernzeit für jeden Teilnehmer verglichen wurde, durchgeführt (s. Tab. 35). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Lernzeit (über vs. unter dem Median) ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,108) = 12.67$; $p < .001$; $\eta^2 = .26$). Dieser entspricht dem bei der Erinnerungsleistung berichteten Ergebnis, dass die Grundschüler und die älteren

Erwachsenen etwas schlechter abschnitten als die Jugendlichen und die jüngeren Erwachsenen. Weiterhin zeigte sich ein signifikanter Haupteffekt des Faktors Lernzeit ($F(1,108) = 34.43; p < .001; \eta^2 = .24$). Kürzer gelernte Items (64.75%) wurden besser erinnert als länger gelernte (53.96%).

Tabelle 36: Mittlere Gammakorrelationen zwischen Lernzeit und Erinnerungsleistung in Abhängigkeit von Altersgruppe, Lerndurchgang und Anreizwert

	Lerndurchgang 1	Lerndurchgang 2	
		1 Pkt.	5 Pkt.
Grundschüler	-.20 (.23)	-.12 (.37)	-.19 (.46)
Jugendliche	-.09 (.35)	-.20 (.37)	-.19 (.32)
Jüngere Erwachsene	-.15 (.25)	-.16 (.44)	-.06 (.36)
Ältere Erwachsene	-.25 (.34)	-.27 (.44)	-.31 (.34)

Standardabweichungen in Klammern.

6.7.3.2 Lerndurchgang 2

6.7.3.2.1 Erinnerungsleistung

Im Lerndurchgang 2 wurde zusätzlich der Faktor Anreizwert (1 vs. 5 Punkte) in die Analyse aufgenommen, um in Folge daten- und zielorientierte Lernvorgänge innerhalb derselben Aufgabe voneinander trennen zu können (s. Tab. 37). In die ANOVA flossen also nun die Altersgruppe als Zwischensubjektfaktor sowie die Itemschwierigkeit und der Anreizwert als Innersubjektfaktoren ein. Es zeigte sich ein signifikanter Haupteffekt des Faktors Anreizwert ($F(1,108) = 36.92; p < .001; \eta^2 = .26$). Höher bewertete Items (69.59%) wurden besser erinnert als niedriger bewertete (60.12%). Weiterhin ergaben sich signifikante Effekte der Faktoren Altersgruppe ($F(3,108) = 5.58; p < .01; \eta^2 = .13$) und Itemschwierigkeit ($F(3,108) = 352.80; p < .001; \eta^2 = .77$) sowie eine signifikante Wechselwirkung zwischen diesen beiden Faktoren ($F(3,108) = 12.21; p < .001; \eta^2 = .25$). Insgesamt erbrachten die jüngeren Erwachsenen (75.00%) eine bessere Erinnerungsleistung als die Grundschüler (58.18%; $p < .01$ im Post-hoc-Test) und die älteren Erwachsenen (59.08%; $p < .05$). Außerdem wurden die einfachen Items (83.64%) deutlich besser erinnert als die schwierigeren (46.07%), wobei dieser Unterschied bei den Grundschulern und den älteren Erwachsenen deutlich höher ausfiel als bei den anderen

beiden Altersgruppen (Grundschüler: $d = 2.72$; Jugendliche: $d = 1.26$; jüngere Erwachsene: $d = 1.12$; ältere Erwachsene: $d = 2.46$).

Tabelle 37: Erinnerungsleistung in Prozent richtiger Antworten in Abhängigkeit von Altersgruppe, Itemschwierigkeit und Anreizwert im Lerndurchgang 2

	Gesamt	Einfache Items		Schwierige Items	
		1 Pkt.	5 Pkt.	1 Pkt.	5 Pkt.
Grundschüler	58.18 (13.77)	73.81 (19.99)	88.69 (12.87)	27.98 (23.59)	42.26 (28.50)
Jugendliche	67.16 (18.98)	76.98 (19.34)	84.52 (13.97)	49.60 (31.50)	57.54 (27.23)
Jüngere Erwachsene	75.00 (19.48)	84.05 (15.85)	90.24 (15.34)	56.67 (31.14)	69.05 (26.93)
Ältere Erwachsene	59.08 (17.79)	81.55 (18.34)	89.29 (11.31)	30.36 (25.68)	35.12 (32.82)

Standardabweichungen in Klammern.

6.7.3.2.2 Judgments of Learning

Auch hier wurden zur Bestimmung der Güte der JOLs Differenzierungsleistung und Genauigkeit berechnet (s. Tab. 38). Die ANOVA mit dem Zwischensubjektfaktor Altersgruppe und dem Innersubjektfaktor Korrektheit der Antwort ergab einen signifikanten Haupteffekt der Korrektheit der Antwort ($F(1,108) = 856.70$; $p < .001$; $\eta^2 = .89$): Die Teilnehmer gaben im Mittel höhere JOLs vor richtigen (88.97) als vor falschen Antworten (31.28). Die mittleren Gammakorrelationen zwischen JOLs und Erinnerungsleistung waren sämtlich von Null verschieden (alle p 's $< .001$). Die dazugehörige ANOVA fand keinen signifikanten Effekt des Faktors Altersgruppe (s. Tab. 38)

Tabelle 38: Mittlere JOLs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe; mittlere Gammakorrelationen für JOLs in Abhängigkeit der Altersgruppe im Lerndurchgang 2

	Mittlere JOLs		Gamma- korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	87.71 (11.53)	25.57 (15.08)	.94 (.08)
Jugendliche	89.78 (12.50)	37.01 (20.89)	.89 (.16)
Jüngere Erwachsene	89.75 (10.56)	28.00 (18.40)	.90 (.13)
Ältere Erwachsene	88.66 (11.30)	34.54 (21.99)	.79 (.37)

Standardabweichungen in Klammern.

6.7.3.2.3 Sicherheitsurteile

Auch für die Differenzierungsleistung in den SUs (s. Tab. 39) wurde in der ANOVA mit dem Zwischensubjektfaktor Altersgruppe und dem Innersubjektfaktor Korrektheit der Antwort lediglich der Haupteffekt der Korrektheit der Antwort signifikant ($F(1,108) = 1395.44$; $p < .001$; $\eta^2 = .93$). Die mittleren SUs nach richtigen Antworten waren deutlich höher (91.44) als die nach falschen Antworten (21.91). Die mittleren Gammakorrelationen waren ebenfalls in jeder Altersgruppe signifikant von Null verschieden (alle p 's $< .001$). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor ergab keine signifikanten Effekte (s. Tab. 39).

Tabelle 39: Mittlere SUs in Abhängigkeit von Korrektheit der Antwort und Altersgruppe; mittlere Gammakorrelationen für SUs in Abhängigkeit der Altersgruppe im Lerndurchgang 2

	Mittlere SUs		Gamma- korrelationen
	Richtige Antworten	Falsche Antworten	
Grundschüler	89.97 (11.10)	22.24 (17.82)	.95 (.06)
Jugendliche	92.11 (19.52)	26.87 (24.17)	.97 (.07)
Jüngere Erwachsene	93.47 (6.65)	20.54 (16.50)	.93 (.09)
Ältere Erwachsene	90.20 (10.57)	18.00 (12.95)	.97 (.05)

Standardabweichungen in Klammern.

6.7.3.2.4 Auswirkung der Anreizwerte auf die Höhe der Judgments of Learning und die Lernzeit

Um beurteilen zu können, inwiefern die Teilnehmer ihr Lernverhalten an das Ziel, viele Punkte zu erreichen, anpassten (vgl. Koriat, Ackerman et al., 2014), wurden neben der Erinnerungsleistung (s.o.) die Höhe der JOLs sowie die Lernzeit in Abhängigkeit der Punktwerte verglichen.

Für die JOLs (s. Tab. 40) ergab die entsprechende ANOVA mit Altersgruppe als Zwischensubjektfaktor sowie Anreizwert und Itemschwierigkeit als Innersubjektfaktoren signifikante Haupteffekte der Faktoren Altersgruppe ($F(3,108) = 4.60; p < .01; \eta^2 = .11$) und Anreizwert ($F(1,108) = 25.80; p < .001; \eta^2 = .19$) sowie eine signifikante Wechselwirkung zwischen beiden Faktoren ($F(3,108) = 3.38; p < .05; \eta^2 = .09$). Der Haupteffekt des Faktors Altersgruppe war im Post-hoc-Test darauf zurückzuführen, dass die Grundschüler insgesamt niedrigere JOLs (60.90) abgaben als die jüngeren Erwachsenen (73.87; $p < .05$). Bezüglich der Interaktion zeigten t -Tests für die einzelnen Altersgruppen, dass lediglich die Grundschüler, die Jugendlichen (beide p 's $< .01$) und die jüngeren Erwachsenen ($p < .001$), nicht aber die älteren Erwachsenen höhere JOLs bei 5-Punkt- als bei 1-Punkt-Items abgaben. Auch für die Itemschwierigkeit ergaben sich ein signifikanter Haupteffekt ($F(1,108) = 263.75; p < .001; \eta^2 = .71$) sowie eine Wechselwirkung mit dem Faktor Altersgruppe ($F(3,108) = 8.88; p < .001; \eta^2 = .20$). T -Tests für die einzelnen Altersgruppen zeigten, dass in allen Gruppen höhere JOLs vor einfachen als vor schwierigen Items abgegeben wurden (alle p 's $< .001$); der Unterschied war jedoch bei den Grundschülern und den älteren Erwachsenen stärker ausgeprägt als bei den anderen beiden Altersgruppen (Grundschüler: $d = 1.81$; Jugendliche: $d = 1.02$; jüngere Erwachsene: $d = 1.18$; ältere Erwachsene: $d = 2.24$).

Tabelle 40: Mittlere JOLs in Abhängigkeit von Altersgruppe, Anreizwert und Itemschwierigkeit im Lerndurchgang 2

	Einfache Items		Schwierige Items	
	1 Pkt.	5 Pkt.	1 Pkt.	5 Pkt.
Grundschüler	71.33 (16.58)	80.39 (15.48)	41.01 (21.59)	50.87 (22.49)
Jugendliche	77.31 (18.76)	83.67 (11.89)	59.98 (23.94)	65.89 (18.78)
Jüngere Erwachsene	81.22 (15.66)	87.30 (12.82)	58.09 (24.45)	68.88 (19.78)
Ältere Erwachsene	85.26 (12.83)	84.02 (14.45)	44.65 (21.62)	45.56 (26.50)

Standardabweichungen in Klammern.

Für die mittleren Lernzeiten ergab sich ein ähnliches Bild (vgl. Tab. 41). In der ANOVA mit Altersgruppe als Zwischensubjektfaktor und Anreizwert als Innersubjektfaktor wurden beide Haupteffekte (Altersgruppe: $F(3,108) = 6.27$; $p < .01$; $\eta^2 = .15$; Anreizwert: ($F(1,108) = 22.98$; $p < .001$; $\eta^2 = .18$) sowie die Interaktion zwischen beiden Faktoren signifikant ($F(3,108) = 3.21$; $p < .05$; $\eta^2 = .08$). Der Effekt der Altersgruppe ergab sich daraus, dass die jüngeren Erwachsenen (12.28 Sek.) im Mittel längere Lernzeiten aufwiesen als die Grundschüler (8.20 Sek.; $p < .05$ im Post-hoc-Test) und die älteren Erwachsenen (7.43 Sek.; $p < .01$). Wie bei den JOLs wurden t -Tests für die einzelnen Altersgruppen berechnet. Diese zeigten, dass lediglich die Jugendlichen ($p < .05$) und die jüngeren Erwachsenen ($p < .001$) die höher bewerteten Items länger lernten als die niedriger bewerteten. Weiterhin wurden der Haupteffekt des Faktors Itemschwierigkeit ($F(1,108) = 100.46$; $p < .001$; $\eta^2 = .48$) sowie die Interaktion von Itemschwierigkeit und Altersgruppe ($F(3,108) = 4.22$; $p < .01$; $\eta^2 = .11$) signifikant. Auch hier zeigten t -Tests für die einzelnen Altersgruppen, dass einfachere Items insgesamt kürzer gelernt wurden als schwierigere (Grundschüler: $p < .01$; alle anderen p 's $< .001$). Der Unterschied war jedoch bei den älteren Erwachsenen am stärksten ausgeprägt (Grundschüler: $d = 0.37$; Jugendliche: $d = 0.53$; jüngere Erwachsene: $d = 0.74$; ältere Erwachsene: $d = 0.90$).

Tabelle 41: Mittlere Lernzeiten in Sekunden in Abhängigkeit von Altersgruppe, Anreizwert und Itemschwierigkeit im Lerndurchgang 2

	Einfache Items		Schwierige Items	
	1 Pkt.	5 Pkt.	1 Pkt.	5 Pkt.
Grundschüler	7.19 (2.77)	7.99 (4.29)	8.16 (3.17)	9.46 (5.27)
Jugendliche	7.72 (4.45)	9.44 (5.87)	11.08 (7.02)	12.39 (7.62)
Jüngere Erwachsene	8.77 (4.14)	11.76 (5.93)	12.74 (6.36)	15.86 (7.61)
Ältere Erwachsene	5.63 (2.33)	5.92 (2.32)	8.82 (5.52)	9.34 (4.14)

Standardabweichungen in Klammern.

6.7.3.2.5 Lernzeit als Hinweisreiz für die Judgments of Learning

Auch im zweiten Lerndurchgang wurde untersucht, inwiefern die Lernzeit als Hinweisreiz für die JOLs genutzt wurde, um Hinweise auf das CM-Modell zu erhalten. Dies wurde getrennt für die beiden Anreizwerte analysiert, um die Zielorientierung der Teilnehmer konstant zu halten.

Zunächst wurden für jeden Behaltenswert Pearson-Korrelationen zwischen Lernzeit und JOLs für jeden Teilnehmer berechnet (s. Tab. 34). Diese waren bei den Jugendlichen (beide p 's < .01) und bei den älteren Erwachsenen (beide p 's < .001) signifikant von Null verschieden. Für die jüngeren Erwachsenen traf dies nur auf die mit einem Punkt bewerteten Wortpaare zu (p < .001). Eine ANOVA mit Altersgruppe als Zwischensubjektfaktor und Anreizwert als Innersubjektfaktor ergab weder einen signifikanten Haupteffekt noch signifikante Interaktionen. Auch hier wurde eine zusätzliche Analyse, bei der die JOLs für jeden Anreizwert nach dem Median der Lernzeit in lang und kurz gelernte Items aufgeteilt wurden, durchgeführt (s. Tab. 42). Die dazugehörige ANOVA mit Altersgruppe als Zwischensubjektfaktor und Lernzeit (über vs. unter dem Median) sowie Anreizwert als Innersubjektfaktoren ergab einen signifikanten Haupteffekt des Faktors Lernzeit ($F(1,108) = 30.30$; $p < .001$; $\eta^2 = .22$). Kürzer gelernte Items erhielten im Schnitt höhere JOLs (72.15) als länger gelernte (64.93). Weiterhin wurden die Haupteffekte der Faktoren Altersgruppe $F(3,108) = 4.50$; $p < .01$; $\eta^2 = .11$) und Anreizwert ($F(1,108) = 23.81$; $p < .001$; $\eta^2 = .18$) sowie die Interaktion beider Faktoren signifikant ($F(3,108) = 3.58$; $p < .05$; $\eta^2 = .09$). Diese beiden Haupteffekte und die Interaktion entsprechen den bei der Zielorientierung berichteten Effekten, dass in allen Altersgruppen außer bei den älteren Erwachsenen höhere JOLs bei 5-Punkt-Items als bei 1-Punkt-Items abgegeben wurden.

Tabelle 42: Mittlere JOLs für Items mit einer Lernzeit über bzw. unter dem Median in Abhängigkeit von Altersgruppe und Anreizwert im Lerndurchgang 2

	Unter dem Median		Über dem Median	
	1 Pkt.	5 Pkt.	1 Pkt.	5 Pkt.
Grundschüler	57.52 (22.10)	69.62 (18.64)	54.78 (22.41)	61.17 (21.19)
Jugendliche	72.11 (22.86)	77.98 (14.42)	65.00 (19.68)	71.43 (15.85)
Jüngere Erwachsene	75.50 (19.64)	80.27 (17.00)	64.18 (20.19)	75.96 (15.26)
Ältere Erwachsene	73.09 (18.22)	71.15 (18.08)	57.29 (19.26)	57.94 (22.27)

Standardabweichungen in Klammern.

6.7.3.2.6 Validität der Lernzeit als Prädiktor für die Erinnerungsleistung

Auch in diesem Lerndurchgang wurde zusätzlich überprüft, ob die Lernzeit einen validen Hinweisreiz für die spätere Erinnerungsleistung darstellt. Zu diesem Zweck wurden

Gammakorrelationen für jeden Teilnehmer zwischen Lernzeit und Erinnerungsleistung – dieses Mal jedoch getrennt für die beiden Anreizwerte – berechnet (vgl. Tab. 36). *T*-Tests zeigten, dass die gemittelten Korrelationen bei den Grundschulern nur für die 5-Punkt-Items signifikant von Null verschieden waren ($p < .05$), bei den Jugendlichen für beide Anreizkategorien (beide p 's $< .01$), bei den jüngeren Erwachsenen in keinem der beiden Fälle und bei den älteren Erwachsenen sowohl für die 1-Punkt-Items ($p < .01$) als auch für die 5-Punkt-Items ($p < .001$). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Anreizwert als Innersubjektfaktor ergab weder signifikante Haupteffekte noch Interaktionen. Auch hier wurden die Ergebnisse durch eine zweite Analyse überprüft, indem die Erinnerungsleistung für kurz und lang gelernte Items kontrastiert wurde (s. Tab. 43). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor und Lernzeit (über vs. unter dem Median) und Anreizwert als Innersubjektfaktoren ergab einen signifikanten Haupteffekt des Faktors Altersgruppe ($F(3,108) = 5.68$; $p < .01$; $\eta^2 = .14$) in dem Sinne, dass die jüngeren Erwachsenen im Schnitt mehr Items erinnerten als die Grundschüler und die älteren Erwachsenen (s. o.). Weiterhin zeigten sich ein signifikanter Haupteffekt des Faktors Lernzeit ($F(1,108) = 31.62$; $p < .001$; $\eta^2 = .23$), wonach kürzer gelernte Items (70.49%) besser erinnert wurden als länger gelernte (59.18%), sowie ein signifikanter Haupteffekt des Faktors Anreizwert ($F(1,108) = 34.76$; $p < .001$; $\eta^2 = .24$), wonach 1-Punkt-Items (60.10%) schlechter erinnert wurden als 5-Punkt-Items (69.57%; s. o.).

Tabelle 43: Erinnerungsleistung in Prozent für Items mit einer Lernzeit über bzw. unter dem Median in Abhängigkeit von Altersgruppe und Anreizwert im Lerndurchgang 2

	Unter dem Median		Über dem Median	
	1 Pkt.	5 Pkt.	1 Pkt.	5 Pkt.
Grundschüler	53.10 (25.61)	70.83 (22.29)	47.98 (21.24)	59.52 (25.33)
Jugendliche	68.75 (25.31)	75.40 (17.47)	57.54 (25.67)	66.72 (22.67)
Jüngere Erwachsene	76.40 (22.31)	81.49 (20.21)	64.89 (26.91)	77.87 (22.11)
Ältere Erwachsene	64.64 (23.99)	73.33 (19.07)	47.50 (29.78)	51.43 (28.67)

Standardabweichungen in Klammern.

6.7.3.2.7 Fragebogen zum Lernanreiz

Zur Überprüfung der Anreizwirkung der Belohnung wurden die Ergebnisse des Fragebogens zum Lernanreiz mit ANOVAs bzw. einem *t*-Test für die einzelnen Fragen ausgewertet. Für die

Frage 1 („Wie wichtig war es dir/Ihnen, im zweiten Lerndurchgang möglichst viele Punkte zu sammeln?“), zeigte sich, dass die Mittelwerte auf der Skala von 1 bis 7 in allen Altersgruppen recht hoch ausfielen (s. Tab. 44). Die ANOVA mit Altersgruppe als Zwischensubjektfaktor ergab keinen signifikanten Haupteffekt. Bei Frage 2 („Wie wichtig war es dir/Ihnen, möglichst viel von den Extrageschenken/von dem Geld zu bekommen?“) wurde in der ANOVA ein signifikanter Haupteffekt des Faktors Altersgruppe gefunden ($F(3,108) = 14.65$; $p < .001$; $\eta^2 = .29$). Die Post-hoc-Tests zeigten, dass die älteren Erwachsenen hier im Mittel einen deutlich niedrigeren Wert angaben als die drei übrigen Altersgruppen (Grundschüler und jüngere Erwachsene: beide p 's $< .001$; Jugendliche: $p < .01$). Ein ähnliches Muster zeigte sich auch bei Frage 3 („Denkst du/denken Sie, dass die Extrabelohnung dich/Sie motiviert hat, die Wortpaare mit den fünf Punkten besser und länger zu lernen?“). Auch hier gaben die älteren Erwachsenen niedrigere Werte an als die anderen Altersgruppen ($F(3,108) = 10.44$; $p < .001$; $\eta^2 = .23$; Post-hoc-Test: Grundschüler: $p < .001$; Jugendliche: $p < .05$; jüngere Erwachsene: $p < .01$). Der t -Test für die Frage 4 („Wie gut haben dir die Extrageschenke gefallen?“) ergab keine Unterschiede zwischen den Grundschulern und den Jugendlichen.

Tabelle 44: Fragebogen zum Lernanreiz

	Frage 1 „Wichtigkeit der Punkte“	Frage 2 „Wichtigkeit der Belohnung“	Frage 3 „Motivation durch Belohnung“	Frage 4 „Attraktivität der Belohnung“
Grundschüler	5.14 (1.65)	4.04 (1.67)	4.75 (2.37)	5.67 (1.52)
Jugendliche	4.43 (1.85)	3.21 (1.62)	3.43 (1.75)	5.86 (1.16)
Jüngere Erwachsene	5.15 (1.41)	4.00 (1.92)	3.71 (2.34)	---
Ältere Erwachsene	4.25 (2.27)	1.54 (1.17)	1.82 (1.28)	---

Standardabweichungen in Klammern.

6.7.4 Diskussion

In der letzten Studie dieser Arbeit ging es darum, die Befunde aus Studie 5 zum CM-Modell in den vier beteiligten Altersgruppen zu replizieren und erstmals zu untersuchen, inwieweit ein flexibler Wechsel zwischen MC- und CM-Modell im Verlauf der Lebensspanne möglich ist.

Auch in dieser Studie musste die grundlegende Aufgabenschwierigkeit zwischen den Altersgruppen angepasst werden. Es wurde dabei auf die Listen der Studie 5 (Lerndurchgang

2) zurückgegriffen. Lediglich die Liste für die jüngeren Erwachsenen wurde um 12 Wortpaare erweitert mit dem Ziel, den Unterschied zwischen dieser und den anderen Gruppen weiter anzugleichen. Auf diese Weise konnten wieder Boden- und Deckeneffekte in der Erinnerungsleistung vermieden werden. Die Erinnerungsleistung lag bei den Grundschulern und den älteren Erwachsenen im Lerndurchgang 2 etwas höher als im ersten Durchgang, was zum einen durch Übungseffekte, zum anderen mit einer erhöhten Motivation durch die Punkte bzw. die Belohnung erklärt werden kann. Nicht im Zentrum der Fragestellung, aber dennoch relevant für den Vergleich mit den Überwachungsleistungen in den anderen Experimenten dieser Arbeit ist die Differenzierungsfähigkeit und die Genauigkeit in den JOLs und SUs. Für beide Lerndurchgänge und für sämtliche Altersgruppen waren beide Aspekte der Überwachungsleistung sehr gut ausgeprägt; bei den SUs ergaben sich tendenziell noch bessere Leistungen als bei den JOLs (vgl. Dunlosky & Metcalfe, 2009; Schneider, 2015).

Die Ergebnisse im Lerndurchgang 1 zum CM-Modell konnten im Wesentlichen die Befunde des zweiten Lerndurchgangs in Studie 5 bestätigen. Es wurden in der Auswertung nach dem Mediansplit der Lernzeit in allen Altersgruppen signifikante Unterschiede in den JOLs zwischen länger und kürzer gelernten Items gefunden, was darauf hindeutet, dass die Lernzeit einen wichtigen Hinweisreiz für die JOLs darstellt. Die Analyse nach Korrelationen zwischen Lernzeit und JOLs ergab dagegen Alterseffekte in dem Sinn, dass nur bei den älteren drei Gruppen signifikante Korrelationen gefunden wurden. Die Werte der älteren Erwachsenen lagen dabei am höchsten. Der in Experiment 5 gefundene Alterstrend, wonach das CM-Modell bei den älteren Erwachsenen weniger stark ausgeprägt ist als bei den jüngeren Erwachsenen, konnte also hier nicht belegt werden. Dieser Unterschied war jedoch auch in der vorherigen Studie nur geringfügig ausgeprägt. Eine breitere Datenbasis mit mehr Teilnehmern und mehr Trials könnte klären, ob exekutive Defizite bei älteren Erwachsenen datenorientiertes Lernen erschweren und welche Faktoren interindividuelle Unterschiede in dieser Altersgruppe bedingen (Pansky et al., 2009; Souchay & Isingrini, 2004). Auch bei den Grundschulern wurden große Streuungen in der Korrelation zwischen JOLs und Lernzeit im ersten Lerndurchgang gefunden, sodass auch hier weitere Studien nötig sind, um feststellen zu können, was Personen, die höhere Zusammenhänge aufweisen, von denen mit niedrigeren unterscheidet. Generell entsprechen die Ergebnisse für die Grundschüler und Jugendlichen denen der Vorstudie und der bislang vorliegenden Literatur, dass ab etwa der dritten Klasse Nachweise für das CM-Modell gefunden werden können (Koriat et al., 2009b) und bis zum Jugendalter die Fähigkeit, Informationen aus Kontrollprozessen für die Anpassung von Überwachungsvorgängen zu nutzen, zunimmt (Hoffmann-Biencourt et al., 2010).

Im Lerndurchgang 2 zeigte sich bezüglich der Zielorientierung, dass die Teilnehmer aller Altersgruppen bessere Erinnerungsleistungen für höher als für niedriger bewertete Items erreichten. Die Lernzeit war jedoch nur bei den Jugendlichen und den jüngeren Erwachsenen für die 5-Punkte-Items länger als für die 1-Punkt-Items. In der Anpassung der JOLs nach den Punktwerten zeigten sich deutliche Alterseffekte: Die jüngeren Erwachsenen vergaben für die höher bewerteten Items hoch signifikant höhere mittlere JOLs als für die niedriger bewerteten. Dieser Unterschied war bei den Grundschulern und den Jugendlichen ebenfalls statistisch bedeutsam, jedoch geringer ausgeprägt. Die älteren Erwachsenen unterschieden in ihren JOLs nicht zwischen 1- und 5-Punkte-Items. Das heißt, während in Studien, die zielorientiertes Lernen separat untersuchen, sowohl für Grundschüler als auch für ältere Erwachsene mehrheitlich Belege für das MC-Modell gefunden wurden (z.B. Froger et al., 2012; Hines et al., 2012; Lockl & Schneider, 2003; vgl. auch Studie 5), fällt dies besonders den Senioren schwerer, wenn parallel datenorientiertes Lernen erfasst wird. Dieser Befund kann nicht dadurch erklärt werden, dass die Teilnehmer nicht ausreichend motiviert waren, viele Punkte zu erreichen. Zwar gaben die älteren Erwachsenen an, durch die monetäre Belohnung nicht oder nur wenig beeinflusst worden zu sein, jedoch waren die Teilnehmer überwiegend und im gleichen Ausmaß wie die anderen Gruppen daran interessiert, möglichst viele Punkte zu erreichen. Bei den Grundschulern zeigte sich zusätzlich, dass die Geschenke für die meisten Teilnehmer sehr attraktiv waren. Die älteren Probanden schienen vielmehr wie bei Price et al. (2010) weniger sensitiv auf die Anreizwerte zu reagieren, wenn gleichzeitig die Itemschwierigkeit manipuliert wurde. Dies könnte auch für die vorliegende Studie mit den Zielen der älteren Erwachsenen erklärt werden: Wenn es ihnen insgesamt schwerer fiel, sich schwierige Wortpaare zu merken, ist es eine adaptive Strategie, sich unabhängig vom Behaltenswert auf die einfacheren Wortpaare zu konzentrieren. Weiterhin galt auch für die Grundschüler, wenn auch in einem geringeren Maße als für die älteren Erwachsenen, dass diese stärker auf den deutlich salienteren Aspekt der Itemschwierigkeit als auf die Anreizwerte in der Kontrolle ihres Lernverhalten reagierten (vgl. Castel, Lee et al., 2011; Koriat et al., 2006).

Bezüglich des datenorientierten Lernens zeigte der zweite Lerndurchgang analog zu Studie 5 und zum ersten Lerndurchgang, dass die Teilnehmer aller Altersgruppen prinzipiell höhere JOLs nach kürzer gelernten Items abgaben als nach länger gelernten. Bei den Korrelationen zwischen Lernzeit und JOLs ergaben sich jedoch leichte Alterseffekte in dem Sinn, dass nur bei den Jugendlichen und den älteren Erwachsenen, bei den jüngeren Erwachsenen für die 1-Punkt-Wortpaare, nicht aber bei den Grundschulern die Korrelationen signifikant von Null verschieden waren. Besonders die Grundschüler schienen demnach Schwierigkeiten beim

datenorientierten Lernen aufzuweisen. In allen Altersgruppen war die Lernzeit im Wesentlichen ein valider Prädiktor für die Erinnerungsleistung.

Insgesamt gesehen zeigen die Ergebnisse, dass die Jugendlichen und die jüngeren Erwachsenen in der Lage waren, flexibel zwischen ziel- und datenorientiertem Lernen auch innerhalb einer einzigen Aufgabe zu wechseln (vgl. Koriat, Ackerman et al., 2014). Die Grundschüler und die älteren Erwachsenen wiesen hier mehr Schwierigkeiten auf. Besonders die älteren Erwachsenen richteten ihre JOLs stark nach Hinweisreizen, die sich aus Charakteristika der Items und der Lernzeit ergaben (CM-Modell) – umgekehrt fiel es ihnen schwerer, gleichzeitig das Ziel, viele Punkte zu erreichen, strategisch umzusetzen. Bei den Drittklässlern wurden im Lerndurchgang 2 etwas weniger Nachweise für ziel-, aber auch für datenorientiertes Lernen gefunden als bei den beiden mittleren Altersgruppen. Der flexible Wechsel zwischen beiden Lernformen schien sich hier auf beide Prozesse auszuwirken. Dies entspricht Vorbefunden, die nahe legen, dass die Sensitivität, sowohl auf ziel-, als auch auf datenorientierte Prozesse gleichzeitig zu reagieren, mit dem Alter ansteigt (Koriat, Ackerman et al., 2014). Die Entwicklungsverläufe vom Kindes- zum jüngeren Erwachsenenalter und vom diesem zum höheren Erwachsenenalter scheinen nicht vollständig spiegelbildlich zu sein. Während bei den älteren Erwachsenen ihr persönliches Lernziel das flexible Reagieren auf beide Prozesse zu erschweren scheint (vgl. Ariel et al., 2009), fällt es den Grundschulern womöglich schwer, Bottom-up- und Top-down-Prozesse (d.h. aus der Lernzeit auf die JOLs zu schließen und vice versa) in einem einzigen metakognitiven Urteil zu integrieren (Koriat, Ackerman et al., 2014). Ob diese Vermutung richtig ist, muss in weiteren Studien mit einer direkten Kontrastierung dieser beiden Altersgruppen geklärt werden. Hierbei könnte es aufschlussreich sein, den Probanden verschiedene Lernziele vorzugeben bzw. ihre eigenen abzufragen, oder den Wechsel zwischen CM- und MC-Modell nur bei mittelschweren Wortpaaren zu untersuchen und somit die Salienz der Punktwerte besonders für diese beiden Altersgruppen zu steigern (vgl. Ariel et al., 2009; Koriat et al., 2006; Koriat, Ackerman et al., 2014).

Selbst wenn durch die Studien 5 und 6 sowohl das MC- als auch das CM-Modell prinzipiell für alle Altersgruppen – mit Verbesserungen zum jüngeren Erwachsenenalter und leichten Verschlechterungen zum älteren Erwachsenenalter hin – nachgewiesen werden konnten, scheint ein flexibler Wechsel zwischen beiden Modellen besonders den jüngsten und den ältesten Probanden schwer zu fallen. Für die Analyse genauerer Entwicklungsprozesse ist der Einschluss von mehr Altersgruppen in feineren Abstufungen nötig sowie weitere Studien zu detaillierteren Grundlagen beider Prozesse.

6.8 Zusammenfassung

In sechs umfangreichen Studien wurden Entwicklungsverläufe in metakognitiven Überwachungsprozessen sowie der wechselseitige Zusammenhang zwischen Überwachung und Kontrolle über die Lebensspanne hinweg untersucht. Erstmals wurden in diesem Forschungsgebiet die Leistungen von Grundschulern, Jugendlichen, jüngeren und älteren Erwachsenen in einer Vielzahl an Aufgaben und Versuchsdesigns unmittelbar verglichen.

In Studie 1 zeigte sich, dass Befunde zur metakognitiven Überwachung, die aus Studien mit enger gefassten Altersvergleichen stammen, auch auf ein lebensspannenübergreifendes Paradigma übertragbar sind. Weiterhin konnte nachgewiesen werden, dass prospektive und retrospektive Überwachungsprozesse bei einfacher strukturiertem Material altersunabhängig differenzierter ausfallen als solche bei komplexerem, alltagsnäherem Material.

Die Studien 2 und 3 untersuchten den Einfluss des bereichsspezifischen Vorwissens auf die Überwachungsleistung. Es stellte sich heraus, dass der Vorteil der Experten gegenüber den Novizen, der für die SU-Genauigkeiten in Studie 2 beim Lernen von fußballbezogenen Wortpaaren gefunden wurde, deskriptiv auch für die SUs in Studie 3 beim Lernen aus einem Text galt. Die Differenzierungsleistung in den JOLs war bei den Experten unter den Drittklässlern und den älteren Erwachsenen für das Beantworten von Fragen aus dem Fußballtext sogar etwas schlechter als bei den Novizen. Insgesamt gesehen zeigten beide Studien, dass die Teilnehmer mit einem höheren Vorwissen ihre eigenen Leistungen optimistischer einschätzten als diejenigen mit einem geringeren Vorwissen.

In Studie 4 wurde geprüft, inwiefern ein Training in einer Elaborationsstrategie beim Lernen von Wortpaaren positive Auswirkungen auf die Überwachungsleistung hat. Die positiven Effekte dieser Intervention waren für die SUs etwas stärker ausgeprägt als für die JOLs. Weiterhin profitierten die Grundschüler und die älteren Erwachsenen am meisten vom Training. Es konnte gezeigt werden, dass die Anwendung der Strategie die Nutzung von vertrautheitsbezogenen und somit weniger validen Hinweisreizen bedeutsam reduzierte.

Die Studien 5 und 6 gingen der Frage nach, wie Überwachungs- und Kontrollvorgänge sich im Lauf des Lebens wechselseitig beeinflussen. In Studie 5 wurden in einem sequenziellen Design in allen Altersgruppen sowohl Nachweise für das MC-Modell (d.h. Personen nutzen beim zielorientierten Lernen Informationen aus den JOLs, um ihre Lernzeit anzupassen) als auch für das CM-Modell (d.h. Personen nutzen beim datenorientierten Lernen Informationen aus der Lernzeit, um ihre JOLs anzupassen) erbracht. Alterseffekte zeigten sich dahingehend, dass die

jüngeren Erwachsenen in beide Wirkrichtungen zumindest deskriptiv etwas höhere Zusammenhänge aufwiesen. In Studie 6 konnten im ersten Lerndurchgang die Ergebnisse aus Studie 5 zum CM-Modell im Wesentlichen bestätigt werden. Für den simultanen Wechsel zwischen ziel- und datenorientiertem Lernen zeigte sich, dass hier die Gruppe der Grundschüler und der älteren Erwachsenen schlechter abschnitt als die Jugendlichen und die älteren Erwachsenen.

Da insgesamt eine Vielzahl einzelner Fragestellungen untersucht wurde, sollen die Ergebnisse in der folgenden Gesamtdiskussion sowohl unter dem Gesichtspunkt der Verortung im Modell des prozeduralen Metagedächtnisses von Nelson und Narens (1990, 1994) als auch im Hinblick auf den Einfluss der untersuchten Kontextfaktoren zusammenfassend betrachtet werden.

7. GESAMTDISKUSSION

Die vorliegende Arbeit verfolgte mehrere Ziele. Übergeordnet war das Anliegen, bereits bestehende Forschungsparadigmen und neuere Fragestellungen zum prozeduralen Metagedächtnis über die Lebensspanne hinweg zu erfassen. Es sollten also Entwicklungsverläufe über einen großen Altersbereich hinweg betrachtet werden, anstatt – wie oft üblich – lediglich zwei oder drei enger gefasste Altersgruppen zu kontrastieren. Aus diesem Grund wurden Grundschüler der dritten Klasse, Jugendliche der siebten und achten Jahrgangsstufe, jüngere und ältere Erwachsene in die Erhebungen aufgenommen.

Die einzelnen Studien hatten spezifischere Zielsetzungen. So sollte im Experiment 1 grundsätzlich geklärt werden, inwiefern Überwachungsprozesse bei einfacher strukturierten (hier: Bildpaare) und komplexeren, alltagsnäheren Stimuli (hier: Filmmaterial) in den unterschiedlichen Altersgruppen vergleichbar sind. Die Studien 2 und 3 nahmen die Frage nach dem Vergleich von verschiedenen komplexen Materialien (Wortpaare und Text) auf. Hier lag jedoch der Fokus auf dem Einfluss des bereichsspezifischen Vorwissens am Beispiel der Domäne Fußball auf die Überwachungsleistungen. Studie 4 bezog sich auf einen weiteren möglichen Einflussfaktor auf die Güte der Überwachung: Es wurde untersucht, inwiefern ein Strategietraining in den verschiedenen Altersgruppen die Leistungen im prozeduralen Metagedächtnis verändert. In den Studien 5 und 6 wurden zusätzlich Steuerungsprozesse erhoben, um erstmals in einer breiten Altersspanne wechselseitige Zusammenhänge zwischen Überwachung und Steuerung sowohl sequenziell als auch simultan zu untersuchen.

Das Versuchsdesign orientierte sich im Wesentlichen am Modell zum prozeduralen Metagedächtnis von Nelson und Narens (1990, 1994): Der Ablauf folgte den einzelnen Phasen des Lernens mit den jeweils zugehörigen pro- und retrospektiven Überwachungsurteilen (Globales Verständnisurteil, EOLs, JOLs, SUs). In den Experimenten 5 und 6 wurde als Maß für die Steuerung die Lernzeitallokation gewählt.

7.1 Überwachungsleistungen im Entwicklungsverlauf

In diesem Abschnitt soll zusammenfassend diskutiert werden, welche Rückschlüsse die Ergebnisse der sechs Experimente auf Entwicklungsprozesse in den verschiedenen erhobenen Überwachungsmaßen zulassen.

7.1.1 Globales Verständnisurteil

Überwachungsurteile wie EOLs und ein globales Verständnisurteil, die vor dem eigentlichen Lernvorgang erhoben werden, kamen nur in den Studien 2 und 3 zum Einfluss des Vorwissens auf die Metagedächtnisleistung zum Einsatz. Die Entscheidung, diese zusätzlich zu den häufiger erfassten Maßen JOL und SU aufzunehmen, war in der Überlegung begründet, dass sich gerade bei diesen Urteilen Unterschiede zwischen Personen mit viel bzw. wenig domänenspezifischem Wissen abbilden lassen müssten. Denn zu diesem Zeitpunkt sollten sich Annahmen über die eigene Leistung, die auf der selbst eingeschätzten Expertise beruhen, am stärksten auswirken, da hier im Vergleich zu JOLs und SUs weniger Informationen aus dem Lern- und Erinnerungsvorgang einfließen.

In Studie 3 wurde darauf verzichtet, wie in anderen einschlägigen Studien ein globales Urteil nach der Präsentation der Überschrift bzw. nach einem ersten Überfliegen des Textes zu erheben (Shiu & Chen, 2013; Thiede et al., 2003), um den Teilnehmern eine validere Basis für ihr Urteil zu bieten. Das globale Verständnisurteil wurde hier nach der zweifachen Präsentation des Textes (visuell und auditiv) erfasst.

Es konnten für dieses Urteil weder signifikante Beziehungen zur späteren Leistung noch einheitliche Ergebnisse zu Altersverläufen gefunden werden. Nur eine einzige Korrelation war bedeutsam von Null verschieden. Die Ergebnisse sind damit mit anderen Befunden zur Erhebung von Globalurteilen aus der Metacomprehension-Forschung vergleichbar (z.B. Burkett & Azevedo, 2012; Shui & Shen, 2013; Thiede et al., 2003). Ergebnisse aus diesen Arbeiten mit jüngeren Erwachsenen scheinen demnach auch für andere Altersgruppen zu gelten. Da davon auszugehen ist, dass eine Einschätzung des späteren Lernerfolgs wesentliche Auswirkungen auf das Lernverhalten hat (Nelson & Leonesio, 1988), ist es wichtig, auch weiterhin Anstrengungen zu unternehmen, validere Formulierungen zu finden, die genauere globale Vorhersagen bei komplexeren Materialien ermöglichen.

7.1.2 Ease-of-Learning-Urteile

Für EOLs existiert bislang ähnlich wie bei den Globalurteilen keine einheitliche Operationalisierung (vgl. Abschnitt 3.1.2); so sollten beispielsweise die Teilnehmer bei Leonesio und Nelson (1990) die Items in eine Rangfolge nach subjektiver Schwierigkeit bringen, bei Kelemen et al. (2000) fand die Abfrage dagegen vergleichbar mit der Erfassung in den eigenen Studien statt, also mit der Frage „Wie schwer wäre dieses Item für Sie zu lernen?“.

Für die EOLs in Studie 2 zeigte sich, dass die Teilnehmer aller Altersgruppen prinzipiell in der Lage waren, zwischen später richtigen und später falschen Antworten bereits vor dem eigentlichen Lernvorgang zu differenzieren, d.h. sie gaben im Mittel höhere EOLs vor richtigen als vor falschen Antworten. Im Vergleich zu den in den eigenen Experimenten festgestellten und in der Literatur (z.B. von der Linden et al., 2011) berichteten Differenzierungsleistungen bei JOLs und SUs fällt auf, dass die durchschnittlichen Werte bei den EOLs vor falschen Antworten höher auszufallen scheinen (4.53 auf einer 7-Punkte-Skala): Auch bei später falschen Antworten waren die Teilnehmer – obwohl die Werte substanziell niedriger lagen als bei richtigen Antworten – relativ zuversichtlich, diese erinnern zu können. Es fällt jedoch schwer, diese optimistische Einschätzung der Teilnehmer mit der Literatur zu EOLs in Bezug zu setzen. Dies liegt zum einen an der genannten Vielfalt der Operationalisierungen, zum anderen daran, dass bislang keine Studie vorliegt, die Überwachungsleistungen in EOLs mithilfe der Differenzierung auswertet. Die hohen Einschätzungen vor falschen Antworten decken sich jedoch mit Studien, die die Kalibrierung von EOLs untersuchen: Besonders bei Kindern, aber auch bei Jugendlichen wurde eher eine Überschätzung der eigenen Leistung gefunden (Labuhn et al., 2010; Visé & Schneider, 2000). Bezüglich der älteren Erwachsenen ist die Datenlage zu EOLs bislang spärlich: Hier finden Kalibrierungsstudien sowohl Ergebnisse, die eher für eine Überschätzung der eigenen Leistung (z.B. Rabbit & Abson, 1991), als auch solche, die eher für eine Unterschätzung der eigenen Leistung sprechen (z.B. Hertzog et al., 1994). Die Diversität der Befunde im höheren Erwachsenenalter scheint durch die generelle Tendenz von Probanden, ihre Vorhersagen an einem mittleren Wert zu orientieren, erklärbar zu sein (vgl. Hertzog et al., 1994): Die Altersgruppe, deren tatsächliche Leistung näher an der Mitte liegt, schätzt sich folglich genauer ein. Da in der eigenen Studie Item-by-Item-Urteile abgegeben wurden und der Versuch unternommen wurde, die Aufgabenschwierigkeit zwischen den Altersgruppen so weit wie möglich anzupassen, sollte diese „Tendenz zur Mitte“ hier recht gering ausgeprägt sein, sodass die hohen EOLs vor falschen Antworten bei der Gruppe der älteren Erwachsenen eher für die Schwierigkeit spricht, den eigentlichen Lernprozess zu antizipieren und sich eher auf Faktoren wie die Vertrautheit mit den Stimuli zu konzentrieren (vgl. Hertzog et al., 1990). Für diese Annahmen spricht auch, dass die jüngeren Erwachsenen, also die Gruppe, die am meisten Erfahrung mit dem Bearbeiten eher abstrakter Aufgaben hat, im Vergleich zu den anderen Altersgruppen bessere Differenzierungsleistungen erbrachten.

Die Gammakorrelationen lassen sich ebenfalls nur mit wenigen Studien direkt vergleichen. Im eigenen Experiment lagen die Leistungen zwischen .25 (Experten und Novizen unter den

Grundschulern) und .63 (Experten unter den jüngeren Erwachsenen), was in etwa den von Kelemen et al. (2000) für jüngere Erwachsene berichteten Werten entspricht. Insgesamt fällt die prädiktive Validität von EOLs also gering bis mittelhoch, wenn auch in den meisten Fällen substanziell aus. Da die eigene Studie eine der wenigen ist, die den Altersverlauf bei EOLs mithilfe von Gammakorrelationen bestimmt, muss die Interpretation mit der entsprechenden Vorsicht vorgenommen werden. Alterseffekte zeigten sich in der Hinsicht, dass die jüngeren Erwachsenen signifikant höhere mittlere Korrelationen zwischen EOLs und Erinnerungsleistung aufwiesen als die Grundschüler. Die Erklärung, dass jüngere Erwachsene weniger Schwierigkeiten haben, den zukünftigen Lernprozess gedanklich vorwegzunehmen, könnte auch hier greifen. Jedoch können auch andere Erklärungen wie z.B. die Vertrautheit mit dem Material, die wahrgenommene Itemschwierigkeit oder die aufgabenunabhängige Selbsteinschätzung eigener Fähigkeiten als Einflussfaktoren nicht ausgeschlossen werden.

7.1.3 Judgments-of-Learning

Im Gegensatz zu den EOLs ist die Abfrage von JOLs in der Literatur einheitlicher. In vielen Fällen hat sich die Frage „Wie sicher sind Sie, dass Sie diese Information später erinnern werden?“ oder eine vergleichbare Formulierung durchgesetzt (z.B. Hertzog et al., 2010; Hoffmann-Biencourt et al., 2010; von der Linden et al., 2011). Da in dieser Arbeit in allen sechs Experimenten JOLs erhoben wurden, ist ein umfassenderer Vergleich möglich als für die EOLs. Alle JOLs wurden zeitlich verzögert erfasst, da diese deutlich genauer sind als unmittelbar erhobene (Nelson & Dunlosky, 1991), um so mögliche zusätzliche Effekte von Kontextbedingungen auf die Qualität der JOLs zu prüfen. Die JOLs beziehen sich auf Erinnerungsleistungen, die das Ergänzen eines zweiten Wortpaares (Studien 1, 2, 5 und 6), das Wiedererkennen von Wortpaaren (Studie 4) oder die freie Beantwortung von Fragen (Studien 1 und 3) umfassen.

Über alle Studien hinweg ergaben sich sehr gute Differenzierungsleistungen für alle Altersbereiche. In jedem Experiment wurde der Haupteffekt des Faktors Korrektheit der Antwort hoch signifikant (alle p 's < .001). Auch die partiellen Eta-Quadrat-Werte lagen sämtlich im sehr hohen Bereich (vgl. Cohen, 1988): Der im Vergleich niedrigste Wert von $\eta^2 = .69$ für den Haupteffekt dieses Faktors wurde in Studie 3 gefunden, der höchste Wert von $\eta^2 = .90$ in Studie 1. Die Werte vor später richtigen Antworten waren über alle Experimente hinweg im oberen Bereich der jeweiligen Skala angesiedelt. Die Mittelwerte vor falschen Antworten lagen in Studie 3 mit 4.05 von 7 Punkten eher hoch und vergleichbar mit den EOLs,

für alle anderen Studien jedoch eher am unteren Ende der Skala (im Mittel um 3.20 bei den 7-Punkte-Skalen und um 36.00 bei den 100-Punkte-Skalen). Die Teilnehmer waren also prinzipiell sehr gut und deutlich besser als bei den EOLs in der Lage, bei der verzögerten Abfrage zwischen später erinnerten und nicht bzw. falsch erinnerten Informationen zu unterscheiden. Da insgesamt nur recht wenig Arbeiten die Differenzierungsfähigkeit bei JOLs untersuchen, fällt es schwer, eine Einordnung vorzunehmen. Zwei Studien mit Grundschulern und jüngeren Erwachsenen zum inzidentellen Lernen weisen geringere F-Werte und Effektstärken zum Einfluss des Faktors Korrektheit der Antwort auf; aber auch hier wurden alle einschlägigen Haupteffekte hoch signifikant (von der Linden et al., 2011; Roebers, von der Linden, Schneider & Howie, 2007). Die höheren Werte in den eigenen Studien lassen sich dadurch erklären, dass schon beim Lernvorgang Überwachungsvorgänge bei den Probanden stärker aktiviert waren, da diese hier im Gegensatz zu den zitierten Arbeiten wussten, dass ihre Erinnerungsleistung später abgefragt werden würde.

Die Gammakorrelationen zwischen JOL und Erinnerungsleistung fielen über alle sechs Studien hinweg recht hoch aus. Von den insgesamt 48 Einzelwerten in den untersuchten Zellen waren bis auf eine Ausnahme sämtliche Werte signifikant größer als Null, davon 45 Werte mit einer Wahrscheinlichkeit von $p < .001$. Je nach Art der Aufgabe (s. unten) waren mittelhohe bis nahezu perfekte Korrelationen zu beobachten. Die einzelnen Ergebnisse stehen damit im Einklang mit der Literatur zur JOL-Genauigkeit, wonach die zeitlich verzögerte Erhebung von JOLs sehr zutreffende Leistungsvorhersagen liefert (z.B. Connor et al., 1997; Nelson & Dunlosky, 1991; Schneider et al., 2000). Dies zeigte bislang vor allem die Literatur zum Lernen von Paarassoziationen (für einen Überblick s. Rhodes & Tauber, 2011). So wurden auch in den eigenen Studien bei den Aufgaben zum Lernen von Wortpaaren Korrelationen zwischen .66 und .94 gefunden. Die Aufgaben zum Lernen aus Texten und Filmen und zum Wiedererkennen erbrachten niedrigere Korrelationen zwischen .33 und .90. Diese Abweichungen von den Genauigkeiten in klassischen Paar-Assoziationsdesigns sollen in Abschnitt 7.2.1 ausführlicher diskutiert werden.

Alterseffekte, die sich ausschließlich auf die Höhe der JOLs oder die Differenzierungsleistung beziehen und nicht in einer Dreifachinteraktion mit einem zusätzlichen Faktor (Vorwissen, Strategieinstruktion) aufgehen, kamen kaum vor. In Studie 1 gaben die Jugendlichen – unabhängig von der Korrektheit der Antwort – etwas niedrigere JOLs ab als die Grundschüler und die älteren Erwachsenen. Da dieser Effekt in den fünf weiteren Studien nicht auftrat, ist nicht davon auszugehen, dass die Jugendlichen generell weniger optimistisch in ihrer Leistungseinschätzung sind als andere Altersgruppen. In Studie 5 zeigte sich im Lerndurchgang

1 zum MC-Modell, dass die älteren Erwachsenen etwas besser differenzierten als die jüngeren. Auch diese Tendenz blieb auf ein einziges Teilergebnis beschränkt und scheint deshalb nicht generalisierbar zu sein. Alterseffekte in den Gammakorrelationen zeigten sich lediglich in einer von acht Analysen (Studie 3), sodass insgesamt davon ausgegangen werden kann, dass die Überwachungsgenauigkeit in den JOLs im Gegensatz zu den EOLs zumindest ab dem mittleren Grundschulalter für verzögerte Urteile über die Lebensspanne hinweg recht konstant bleibt.

Die weitgehend fehlenden Alterseffekte sind in Einklang mit Vorbefunden, die schon bei Grundschulern von einer guten Überwachungsfähigkeit in den JOLs ausgehen (z.B. Roebbers, von der Linden, Schneider & Howie, 2007; Schneider et al., 2000) und für das höhere Erwachsenenalter in der Regel keine Verschlechterungen nachweisen (z.B. Hertzog et al., 2002). Die hier im Gegensatz zu anderen Studien nicht gefundenen Verbesserungen in der JOL-Qualität vom Grundschul- zum Jugendalter (Hoffmann-Biencourt et al., 2010; Koriat et al., 2009a) können darauf zurückgeführt werden, dass die Bemühungen, das Niveau bzw. die Anforderung der Lernaufgaben durch die Anpassung von Präsentationszeiten, Zahl der Items, Länge der Texte und Schwierigkeiten der Fragen zwischen den Altersgruppen möglichst konstant zu halten, insgesamt gelungen sind. Alterskorrelierte Unterschiede können in den anderen Studien auf eine erschwerte Verarbeitung und somit weniger kognitive Kapazität für die Überwachungsvorgänge in den älteren Versuchsgruppen zurückgeführt werden (vgl. de Bruin et al., 2007).

7.1.4 Sicherheitsurteile

Die SUs wurden in jeder Studie mit der Frage „Wie sicher sind Sie, dass Ihre Antwort richtig ist?“ erfasst. Diese Formulierung entspricht ebenfalls den gängigen Formaten in der Literatur (z.B. Destan et al., 2014; Roebbers, 2002; Roebbers et al., 2004).

Die Differenzierungsfähigkeit bei den SUs lag tendenziell noch höher als bei den JOLs. Der Haupteffekt des Faktors Korrektheit der Antwort wurde wieder in allen Fällen hoch signifikant (alle p 's < .001), in sechs von acht Auswertungen waren die zugehörigen η^2 -Werte größer als .86. In Studie 3 und in Studie 4 lagen die Effektstärken etwas niedriger ($\eta^2 = .72$ bzw. $\eta^2 = .56$), aber dennoch im sehr hohen Bereich. Die Ergebnisse entsprechen im Wesentlichen denen der wenigen Studien, die die Differenzierungsfähigkeit bei SUs im Grundschul- und jüngeren Erwachsenenalter analysieren und ebenfalls hohe Effektstärken finden (von der Linden & Roebbers, 2006; Roebbers, von der Linden & Howie, 2007; Roebbers, von der Linden, Schneider

& Howie et al., 2007). Zu einer besseren Einordnung der eigenen Ergebnisse sind jedoch auch für die SUs noch weitere Vergleichsstudien zur Differenzierungsfähigkeit nötig, besonders für die Altersgruppen der Jugendlichen und der älteren Erwachsenen.

Die Gammakorrelationen in den SUs waren ebenfalls mehrheitlich höher als die der EOLs und JOLs. In den Studien 1, 2, 5 und 6 lagen die mittleren Werte in den einzelnen Bedingungen zwischen .67 und der nahezu perfekten Korrelation von .97. In den Studien 3 und 4 lagen die mittleren Gammakorrelationen etwas niedriger als in den anderen Studien, jedoch bis auf die Werte in einzelnen Altersgruppen sämtlich im hoch signifikanten Bereich (p 's < .001). Da die Resolution von SUs bislang deutlich seltener erfasst wurde als die Kalibrierung, liegen auch hier wenige Studien zum intentionalen Lernen vor, mit denen die eigenen Genauigkeitsparameter verglichen werden können. Da diese jedoch in großer Mehrheit hohe bis sehr hohe Gammakorrelationen finden (z.B. Dunlosky, Rawson & Hacker, 2002; Koriat & Goldsmith, 1996; von der Linden et al., 2011; M. D. Robinson et al., 2000) sind die eigenen Befunde dort gut einzuordnen.

Alterseffekte waren auch für die SUs über alle Studien hinweg selten und nicht konsistent. Sie lassen sich also nicht einem einheitlichen Entwicklungsverlauf in den SUs zuordnen. Die vorliegende Literatur zu SUs lässt hingegen eher vermuten, dass sich stärkere Alterseffekte als bei den JOLs zeigen. So finden einige Studien eine Verbesserung der Differenzierungsleistung in den SUs über das Grundschulalter bis zum Jugendalter (Krebs & Roebbers, 2012; von der Linden & Roebbers, 2006; Roderer & Roebbers, 2010; Roebbers, 2002). Diese Studien sind mit den eigenen nur bedingt vergleichbar, da es sich hier mehrheitlich um Studien aus dem Feld der Augenzeugengedächtnisforschung handelt, d.h. es kommen eine inzidentelle Lerninstruktion, irreführende Fragen und ausschließlich komplexere Materialien zum Einsatz. Für die Gruppe der Grundschüler womöglich am ehesten ist die Aufgabe in Studie 3 zum Wiedererkennen von Sätzen den „irreführenden Fragen“ zuzuordnen, da in der Regel nur wenige Wörter im Vergleich zum Originaltext geändert wurden. Hier zeigten die Grundschüler entsprechend auch schlechtere Leistungen als die Jugendlichen und die jüngeren Erwachsenen. Die Befundlage zu den älteren Erwachsenen divergiert in der Literatur insgesamt deutlicher: Je nach Aufgabenstellung und Design werden keine altersbezogenen Defizite in der Überwachung der Sicherheit gefunden (z.B. Dodson et al., 2007; Hines et al., 2009) oder aber Verschlechterungen bei älteren im Vergleich zu jüngeren Erwachsenen (z.B. Kelley & Sahakyan, 2003; Shing et al., 2009). Auch bei den älteren Erwachsenen beruhen Verschlechterungen in der SU-Qualität zum Teil auf dem Einsatz irreführender Fragen und der Tatsache, dass sich die Senioren eher an der Vertrautheit mit den Items orientieren, nicht aber an tatsächlichen Erinnerungsspuren (z.B.

Kelley & Sahakyan, 2003). Diese Tatsache konnte in der Studie 3 beim Wiedererkennen der Sätze aus dem Fußballtext beobachtet werden: Die älteren Erwachsenen konnten hier nicht mit Sicherheit unterscheiden, welche Sätze wörtlich im Originaltext vorkamen und welche nicht. Dies drückte sich ähnlich wie in der Studie von Shing et al. (2009) in hohen SUs nach falschen Antworten aus ($M = 5.48$ auf einer 7-Punkte-Skala).

7.1.5 Fazit

Insgesamt korrespondieren die Überwachungsleistungen in den einzelnen Urteilen im Wesentlichen mit der einschlägigen Literatur. Es ist also gelungen, Vorbefunde zur metakognitiven Überwachung in einem lebensspannenübergreifenden Design zu replizieren und bedeutsam zu erweitern.

Studienübergreifend zeigte sich, dass die Überwachungsleistungen umso besser wurden, je weiter fortgeschritten der Lern- und Erinnerungsprozess war. Das heißt, dem Cue-utilization-Ansatz entsprechend (Koriat, 1997) fließen in prospektive Urteile, die vor dem eigentlichen Lernvorgang stattfinden (EOLs), Vorannahmen über die eigene Leistungsfähigkeit (Zhao & Linderholm, 2008) und intrinsische Hinweisreize wie die Itemschwierigkeit, die Bildhaftigkeit oder die Auftretenshäufigkeit (Jönsson und Lindström, 2010; Underwood, 1966) ein. JOLs, die mit einer zeitlichen Verzögerung nach dem Lernvorgang erfasst werden, beruhen dann zusätzlich auf extrinsischen Hinweisreizen, d.h. Informationen aus der Lernsituation wie z.B. der Präsentationsdauer, und auf mnemonischen Hinweisreizen, die aus der Lernerfahrung resultieren, also beispielsweise der Leichtigkeit der Verarbeitung (Koriat, 1997). In SUs schließlich gehen dann auch Informationen aus dem Speicher- und Abrufprozess ein. Hier sind ebenfalls mnemonische Hinweisreize von Bedeutung, wie z.B. die Leichtigkeit des Abrufs, die Antwortlatenz oder die Vollständigkeit des Abrufs (Koriat, 2012). Die eigenen Ergebnisse legen nahe, dass die Heuristiken, nach denen die Überwachungsurteile getroffen werden, immer mehr der tatsächlichen Erinnerungsleistung entsprechen, je mehr Informationen einfließen und je valider die genutzten Hinweisreize sind. Da die große Mehrzahl der einschlägigen Studien bisher einzelne ausgewählte Urteile untersucht hat, bedarf diese Annahme jedoch noch weiterer empirischer Bestätigung. Ausnahmen bilden die Studien von Leonesio und Nelson (1990) sowie von Kelemen et al. (2000). In beiden Studien wurden EOLs, JOLs und FOKs, bei Kelemen et al. (2000) zusätzlich ein Maß für das Überwachen des Textverständnisses erhoben. Übereinstimmend mit den eigenen Ergebnissen zeigte sich, dass die JOLs eine deutlich höhere prädiktive Validität aufwiesen als die EOLs. Die Metacomprehension-Leistung lag bei

Kelemen et al. (2000) in etwa auf dem Niveau der EOLs. Mögliche Gründe für diese Befunde bleiben in beiden Studien spekulativ. Die niedrigen Korrelationen zwischen den Gammakorrelationen lassen jedoch beide Autorengruppen davon ausgehen, dass es keine eindimensionale metakognitive Überwachungsfähigkeit gibt, sondern zu verschiedenen Zeitpunkten des Lernvorgangs verschiedene Informationen in die Urteile einfließen. Weitere Studien, die mehrere Maße vergleichen und insbesondere auch SUs außerhalb der Augenzeugengedächtnisforschung einsetzen, sind notwendig. Auch ein integratives Modell, das die Grundlage von Überwachungsurteilen über die gesamte Lern-, Speicher- und Abrufphase differenziert beschreibt und die einzelnen Hinweisreize gewichtet, steht bislang noch aus.

Alterseffekte waren über alle sechs Studien hinweg recht gering ausgeprägt. Vor allem bei den EOLs zeigte sich, dass die jüngeren Erwachsenen besser zwischen später richtigen und falschen Antworten differenzierten als die älteren Erwachsenen und höhere Gammakorrelationen aufwiesen als die Grundschüler. Dies kann mit einem etwas schlechteren Abstraktions- und Antizipationsvermögen in diesen beiden Altersgruppen erklärt werden oder auch mit der Tatsache, dass sich beide Altersgruppen, besonders aber die älteren Erwachsenen, bei der Abgabe der Überwachungsurteile vermehrt auf die Vertrautheit mit dem Itemmaterial stützen (vgl. Shing et al., 2009). Bei den Grundschulern ist nicht auszuschließen, dass zu diesem Zeitpunkt im Lernprozess noch vermehrt Wünsche an die eigene Leistung die Urteile beeinflussen (vgl. Visé & Schneider, 2000). In den JOLs und SUs waren keine konsistenten Alterseffekte nachzuweisen. Die Teilnehmer aller Altersgruppen wiesen hier insgesamt recht gute Überwachungsleistungen auf. Dies entspricht im Wesentlichen aktuellen Befunden, die für Kinder ab etwas der Mitte der Grundschule bereits angemessene und nahezu auf Erwachseneniveau liegende JOLs und SUs nachweisen und nur unter wenigen Bedingungen Defizite im höheren Erwachsenenalter feststellen (für einen Überblick s. Dunlosky & Metcalfe, 2009; Schneider, 2015). Ein methodischer Aspekt der eigenen Studien soll an dieser Stelle noch einmal hervorgehoben werden: Es wurde – im Gegensatz zur Mehrzahl der entwicklungspsychologisch ausgerichteten Studien zum Metagedächtnis – großer Wert darauf gelegt, dass sich die Aufgabenschwierigkeit zwischen den einzelnen Altersgruppen möglichst stark ähnelt. Da dies weitgehend gelungen ist, kann aus den fehlenden Alterseffekten in den JOLs und den SUs geschlossen werden, dass bei ähnlichen Anforderungen auch ähnliche Überwachungsleistungen zu erwarten sind. Bisherige Befunde, die Entwicklungseffekte aufzeigen, müssten dementsprechend auf die grundlegende Aufgabenschwierigkeit überprüft werden (vgl. Connor et al., 1997). Der hier erfolgte Vergleich einer großen Altersspanne erlaubt

es, Entwicklungsprozesse zu kontrastieren und gibt Anhaltspunkte für mögliche Trainings- oder Fördermaßnahmen. So scheint es besonders in Hinblick auf die EOLs von Bedeutung zu sein, bei Grundschulern und älteren Erwachsenen realistischere Einschätzungen der Schwierigkeit von Lernmaterial zu fördern, um nachfolgend effektives Lernen zu ermöglichen. Weiterhin trägt die eigene Arbeit wesentlich dazu bei, die bislang eher spärliche Literatur zur Überwachungsleistungen von Jugendlichen zu erweitern.

7.2 Einfluss von Kontextfaktoren auf die Überwachungsleistung

Nachdem die Leistungen in den einzelnen Überwachungsurteilen zusammenfassend verglichen wurden, soll im Folgenden diskutiert werden, inwiefern verschiedene Kontextfaktoren, nämlich die Komplexität des Lernmaterials, die Modalität der Erinnerungsabfrage, das Vorwissen sowie die verwendete Strategie in den verschiedenen Altersgruppen auf die Qualität der Überwachung Einfluss nehmen.

7.2.1 Einfluss der Komplexität des Lernmaterials

Inwieweit sich die Überwachungsurteile in den verschiedenen Altersgruppen in Abhängigkeit der Komplexität des Lernmaterials unterscheiden, lässt sich im Vergleich der Studien 1 und 3 mit den restlichen Studien zum Paar-Assoziationslernen (Studien 2, 5 und 6) nachvollziehen. Das Wiedererkennen von Sätzen in Studie 3 sowie die Studie 4 sollen nicht berücksichtigt werden, da hier im Gegensatz zu den übrigen Studien ein Rekognitionsparadigma angewandt wurde.

In Studie 1 wurden JOLs und SUs beim Lernen von Bildpaaren und von Informationen aus einem Film erstmals direkt kontrastiert. Die Genauigkeiten der Überwachung waren in allen Bedingungen und Altersgruppen sehr hoch (im Mittel bei .87 bzw. .86), sodass hier keine signifikanten Effekte resultierten. Hinsichtlich der Differenzierung zeigten sich bei beiden Maßen signifikante Wechselwirkungen zwischen Korrektheit der Antwort und Komplexität des Lernmaterials: Die Teilnehmer erreichten also bessere Differenzierungsleistungen bei den einfacheren Materialien. Insgesamt waren die Leistungen sowohl für die Genauigkeiten als auch für die Differenzierung sehr gut. Eine Besonderheit in dieser Studie stellte die Tatsache dar, dass die Antwort „weiß nicht“ möglich war; diese Option wurde bei Unsicherheit recht häufig von den Probanden genutzt, so in 19% der Fälle bei den Bildpaaren und in 20% der Fälle

bei den Fragen zum Film. Dies kann unter Umständen dazu geführt haben, dass auch schon für das komplexere Material gute Überwachungsleistungen gefunden wurden. Denn auch in anderen Studien mit Filmmaterial, bei denen „weiß nicht“-Antworten möglich waren, zeigten sich recht gute Überwachungsleistungen (von der Linden & Roebbers, 2006; Roebbers, von der Linden & Howie, 2007).

Abgesehen davon, dass die Antwortoption „weiß nicht“ in Studie 3 nicht mehr möglich war, unterschied sich das Vorgehen auch dahingehend, dass die Fragen zum Text schwieriger zu beantworten waren als die Fragen zum Film, da z.T. auch Inferenzen gezogen werden mussten und die Fragen gerade für die Novizen deutlich anspruchsvoller waren. Auch nach zweimaliger Darbietung des Textes waren die Erinnerungsleistungen etwas schlechter als bei den Fragen zum Film. Es war demnach mehr kognitive Energie für das Beantworten der Fragen nötig, die dann wahrscheinlich von der Überwachung abgezogen wurde (vgl. de Bruin et al., 2007). Diese Faktoren können erklären, dass die Überwachungsqualität in Studie 3 schlechter ausfiel als in Studie 1 und in den Experimenten zum Paar-Assoziationslernen. So lagen die mittleren Gammakorrelationen bei den JOLs bei .55, bei den SUs bei .64. In den übrigen Studien wurden dagegen mittlere Genauigkeitswerte von .75 bis .79 für die JOLs (Ausnahme: Lerndurchgang 1 in Studie 5) bzw. Werte von .88 bis .96 für die SUs erreicht. Damit liegen die Ergebnisse in Studie 3 deutlich näher an den bisherigen Ergebnissen zur Metacomprehension als die in Studie 1, jedoch immer noch höher als die in Überblicksarbeiten gefundenen Werte von etwa .25 für JOLs (Dunlosky & Lipko, 2007; Maki, 1998a). Dies kann für die Studie 3 damit erklärt werden, dass die prädiktiven Urteile auf Fragen bezogen waren, die in der Erinnerungsabfrage wörtlich wieder gestellt wurden. Die Übereinstimmung war also für beide Formate sehr hoch, d.h. Probeabrufe, die während der JOL-Abgabe stattfanden, hatten eine hohe prädiktive Validität für die spätere Erinnerungsleistung (vgl. Dunlosky, Rawson & McDonald, 2002). Weiterhin wurden verhältnismäßig viele Fragen zum Text gestellt (18 bis 20 Fragen), ein Faktor, der ebenfalls zu besseren Metacomprehension-Leistungen beiträgt, da die Berechnung der Korrelationen dann sensibler für Abweichungen wird (Weaver, 1990). Es konnte auch andernorts gezeigt werden, dass für JOLs unter günstigen Bedingungen auch beim Lernen aus Texten recht hohe Gammakorrelationen gefunden werden können (z.B. Anderson & Thiede, 2008; Roebbers, von der Linden, Schneider & Howie, 2007).

Für Gammakorrelationen bei SUs beim Lernen aus Texten existieren divergierende Ergebnisse. So lagen die mittleren Werte beispielsweise in der Studie von Pressley und Ghatala (1988) bei .31 und .46, bei Mengelkamp und Bannert (2010) bei .62. Für Kinder zwischen sieben und neun Jahren fanden von der Linden et al. (2011) mittlere Korrelationen zwischen .50 und .80. Die im

Vergleich zu Pressley und Ghatala (1988) höheren eigenen Gammakorrelationen sind aus dem Antwortformat erklärbar: Die Autoren stellten Mehrfachwahlaufgaben, bei denen Raten eine größere Rolle spielt als bei offenen Fragen und damit auch die Sicherheit in die eigenen Antworten abnimmt. Insofern korrespondieren die eigenen Ergebnisse gut mit der bestehenden Literatur zu SUs. Ähnlich wie für die JOLs scheint auch hier zu gelten, dass günstige Untersuchungsbedingungen verhältnismäßig genaue Urteile hervorbringen.

Da der Vergleich von Überwachungsleistungen in komplexeren und einfacher strukturierten Materialien keine Alterseffekte ergab, ist davon auszugehen, dass sich die diskutierten Mechanismen bei Anpassung der Aufgabenschwierigkeit in den untersuchten Lebensabschnitten ähneln bzw. konstant bleiben. Es scheint sich also um eine Kompetenz zu handeln, die schon früh im Leben erworben wird und auch bis ins hohe Erwachsenenalter erhalten bleibt.

Insgesamt kann festgestellt werden, dass die Überwachung für komplexere Materialien nur dann auf dem gleichen Niveau liegt wie die für Paar-Assoziationen, wenn die Zurückhaltung von Antworten gewährleistet ist, wenn die Fragen sich auf unmittelbar dem Text oder Film entnehmbare Informationen beziehen, wenn möglichst viele Fragen gestellt werden und wenn das Format sowohl für das Überwachungsurteil als auch für die Erinnerungsabfrage identisch ist.

7.2.2 Einfluss der Modalität der Erinnerungsabfrage

Die Frage danach, inwiefern die Qualität des Überwachungsurteile davon beeinflusst wird, ob es sich um eine offene Erinnerungsabfrage bzw. cued-recall oder um eine Rekognitionsaufgabe handelt, kann ebenfalls nur für die JOLs und die SUs diskutiert werden, da EOLs nur in einer Studie erfasst wurden und dadurch eine Vergleichsmöglichkeit fehlt.

In Studie 3 wurden SUs bei der Aufgabe zum Wiedererkennen von Sätzen aus dem Text erhoben; in Studie 4 wurden sowohl die JOLs als auch die SUs in Bezug auf das Wiedererkennen von Wortpaaren erfasst. In allen übrigen Studien wurden offene Fragen gestellt bzw. die Erinnerung an Wort- und Bildpaare im cued-recall-Paradigma erfasst.

In Studie 3, d.h. beim Wiedererkennen der Sätze, wurde über alle Gruppen hinweg eine mittlere Gammakorrelation von .26 gefunden. Die mittleren SUs nach richtigen vs. falschen Antworten unterschieden sich kaum voneinander (etwa 0.20 auf der 7-Punkte-Skala). In Studie 4 erreichten die Probanden im Mittel über alle Altersgruppen und Versuchsbedingungen hinweg eine

mittlere Gammakorrelation zwischen Rekognitionsleistung und JOLs von .23. Bei den SUs lag der Wert mit .37 geringfügig höher. Dies spiegelte sich auch in den Differenzierungsleistungen wider: In den JOLs unterschieden sich die Mittelwerte zwischen später richtigen und falschen Antworten insgesamt lediglich um 10 Punkte auf der 100-Punkte-Skala, in den SUs nur um etwa 12 Punkte.

In allen anderen Telexperimenten wurden für JOLs und SUs wesentlich bessere Überwachungsleistungen gefunden, sowohl was die Genauigkeiten als auch was die Differenzierung angeht. Das heißt, obwohl die Gedächtnisleistung in der Regel bei Aufgaben zur freien Erinnerung schlechter ausfällt als bei Rekognitionsaufgaben (Thiede & Dunlosky, 1994), scheint es sich mit der Überwachungsleistung genau umgekehrt zu verhalten (vgl. auch Rhodes & Tauber, 2011). Bezüglich der JOLs in Studie 4 kann dies daran liegen, dass das Format für die Erhebung der JOLs (nur Stimulus) nicht mit dem für die Wiedererkennensleistung (Stimulus und Target) übereinstimmte. Die Entscheidung, in der JOL-Phase nur das linke Wortpaar darzubieten, wurde vor allem von der Überlegung geleitet, ein erneutes Lernen des Itempaars zu vermeiden. Es konnte jedoch andernorts gezeigt werden, dass die Überwachungsleistung in Rekognitionsparadigmen für JOLs höher liegt, wenn nur der Stimulus in der JOL-Phase gezeigt wird, da in diesem Fall wahrscheinlich weniger Probeabrufe während der Urteilsbildung unternommen werden (Dunlosky & Nelson, 1997). Andere mögliche Gründe für die niedrigeren Überwachungsleistungen bei Rekognitionsaufgaben als bei Aufgaben zur freien Erinnerung ist zum einen die Tatsache, dass ein Item, das während der JOL-Phase als nicht gespeichert klassifiziert wurde, dennoch korrekt geraten werden kann, zum anderen, dass Disktraktoren in der Erinnerungsphase ausgeschlossen werden können, selbst wenn während der JOL-Phase der Probeabruf nicht erfolgreich war (Thiede & Dunlosky, 1994).

Die SUs in Studie 3 und in Studie 4 müssen getrennt betrachtet werden, da es sich in Studie 4 um ein Paar-Assoziationsparadigma, in Studie 3 um ein Metacomprehension-Design handelt und hier z.T. unterschiedliche Wirkmechanismen zugrunde liegen. In Studie 4 sind die höheren mittleren Gammakorrelationen bei den SUs im Vergleich zu den JOLs wahrscheinlich dem Umstand geschuldet, dass die SUs sich konkret auf die dargebotenen Wortpaare bezogen, während den Teilnehmern bei den JOLs noch nicht bekannt war, in welcher Form (identisch, neu kombiniert, neu) die Items in der Erinnerungsphase dargeboten werden würden. Dass die SUs dennoch in Studie 4 niedriger ausfielen als in sämtlichen anderen Experimenten zur freien Erinnerung liegt – ähnlich wie bei den JOLs – daran, dass Unschärfen in der Überwachung durch korrektes oder inkorrektes Raten auftreten (s.o.). Zusätzlich ist die Abrufleichtigkeit bei SUs in Aufgaben zur freien Erinnerung ein wesentlicher und valider Hinweisreiz, der in

Wiedererkennensparadigmen nicht genutzt werden kann, was die Überwachungsleistung ebenfalls verschlechtert (vgl. M. D. Robinson & Johnson, 1996). In Studie 3 kommt als weitere mögliche Ursache für die niedrigen Genauigkeiten und Differenzierungsleistungen in den SUs hinzu, dass die Aufgabe sich auf die Textoberfläche bezieht. Es ist aufgrund der Wiedererkennensleistung, die nur geringfügig über der Ratewahrscheinlichkeit liegt, davon auszugehen, dass die Probanden mehrheitlich ein Situationsmodell bildeten und die wörtliche Enkodierung des Textes eine untergeordnete Rolle spielte (vgl. Kintsch, 1998). Da die jüngsten Teilnehmer bereits in der dritten Klasse waren, ist auch bei ihnen schon anzunehmen, dass sich ihre Textrepräsentationen auf der gist-Ebene befinden (vgl. Nieding, 2006). Die fehlende Passung zwischen Repräsentationsebene und Erinnerungsabruf trug mit hoher Wahrscheinlichkeit zusätzlich zu den schlechten Überwachungsleistungen in dieser Aufgabe bei.

Insgesamt konnte also bestätigt werden, dass die Überwachung bei Rekognitionsaufgaben gegenüber freier Erinnerung erschwert ist. Dies gilt sowohl für JOLs als auch für SUs, für EOLs kann aufgrund mangelnder Vergleiche keine Aussage getroffen werden. Alterseffekte waren auch für diesen Kontextfaktor nicht zu beobachten. Die gerade für die SUs postulierte bessere Leistung von Kindern bei Rekognitionsaufgaben im Vergleich zur freien Erinnerung (Roebers, 2002) und die bei älteren Erwachsenen gefundenen Verschlechterungen für Rekognitionsaufgaben im Vergleich zu jüngeren Erwachsenen (z.B. Karpel et al., 2001) wurden hier nicht festgestellt. Dies liegt wahrscheinlich daran, dass sich das Design der eigenen Studien zum Wiedererkennen nicht an Paradigmen zur Augenzeugengedächtnisforschung orientierte: Es kamen also weder inzidentelles Lernen wie bei Roebers (2002) noch suggestive Fragen wie bei Karpel et al. (2001) zum Einsatz – also Faktoren, die auf die kognitive Beanspruchung Einfluss nehmen und sich damit auch auf die Überwachungsleistung auswirken.

7.2.3 Einfluss des Vorwissens

Der Einfluss des Vorwissens auf die Überwachungsleistungen in den beteiligten Altersgruppen wurde bereits in der zusammenfassenden Diskussion zu Studie 2 und 3 ausführlich beleuchtet und soll deshalb an dieser Stelle nur kurz dargestellt werden. Insgesamt gesehen wurden in beiden Studien eindeutige Vorteile für die Experten im Vergleich zu den Novizen bezüglich der Gedächtnisleistung gefunden (vgl. Feltovich et al., 2006). Diese Überlegenheit konnte für die Überwachungsleistung nur in einigen Teilbereichen gefunden werden.

In den Differenzierungsleistungen zeigte sich über alle Maße und beide Studien hinweg, dass die Experten ihre eigenen Leistungen optimistischer einschätzten als die Novizen. Die für die Urteile genutzten Hinweisreize scheinen sich also zwischen Probanden mit einem höheren Vorwissen und solchen mit weniger Vorwissen unterschieden zu haben. Es ist anzunehmen, dass bei den Experten neben Hinweisen aus den Itemcharakteristika oder aus dem Lern- und Abrufprozess auch (positive) Annahmen über den eigenen Expertenstatus herangezogen wurden, die zu einer optimistischeren Einschätzung im Vergleich zu den Novizen führten (vgl. Zhao & Linderholm, 2008). Da die Effektstärken des Haupteffekts Expertise für die Differenzierungsleistungen jedoch im kleinen bis mittleren Bereich lagen (.03 - .18), ist zu vermuten, dass andere Hinweisreize einen größeren Einfluss auf die Höhe der Urteile hatten als Vorannahmen der Probanden zur eigenen Expertise.

In den Gammakorrelationen ergaben sich für die EOLs und JOLs keine bedeutsamen Unterschiede zwischen den Experten und den Novizen; bezüglich der SUs waren die Teilnehmer mit einem höheren Vorwissen der Vergleichsgruppe bei den Wortpaaren in der Genauigkeit überlegen. Dieser Effekt zeigte sich auch bei den Fragen zum Fußballtext für die Jugendlichen. Bei den jüngeren und älteren Erwachsenen verfehlte der Unterschied das Signifikanzniveau. Die Tatsache, dass sich das Vorwissen nur bei den SUs günstig auswirkte, könnte darin begründet sein, dass Experten zu diesem Zeitpunkt des Lernvorgangs u.U. übermäßig optimistische Einschätzungen ihrer Leistung aufgrund der Erfahrungen mit dem Lernmaterial bzw. der zusätzlichen Informationen aus dem Speicher- und Abrufvorgang anpassen konnten. Die Grundschüler waren jedoch nicht in der Lage, diesen Nachteil in den SUs auszugleichen.

Die beiden Studien zum Einfluss des Vorwissens auf die Überwachungsleistungen in verschiedenen Altersgruppen geben also Hinweise darauf, dass die kognitive Überlegenheit von Experten im Vergleich zu Novizen nicht vollständig, sondern vor allem für die retrospektive metakognitive Überwachung gilt.

7.2.4 Einfluss der verwendeten Strategie

Der Effekt der verwendeten Strategie bzw. die Frage danach, inwiefern die Überwachungsleistungen durch ein Strategietraining gesteigert werden können, soll hier nur in Kürze behandelt werden, da dies schon in der Diskussion zu Studie 4 ausführlich beleuchtet wurde. Insgesamt zeigen die Ergebnisse der Studie 4, dass sich die Überwachungsleistungen in

den JOLs und den SUs vor allem bei den Drittklässlern und den älteren Erwachsenen, gefolgt von den Jugendlichen durch ein kurzes Training in einer „visual-imagery“-Strategie verbessern ließen. Die jüngeren Erwachsenen wendeten häufiger spontan eine effiziente Strategie zum Lernen von Paar-Assoziationen an als die Kinder und die Senioren und profitierten daher weniger vom Training (vgl. Naveh-Benjamin et al., 2007; Schneider, 2015). Die Vorteile der Experimentalgruppe in der Überwachung können jedoch nicht ausschließlich durch die verbesserte Erinnerungsleistung in dieser Gruppe erklärt werden (vgl. Dunlosky & Metcalfe, 2009). Denn sonst hätten auch in den Vorwissensexperimenten die Experten, die nahezu durchgängig bessere Erinnerungsleistungen als die Novizen erreichten, durchgängig bessere Überwachungsleistungen zeigen müssen, was nicht der Fall war. Die Annahme, dass das Strategietraining die Nutzung idiosynkratischer Hinweisreize fördert, wird durch diese Ergebnisse unterstützt. Die Teilnehmer in der Trainingsgruppe verknüpften also die zu lernenden Wortpaare mit mehr Details, die ihnen dann auch in der Abrufphase zur Verfügung standen. Auf diese Weise konnte erreicht werden, dass vertrautheitsgestützte Prozesse sowohl beim Erinnerungs- als auch beim Überwachungsvorgang bedeutsam reduziert und die entsprechenden Leistungen gesteigert wurden.

Das Training stellte also eine sehr ökonomische und wirksame Maßnahme dar, Überwachungsprozesse zu verbessern. Dies konnte erstmals auch für JOLs nachgewiesen werden (vgl. Shing et al., 2009 für SUs). Es steht aus, auch dieses Verfahren in Zukunft auf andere, vor allem angewandte Bereiche auszuweiten und das Training zu intensivieren, um die Verallgemeinerbarkeit der eigenen Befunde zu überprüfen.

7.2.5 Fazit

Die Untersuchung des Einflusses verschiedener Kontextfaktoren auf die metakognitive Überwachung im Entwicklungsverlauf macht deutlich, dass eine verbesserte Erinnerungsleistung nicht zwangsläufig mit einer besseren metakognitiven Leistung einhergeht.

Fragen zu komplexeren Materialien wie Filmen oder Texten stellen höhere kognitive Anforderungen während der Verarbeitung als Paar-Assoziationsaufgaben; denn hier ist es in der Regel weniger eindeutig, welche Antwort genau gefordert wird, und auch mehrere Antworten korrekt sein können. Dies wirkt sich häufig auch dahingehend aus, dass die Überwachungsleistung im Gegensatz zum Lernen von Itempaaren schlechter ausfällt (z.B.

Maki, 1998b). In den eigenen Studien zeigte sich jedoch, dass unter günstigen Bedingungen (d.h. eindeutige Fragen, hohe Korrespondenz zwischen metakognitivem Urteil und Art der Erinnerungsabfrage) altersunabhängig keine schlechteren Überwachungsleistungen zu erwarten sind. Für den Vergleich zwischen der Überwachung bei Rekognitionsaufgaben und solchen zur freien Erinnerung ergab sich das gegenteilige Bild: Zwar war hier die Erinnerungsleistung besser als beim freien Erinnern, die pro- und retrospektive Überwachung fiel den Teilnehmern jedoch etwas schwerer, da u.a. Aspekte wie das Raten von Antworten eine Rolle spielt. Für den Einfluss des Vorwissens konnte in den eigenen Studien in Übereinstimmung mit der Literatur belegt werden, dass sich hierdurch die Erinnerungsleistung weitgehend altersunabhängig verbessert (vgl. z.B. auch Chi, 2006), eine daraus resultierende erleichterte Überwachung konnte vor allem bei den retrospektiven Urteilen nachgewiesen werden. Schließlich ergab ein Training in einer Elaborationsstrategie sowohl Steigerungen der Gedächtnis- als auch der Überwachungsleistung. Dies galt besonders für die Grundschüler und die älteren Erwachsenen.

Insgesamt gesehen scheint demnach für eine Verbesserung in der Überwachungsfähigkeit nicht ausschließlich eine erleichterte Verarbeitung des Materials und damit eine erhöhte Gedächtnisleistung von Bedeutung zu sein; vielmehr scheint eine Übereinstimmung zwischen Enkodier- bzw. Abrufvorgang und pro- bzw. retrospektivem Überwachungsurteil relevant zu sein. Nur wenn für beide Prozesse ähnliche und diagnostisch valide Hinweisreize zur Verfügung stehen (vgl. Koriat, 1997), geht eine bessere Gedächtnisleistung auch mit einer besseren Überwachung einher. Die eigenen Studien weisen darauf hin, dass diese Übereinstimmung dann erreicht wird, wenn Fragen zu komplexeren Materialien und Überwachungsurteile vergleichbare Formulierungen haben und dadurch z.B. Probeabrufe bei JOLs eine hohe Validität haben, wenn Raten bei der Lösung von Aufgaben eine untergeordnete Rolle spielt, wenn die Wahrnehmung des eigenen Expertenstatus als Hinweisreiz weniger wichtig und Informationen aus dem Lern- und Abrufprozess wichtiger für die Abgabe von Überwachungsurteilen werden oder wenn viele Details aus dem Enkodiervorgang als Basis für den Überwachungsprozess zur Verfügung stehen.

7.3 Ziel- und Datenorientierung im Entwicklungsverlauf

Die letzten beiden Studien der Arbeit befassten sich mit der wechselseitigen Beeinflussung von Überwachungs- und Kontrollprozessen.

Die Zielorientierung wurde in der vorliegenden Arbeit auf zwei Arten untersucht: Zum einen wurde in Studie 5 nach einem Paradigma von Lockl und Schneider (2003) geprüft, inwieweit die Probanden Informationen aus dem Überwachungsprozess für die Anpassung der Lernzeit nutzen (MC-Modell). Zum anderen wurde in Studie 6 die Anpassung der Lernzeit und der JOLs in Abhängigkeit eines extern durch Anreize gesetzten Zieles untersucht. Das MC-Modell konnte – ohne gleichzeitige Untersuchung des CM-Modells – erstmals über eine große Altersspanne hinweg nachgewiesen werden. Das heißt, die Teilnehmer aller Altersgruppen richteten prinzipiell ihre Lernzeit nach der Höhe der JOLs aus, was auch den Vorbefunden für einzelne Altersbereiche aus der Literatur entspricht (Dunlosky & Connor, 1997; Koriat et al., 2006; Lockl & Schneider, 2003).

Evidenz für das CM-Modell, d.h. für datenorientiertes Lernen, ergab sich über alle Altersgruppen hinweg in etwas geringerem Maße. Die Analysen der JOLs nach dem Mediansplit der Lernzeit zeigten für beide Studien signifikante Haupteffekte des Faktors Lernzeit, aber keine statistisch bedeutsamen Interaktionen mit dem Faktor Altersgruppe. Obwohl die Effektstärken in allen Fällen im mittleren Bereich lagen (η^2 zwischen .16 und .22), war der Unterschied in den JOLs zwischen länger und kürzer gelernten Wortpaaren auf Itemebene nicht groß genug, um sich in allen Altersgruppen konsistent in signifikanten, negativen Korrelationen zwischen JOLs und Lernzeit niederzuschlagen. Bei den Drittklässlern erreichte keine Korrelation das Signifikanzniveau, erst bei den jüngeren und älteren Erwachsenen ergaben sich nahezu durchgängig statistisch bedeutsame Werte. Im Wesentlichen entsprechen diese Befunde Ergebnissen zu Altersverläufen anderer Studien, die besagen, dass etwa ab der dritten Klasse Nachweise für das CM-Modell gefunden werden können (Koriat et al., 2009b), die sich jedoch erst ab etwa der vierten Klasse stabilisieren, d.h. dann auch nur für schwierige Itempaare nachweisen lassen (Koriat et al., 2009a). Hervorzuheben ist, dass das CM-Modell erstmals auch bei älteren Erwachsenen nachgewiesen werden konnte. Die vorliegenden Ergebnisse aus beiden Studien zeigen dabei keine wesentlichen Defizite im datenorientierten Lernen für diese Altersgruppe. Die niedrigeren Zusammenhänge zwischen Lernzeit und JOLs aus Studie 5 ließen sich nicht replizieren, sodass nicht grundsätzlich davon ausgegangen werden kann, dass Senioren durch Verschlechterungen in den exekutiven Funktionen Nachteile im Zusammenspiel zwischen Überwachung und Steuerung erwachsen (vgl. Pansky et al., 2009; Souchay & Isingrini, 2004).

Für die Stabilität des CM-Modells bei älteren Erwachsenen spricht auch die Tatsache, dass dessen Ausprägung durch die gleichzeitige Manipulation der Zielorientierung in Studie 6 (Lerndurchgang 2) nicht tangiert wurde. Der flexible Wechsel zwischen beiden Modellen

konnte jedoch insgesamt am besten bei den Jugendlichen und den jüngeren Erwachsenen beobachtet werden. Die älteren Erwachsenen passten weder die Lernzeiten noch die JOLs an die Punkte an, die Grundschüler ausschließlich die JOLs. Da dieser Befund nicht durch grundsätzliche Defizite im MC-Modell in diesen Altersgruppen erklärt werden kann (s.o.), ist eher zu vermuten, dass der flexible Wechsel zwischen beiden Modellen diesen Altersgruppen Schwierigkeiten bereitete (vgl. Koriat, Ackerman et al., 2014). Neben der mangelnden Flexibilität dieser beiden Altersgruppen kommt als Erklärung auch die – im Vergleich zu den Punktwerten – höhere Salienz der Itemschwierigkeit in Frage (vgl. Ariel et al., 2009; Lockl & Schneider, 2004; Price et al., 2010). Es sollte also in zukünftigen Experimenten untersucht werden, ob Grundschüler und ältere Erwachsene in vollem Ausmaß zielorientiertes Lernen zeigen, wenn die Itemschwierigkeit konstant gehalten wird. Die Gruppe der Grundschüler zeigte zusätzlich beim flexiblen Wechsel zwischen ziel- und datenorientiertem Lernen im Vergleich zu den anderen Altersgruppen keine signifikanten Korrelationen zwischen JOLs und Lernzeit (CM-Modell); lediglich die Analyse durch den Mediansplit der Lernzeit wurde auch in dieser Gruppe statistisch bedeutsam. Da dieses Ergebnis jedoch auch für rein datenorientiertes Lernen (zweiter Lerndurchgang in Studie 5; erster Lerndurchgang in Studie 6) auftrat, ist es wohl nicht durch die Interferenz von zielorientierten Prozessen zu erklären, sondern dadurch, dass die Sensitivität für datenorientierte Prozesse auch noch nach dem Grundschulalter zunimmt (Koriat, Ackerman et al., 2014).

Insgesamt scheint der komplexe Vorgang, sowohl zielorientiertes Lernen als top-down-Prozess als auch datenorientiertes Lernen als bottom-up-Prozess innerhalb eines einzigen metakognitiven Urteils zu bündeln, vor allem im Jugend- und jüngeren Erwachsenenalter erfolgreich umgesetzt werden zu können. Für Grundschüler und Senioren scheint die Abstimmung beider Vorgänge eine größere Herausforderung darzustellen. Die eigene Arbeit liefert somit auch hier wichtige Ansatzpunkte für Fördermaßnahmen in diesen Altersgruppen liefert, denn für gute Lernergebnisse ist auch in angewandten Bereichen eine optimale Anpassung von Überwachungs- und Steuerungsprozessen notwendig.

7.4 Ausblick

Auch wenn die vorliegende Arbeit eine Fülle an Fragestellungen zur Entwicklung des prozeduralen Metagedächtnisses erstmals in der Lebensspanne untersuchte, bleiben dennoch einige Punkte offen, die einer weiteren wissenschaftlichen Auseinandersetzung bedürfen.

In methodischer Hinsicht wurden durch zahlreiche Pilotierungsstudien große Anstrengungen unternommen, die Aufgabenschwierigkeit für alle Altersgruppen möglichst konstant zu halten. Dies unterscheidet die Ergebnisse der vorliegenden Arbeit von der Mehrzahl der einschlägigen Studien (für Ausnahmen s. Hines et al., 2009; Lockl & Schneider, 2003; Shing et al., 2009) und erschwert eine direkte Vergleichbarkeit der Ergebnisse. Es kann nicht ausgeschlossen werden, dass andernorts gefundene Alterseffekte, die sich in den eigenen Studien nicht replizieren ließen, zu einem gewissen Teil auch auf unterschiedliche Erinnerungsleistungen in den verschiedenen Altersgruppen zurückzuführen sind (z.B. Hoffmann-Biencourt et al., 2010; Koriat & Shitzer-Reichert, 2002). Der Zusammenhang zwischen Gedächtnis- und Überwachungsleistung sollte auch in zukünftigen Studien noch genauer untersucht werden. Zwar liegen Befunde vor, die zeigen, dass eine bessere Erinnerungsleistung mit einer genaueren Überwachung einhergeht (z.B. Schneider, 1989), welche kausale Wirkrichtung („Eine bessere Erinnerungsleistung erleichtert und verbessert die Überwachung“ vs. „Bessere Überwachungsleistungen führen zu besseren Erinnerungsleistungen“) in welchem Ausmaß gültig ist, muss jedoch noch genauer analysiert werden. Die eigenen Studien machen deutlich, dass ein komplexer Zusammenhang vorliegt. Denn einerseits konnte durch ein Strategietraining die Überwachungsleistung *und* die Erinnerungsleistung wesentlich gesteigert werden, andererseits zeigten die Studien zum Einfluss des Vorwissens aber auch, dass eine bessere Gedächtnisleistung nicht zwangsläufig zu besseren Leistungen im prozeduralen Metagedächtnis führt. Es ist also sinnvoll, in zukünftigen Studien in verschiedenen Altersstufen eine gezielte Manipulation der Erinnerungsleistung vorzunehmen, um genauere Einblicke in kausale Zusammenhänge zu erhalten (vgl. Connor et al., 1997).

Der Einschluss von vier Altersgruppen, die eine Spanne zwischen 7 und 80 Jahren umfassten, stellte nicht nur in Bezug auf die Aufgabenschwierigkeiten eine methodische Herausforderung dar. Zum einen können Kohorteneffekte in einem querschnittlichen Design dieser Spannweite nicht ausgeschlossen werden. Zum anderen stellen besonders die älteren Erwachsenen eine heterogene Gruppe im Vergleich zu anderen Altersstufen dar (vgl. Baltes & Baltes, 1990). Um auch in den anderen Gruppen eine möglichst große Diversität zu erreichen, wurden die Jugendlichen in verschiedenen Schularten rekrutiert; bei den jüngeren Erwachsenen wurde darauf geachtet, nicht nur Studenten einzubeziehen. Die Ergebnisse des demografischen Fragebogens zeigen, dass zumindest die jüngeren und die älteren Erwachsenen in wesentlichen Aspekten (Bildungsstand, Gesundheit) weitgehend vergleichbar waren. Die einbezogenen Senioren waren insgesamt von guter Gesundheit, was dem Ziel, den „normalen“ Alterungsprozess abzubilden, entspricht. Es kann am Design der eigenen Studien kritisch

gesehen werden, dass die Abstände zwischen den Altersgruppen recht groß waren. Obwohl die wenigen generellen Alterseffekte die Entscheidung für diese vier Gruppen insgesamt klar befürworten, könnte der Einschluss von weiteren Altersgruppen in zukünftigen Experimenten u.U. detaillierteren Aufschluss über Entwicklungsveränderungen geben.

Weiterhin erscheint es nötig, die Ergebnisse der vorliegenden Arbeit auch in größeren Stichproben zu replizieren. Dies würde zum einen die Gültigkeit der Befunde untermauern, zum anderen könnten dann auch Analysen zur Kalibrierung durchgeführt werden (vgl. Keast et al., 2007). Letzterer Aspekt könnte dazu beitragen, Verzerrungen im Sinne einer Überschätzung der eigenen Leistung bei Experten unterschiedlicher Domänen genauer zu erfassen, als dies in den eigenen Vorwissenstudien durch die Auswertung der Differenzierungsleistung möglich war. Außerdem könnten dann auch Tendenzen von älteren Probanden, bei vertrauten Stimuli höhere metakognitive Urteile abzugeben, noch besser abgebildet werden (vgl. Shing et al., 2009).

Auch in theoretischer Hinsicht bleiben Fragen offen. Es muss nach wie vor als ungeklärt gelten, in welchem Maße die verschiedenen metakognitiven Urteile auf der gleichen Grundlage beruhen, d.h. ob es sich um eine einzige Komponente oder um verschiedene Urteile handelt (vgl. Kelemen et al., 2000; Leonasio & Nelson, 1990). Die eigenen Studien weisen darauf hin, dass es große Überschneidungen zwischen den einzelnen Urteilen gibt, die zwischen JOLs und SUs stärker ausgeprägt sind als zwischen EOLs und JOLs bzw. SUs. Welche Anteile jedoch zur gemeinsamen Varianz beitragen und welche einen eigenen Beitrag zur Höhe der Urteile liefern, muss noch genauer geklärt werden. Trotz der inzwischen beeindruckenden Forschungstradition zur Metakognition seit den Anfängen in den 70er-Jahren, fehlt es an Studien, die verschiedene Urteile im Verlauf des Lernprozesses direkt kontrastieren.

Auch die Frage nach den Grundlagen der Überwachungsurteile kann noch nicht als geklärt bezeichnet werden. Der Cue-utilization-Ansatz (Koriat, 1997) bietet eine Möglichkeit, verschiedene Hinweisreize theoretisch zu klassifizieren. Dennoch ist es bislang noch nicht umfassend gelungen, den genauen Anteil der einzelnen Hinweisreize zu identifizieren. Die eigenen Experimente leisten hierzu einen Beitrag. Die Ergebnisse deuten darauf hin, dass neben der Itemschwierigkeit Annahmen über die eigene Expertise und über die Vertrautheit mit den Items, die Fähigkeit, Details mit den zu lernenden Items zu verknüpfen, sowie Anreize, die mit dem Lernerfolg verbunden sind, auf die Höhe der Überwachungsurteile Einfluss nehmen. Weiterführende Analysen, inwiefern sich die Wirkung dieser Hinweisreize zu verschiedenen Zeitpunkten des Lernprozesses und in verschiedenen Altersgruppen unterscheidet, stehen

jedoch noch aus. Auch die Frage danach, inwieweit sich metakognitive Urteile auch nach Informationen aus dem Steuerungsprozess richten (CM-Modell), verdient in Zukunft weiterer Betrachtung. Denn bislang wurde vorrangig die Einteilung der Lernzeit als Maß für die Steuerung herangezogen wurde, sodass eine Ausweitung auf andere Operationalisierungen sinnvoll ist (für eine Ausnahme s. Koriat et al., 2009a).

Insgesamt gesehen erscheint es als wünschenswert, auch weiterhin Fragestellungen aus dem Bereich der Metagedächtnisentwicklung in lebensspannenübergreifenden Designs zu untersuchen. Die gering ausgeprägten Alterseffekte in den Überwachungsleistungen lassen darauf schließen, dass – besonders bei den Grundschulern und den älteren Erwachsenen – bereits bzw. noch hohe kognitive Potenziale vorhanden sind, die zur Kompensation möglicher kognitiver Defizite genutzt werden sollten.

8. ZUSAMMENFASSUNG DER ARBEIT

Der Entwicklungsverlauf metakognitiver Überwachungsprozesse und das Zusammenspiel von Überwachungs- und Kontrollprozessen ist über die gesamte Lebensspanne hinweg nur für isolierte Altersgruppen, nicht aber in Studien, die Teilnehmer vom Kindes- bis zum höheren Erwachsenenalter einschließen, untersucht worden. Diese Lücke sollte mit der vorliegenden Arbeit geschlossen werden, denn gerade solche Designs können dazu beitragen, Aufbau- und Abbauprozesse zu kontrastieren, und Hinweise auf frühzeitig vorhandene sowie im Altersverlauf bestehende Fähigkeiten geben, die dann kompensatorisch genutzt werden können. Die eigene Arbeit befasste sich dabei mit dem Verlauf einer Vielzahl von pro- und retrospektiven Überwachungsvorgängen über die Lebensspanne. Der Schwerpunkt lag auf dem Einfluss verschiedener Kontextfaktoren (z.B. Komplexität des Lernmaterials, Vorwissen, Strategienutzung) auf die Überwachungsleistung in den jeweiligen Altersstufen. Außerdem wurde überprüft, inwieweit wechselseitige Zusammenhänge zwischen Überwachungs- und Steuerungsprozessen in den untersuchten Altersgruppen unterschiedlich stark ausgeprägt sind. Diese Fragestellungen wurden in sechs Experimenten mit insgesamt 816 Teilnehmern untersucht. Es handelte sich dabei um Drittklässler im Alter zwischen 7 und 9 Jahren, Jugendliche zwischen 12 und 14 Jahren, jüngere Erwachsene zwischen ca. 18 und 25 Jahren sowie ältere Erwachsene zwischen ca. 60 und 80 Jahren. Erhoben wurden Ease-of-Learning-Urteile (EOLs) bzw. ein globales Verständniskriterium als Maß der Überwachung vor dem eigentlichen Lernprozess, Judgments of Learning (JOLs) als Maß der Überwachung nach dem Lernvorgang und Sicherheitsurteile (SUs) als Maß der Überwachung nach dem Erinnerungsabruf.

Es zeigte sich, dass die Überwachungsleistung sowohl, was die Differenzierungsfähigkeit zwischen richtigen und falschen Antworten, als auch, was die Genauigkeit betrifft, bezüglich der JOLs und der SUs über die gesamte untersuchte Altersspanne hinweg im Wesentlichen konstant und auf recht hohem Niveau blieb. Lediglich bei den EOLs ergaben sich Alterseffekte: Die jüngeren Erwachsenen schnitten besser ab als die anderen Altersgruppen, was mit besseren Fähigkeiten, spätere Lern- und Erinnerungsvorgänge zu antizipieren, erklärt werden kann.

In Bezug auf den Einfluss von Kontextfaktoren konnte nachgewiesen werden, dass sich die Überwachungsleistungen bei verschiedenen komplexen Materialien (Paar-Assoziationen versus Film- oder Textmaterial) unter günstigen Bedingungen, z.B. wenn die Leistungsvorhersagen stark mit der Erinnerungsabfrage korrespondieren, kaum unterscheiden. Bei Rekognitionsaufgaben fielen die Überwachungsleistungen im Vergleich zu Aufgaben zur freien Erinnerung insgesamt schlechter aus. Ein großes bereichsspezifisches Vorwissen resultierte

über alle Maße hinweg eher in einer Überschätzung der eigenen Leistung, bei den SUs jedoch auch in einer verbesserten Leistung im Vergleich zu Personen mit weniger Vorwissen. Ein Strategietraining wirkte sich besonders bei den Grundschulern und den älteren Erwachsenen positiv auf die Überwachungsleistung aus. Die eher gering ausgeprägten Alterseffekte weisen darauf hin, dass die einzelnen Kontextfaktoren über die Lebensspanne hinweg einen vergleichbaren Einfluss zu haben scheinen.

Hinsichtlich sequenzieller Zusammenhänge zwischen Überwachungs- und Steuerungsprozessen (hier operationalisiert durch JOLs und die selbst gesteuerte Lernzeiteinteilung) zeigte sich, dass die Teilnehmer aller Altersgruppen in der Lage waren, sowohl Informationen aus den JOLs für die Anpassung der Lernzeit (Monitoring-affects-control-Modell) als auch – in etwas geringerem Ausmaß – Informationen aus der Lernzeit für die Anpassung der JOLs zu nutzen (Control-affects-monitoring-Modell). Der simultane Wechsel zwischen beiden Modellen stellt einen deutlich komplexeren Vorgang dar und konnte deshalb vor allem bei den Jugendlichen und den älteren Erwachsenen nachgewiesen werden.

Insgesamt gesehen belegen die Ergebnisse der sechs Experimente, dass metakognitive Überwachungsfähigkeiten bereits recht früh, d.h. im mittleren Grundschulalter, gut ausgeprägt sind und auch bei älteren Erwachsenen noch lange auf gutem Niveau erhalten bleiben. Lediglich der flexible Wechsel zwischen Überwachungs- und Kontrollprozessen scheint in diesen beiden Altersgruppen noch Schwierigkeiten zu bereiten. Die ähnliche Wirkweise der Kontextfaktoren in den einzelnen Altersgruppen weist auf vergleichbare zugrunde liegende Prozesse hin. Die grundsätzlich guten metakognitiven Leistungen bei Kindern und älteren Erwachsenen sollten demnach genutzt werden, um Gedächtnisprozesse insbesondere in diesen Altersgruppen zu fördern.

LITERATURVERZEICHNIS

- Allwood, C. M. & Granhag, P. A. (1996). The effects of arguments on realism in confidence judgements. *Acta Psychologica*, *91*, 99-119.
- Allwood, C. M., Jonsson, A.-C. & Granhag, P. A. (2005). The effects of source and type of feedback on child witnesses' metamemory accuracy. *Applied Cognitive Psychology*, *19*, 331-344.
- Anderson, M. C. M. & Thiede, K. W. (2008). Why do delayed summaries improve metacomprehension accuracy? *Acta Psychologica*, *128*, 110-118.
- Annevirta, T. & Vauras, M. (2001). Metacognitive knowledge in primary grades: A longitudinal study. *European Journal of Psychology of Education*, *16*, 237-282.
- Arbuckle, T. Y. & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, *81*, 126-131.
- Ariel, R., Dunlosky, J. & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General*, *138*, 432-447.
- Artelt, C., Neuenhaus, N., Lingel, K. & Schneider, W. (2012). Entwicklung und wechselseitige Effekte von metakognitiven und bereichsspezifischen Wissenskomponenten in der Sekundarstufe. *Psychologische Rundschau*, *63*, 18-25.
- Aster, M. von, Neubauer, A. C. & Horn, R. (2006). *Wechsler Intelligenztest für Erwachsene (WIE)*. Göttingen: Hogrefe.
- Astington, J. W. & Olson, D. R. (1990). Metacognitive and metalinguistic language: Learning to talk about thought. *Applied Psychology: An International Review*, *39*, 77-87.
- Baker, J. M. C., Dunlosky, J. & Hertzog, C. (2010). How accurately can older adults evaluate the quality of their text recall? The effect of providing standards on judgment accuracy. *Applied Cognitive Psychology*, *24*, 77-89.
- Baker, L. (1984). Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology*, *38*, 289-311.
- Baltes, P. B. & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1-34). Cambridge, MA: Cambridge University Press.
- Bannert, M. & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning? *Metacognition and Learning*, *3*, 39-58.
- Begg, I., Duft, S., Lalonde, P., Melnick, R. & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*, 610-632.
- Belmont, J. M. & Borkowski, J. G. (1988). A group-administered test of children's metamemory. *Bulletin of the Psychonomic Society*, *26*, 206-208.

- Brewer, W. F., Sampaio, C. & Barlow, M. R. (2005). Confidence and accuracy in the recall of deceptive and nondeceptive sentences. *Journal of Memory and Language*, 52, 618-627.
- Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in child development* (Vol. 10, pp. 104-152). New York: Academic Press.
- Brown, A. L. (1978). Knowing when, where and how to remember: A problem of metacognition. *Advances in instructional psychology*, 1, 77-165.
- Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. Bruce & W. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 453-481). Hillsdale, NJ: Erlbaum.
- Brown, A. L., Bransford, J. D., Ferrara, R. A. & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & M. E. Markman (Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 77-166). New York: Wiley.
- Brown, A. L. & Smiley, S. S. (1977). Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development*, 48, 1-8.
- Brown, A. L. & Smiley, S. S. (1978). The development of strategies for studying texts. *Child Development*, 49, 1076-1088.
- Buratti, S. & Allwood, C. M. (2012). The accuracy of meta-metacognitive judgments: Regulating the realism of confidence. *Cognitive Processing*, 13, 243-253.
- Butterfield, E. C., Nelson, T. O. & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology*, 24, 654-663.
- Burkett, C. & Azevedo, R. (2012). The effect of multimedia discrepancies on metacognitive judgments. *Computers in Human Behavior*, 28, 1276-1285.
- Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging*, 20, 718-721.
- Castel, A. D. (2007). Aging and memory for numerical information: The role of specificity and expertise in associative memory. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 62B, 194-196.
- Castel, A. D., Benjamin, A. S., Craik, F. I. M. & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition*, 30, 1078-1085.
- Castel, A. D., Farb, N. A. S. & Craik, F. I. M. (2007). Memory for general and specific value information in younger and older adults: Measuring the limits of strategic control. *Memory & Cognition*, 35, 689-700.
- Castel, A. D., Humphreys, K. L., Lee, S. S., Galvan, A., Balota, D. A. & McCabe, D. P. (2011). The development of memory efficiency and value-directed remembering across the life span: A cross-sectional study of memory and selectivity. *Developmental Psychology*, 47, 1553-1564.
- Castel, A. D., Lee, S. S., Humphreys, K. L. & Moore, A. N. (2011). Memory capacity, selective control, and value-directed remembering in children with and without attention-deficit/hyperactivity disorder (ADHD). *Neuropsychology*, 25, 15-24.

- Castel, A. D., McCabe, D. P. & Roediger, H. L., III. (2007). Illusions of competence and overestimation of associative memory for identical items: Evidence from judgments of learning. *Psychonomic Bulletin & Review*, *14*, 107-111.
- Cavanaugh, J. C. & Perlmutter, M. (1982). Metamemory: A critical examination. *Child Development*, *53*, 11-28.
- Chase, W. G. & Ericsson, K.A. (1981). *Skilled memory*. Hillsdale, NJ: Erlbaum.
- Chase, W. G. & Ericsson, K.A. (1982). Skill and working memory. In G. W. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16; pp. 1-58). New York: Academic Press.
- Chi, M. T. (1978). Knowledge structure and memory development. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73-96). Hillsdale, NJ: Erlbaum.
- Chi, M. T. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge Handbook of expertise and expert performance* (pp. 21-30). Cambridge, UK: Cambridge University Press.
- Chua, E. F., Schacter, D. L. & Sperling, R. A. (2009). Neural basis for recognition confidence in younger and older adults. *Psychology and Aging*, *24*, 139-153.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Connor, L. T., Dunlosky, J. & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*, 50-71.
- Cowan, N., Naveh-Benjamin, M., Kilb, A. & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology*, *42*, 1089-1102.
- Craik, F. I. M. & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences*, *10*, 131-138.
- Crawford, J. D. & Stankov, L. (1996). Age differences in the realism of confidence judgements: A calibration study using tests of fluid and crystallized intelligence. *Learning and Individual Differences*, *8*, 83-103.
- Cultice, J. C., Somerville, S. C. & Wellmann, H. M. (1983). Preschoolers' memory monitoring: Feeling-of-knowing judgments. *Child Development*, *54*, 1480-1486.
- Cunningham, J. G. & Weaver, S. L. (1989). Young children's knowledge of their memory span: Effects of task and experience. *Journal of Experimental Child Psychology*, *48*, 32-44.
- Daniels, K. A., Toth, J. P. & Hertzog, C. (2009). Aging and recollection in the accuracy of judgments of learning. *Psychology and Aging*, *24*, 494-500.
- de Bruin, A. B. H., Rikers, R. M. J. P. & Schmidt, H. G. (2007). Improving metacomprehension accuracy and self-regulation in cognitive skill acquisition: The effect of learner expertise. *European Journal of Cognitive Psychology*, *19*, 671-688.
- de Bruin, A. B. H., Thiede, K. W., Camp, G. & Redford, J. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, *109*, 294-310.

- Dember, W. N. (1974). Motivation and the cognitive revolution. *American Psychologist*, 29(3), 161-168.
- Destan, N., Hembacher, E., Ghetti, S. & Roebbers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5- to 7-year-old children. *Journal of Experimental Child Psychology*, 126, 213-228.
- Devolder, P. A., Brigham, M. C. & Pressley, M. (1990). Memory performance awareness in younger and older adults. *Psychology and Aging*, 5, 291-303.
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466-503). New York: Oxford University Press.
- Dixon, R. A. & Hulstsch, D. F. (1983). Structure and development of metamemory in adulthood. *Journal of Gerontology*, 38, 682-688.
- Dodson, C. S., Bawa, S. & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging*, 22, 122-133.
- Dodson, C. S. & Krueger, L. E. (2006). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin & Review*, 13, 770-775.
- Dufresne, A. & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: differential and sufficient aspects. *Journal of Experimental Child Psychology*, 47, 274-296.
- Dunlosky, J. (2005). Why does rereading improve metacomprehension accuracy? Evaluating the levels-of-disruption hypothesis for the rereading effect. *Discourse Processes*, 40, 37-55.
- Dunlosky, J. & Connor, L. T. (1997). Age differences in the allocation of study time account for age differences in memory performance. *Memory & Cognition*, 25, 691-700.
- Dunlosky, J. & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *The educational psychology series* (pp. 249-275). Mahwah, NJ: Erlbaum.
- Dunlosky, J. & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging*, 15, 462-474.
- Dunlosky, J., Kubat-Silman, A. K. & Hertzog, C. (2003). Training monitoring skills improves older adults' self-paced associative learning. *Psychology and Aging*, 18, 340-345.
- Dunlosky, J. & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, 16, 228-232.
- Dunlosky, J. & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage.
- Dunlosky, J. & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language*, 33, 545-565.
- Dunlosky, J. & Nelson, T. O. (1997). Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. *Journal of Memory and Language*, 36, 34-49.

- Dunlosky, J., Rawson, K. A. & Hacker, D. J. (2002). Metacomprehension of science text: Investigating the levels-of-disruption hypothesis. In J. Otero, J. A. Leon & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 255-279). Mahwah, NJ: Erlbaum.
- Dunlosky, J., Rawson, K. A. & McDonald, S. L. (2002). Influence of practice tests on the accuracy of predicting memory performance for paired associates, sentences, and text material. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 68-92). New York: Cambridge University Press.
- Dunlosky, J., Rawson, K. A. & Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language*, 52, 551-565.
- Dunlosky, J. & Tauber, S. K. (2014). Understanding people's metacognitive judgments: An isomechanism framework and its implications for applied and theoretical research. In T. J. Perfect & D. S. Lindsay (Eds.), *The SAGE Handbook of applied memory* (pp. 444-464). Los Angeles, CA: Sage.
- Elischberger, H. B. (2005). The effects of prior knowledge on children's memory and suggestibility. *Journal of Experimental Child Psychology*, 92, 247-275.
- Ericsson, K. A., Krampe, R. T. & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363-406.
- Feltovich, P. J., Prietula, M. J. & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge Handbook of expertise and expert performance* (pp. 41-66). Cambridge, UK: Cambridge University Press.
- Fitts, P. M. & Posner, M. I. (1967). *Human performance*. Oxford, UK: Brooks/Cole.
- Flavell, J. H. (1971). First Discussant's Comments: What is memory development the development of? *Human Development*, 14, 272-278.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Flavell, J. H., Friedrichs, A. G. & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, 1, 324-340.
- Flavell, J. H., Miller, P. H. & Miller, S. A. (2002). *Cognitive Development*. Upper Saddle River, NJ: Prentice Hall.
- Flavell, J. H. & Wellmann, H. M. (1977). Metamemory. In R. Kail (Ed.), *Perspectives on the development of memory and cognition* (Vol. 1, pp. 3-33). Hilldale, NJ: Erlbaum.
- Fleming, S. M. & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B*, 367, 1338-1349.
- Froger, C., Bouazzaoui, B., Isingrini, M. & Tacconat, L. (2012). Study time allocation deficit of older adults: The role of environmental support at encoding? *Psychology and Aging*, 27, 577-588.
- Froger, C., Sacher, M., Gaudouen, M.-S., Isingrini, M. & Tacconat, L. (2011). Metamemory judgments and study time allocation in young and older adults: Dissociative effects of a generation task. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale*, 65, 269-276.

- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases”. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 2, pp. 83-115). Chichester, UK: Wiley.
- Gilewski, M. J., Zelinski, E. M. & Schaie, K. W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging*, 5, 482-490.
- Glenberg, A. M. & Epstein, W. (1985). Calibration of Comprehension. *Journal of Experimental Psychology*, 11, 702-718.
- Glenberg, A. M. & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory and Cognition*, 15, 84-93.
- Gonzales, R. & Nelson, T. O. (1996). Measuring ordinal association in situations that contain tied scores. *Psychological Bulletin*, 119, 159-165.
- Grammer, J. K., Purtell, K. M., Coffman, J. L. & Ornstein, P. A. (2011). Relations between children’s metamemory and strategic performance: Time-varying covariates in early elementary school. *Journal of Experimental Child Psychology*, 108, 139-155.
- Griffin, D. & Brenner, L. (2004). Perspectives on probability judgment calibration. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making*. (pp. 177-198). Malden, MA: Blackwell.
- Griffin, T. D., Jee, B. D. & Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory & Cognition*, 37, 1001-1013.
- Griffin, T. D., Wiley, J. & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, 36, 93-103.
- Groeben, N. (1982). *Leserpsychologie: Textverständnis - Textverständlichkeit*. München: Aschendorff.
- Groninger, L. D. (1976). Predicting recognition during storage: The capacity of the memory system to evaluate itself. *Bulletin of the Psychonomic Society*, 7, 425-428.
- Gruber, H. (2007). Bedingungen von Expertise. In K. A. Heller & A. Ziegler (Hrsg.), *Begabt sein in Deutschland* (Bd. 1, S. 93-112). Berlin: LIT.
- Gruber, H. (2010). Expertise. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 183-189). Weinheim: Beltz.
- Gruber, H., Degner, S. & Lehmann, A. C. (2004). Why do some commit themselves in deliberate practice for many years - and so many do not? Understanding the development of professionalism in music. In M. Radovan & N. Dordevic (Eds.), *Current issues in adults learning and motivation* (pp. 222-235). Ljubljana: Slovenian Institute for Adult Education.
- Haberkorn, K., Lockl, K., Pohl, S., Ebert, S. & Weinert, S. (2014). Metacognitive knowledge in children at early elementary school. *Metacognition and Learning*, 9, 239-263.
- Hambrick, D. Z. & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology*, 44, 339-387.

- Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F. & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34-45.
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, 208-216.
- Heckhausen, J., Dixon, R. A. & Baltes, P. B. (1989). Gains and losses in development throughout adulthood as perceived by different adult age groups. *Developmental Psychology*, *25*, 109-121.
- Hertzog, C., Dixon, R. A. & Hultsch, D. F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging*, *5*, 215-227.
- Hertzog, C. & Dunlosky, J. (2006). Using visual imagery as a mnemonic for verbal associative learning: Developmental and individual differences. In T. Vecchi & G. Bottini (Eds.), *Imagery and spatial cognition: Methods, models and cognitive assessment*. (pp. 263-284). Amsterdam, The Netherlands: John Benjamins.
- Hertzog, C. & Dunlosky, J. (2011). Metacognition in later adulthood: Spared monitoring can benefit older adults' self-regulation. *Current Directions in Psychological Science*, *20*, 167-173.
- Hertzog, C. & Hultsch, D. F. (2000). Metacognition in adulthood and old age. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 417-466). Mahwah, NJ: Erlbaum.
- Hertzog, C., Kidder, D. P., Powell-Moman, A. & Dunlosky, J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired? *Psychology and Aging*, *17*, 209-225.
- Hertzog, C., Kramer, A. F., Wilson, R. S. & Lindenberger, U. (2008). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest*, *9*, 1-65.
- Hertzog, C., Saylor, L. L., Fleece, A. M. & Dixon, R. A. (1994). Metamemory and aging: Relations between predicted, actual and perceived memory task performance. *Aging & Cognition*, *1*, 203-237.
- Hertzog, C., Sinclair, S. M. & Dunlosky, J. (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. *Developmental Psychology*, *46*, 939-948.
- Hines, J., Hertzog, C. & Touron, D. (2012). A prelearning manipulation falsifies a pure associational deficit account of retrieval shift during skill acquisition. *Aging, Neuropsychology, and Cognition*, *19*, 449-478.
- Hines, J. C., Touron, D. R. & Hertzog, C. (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. *Psychology and Aging*, *24*, 462-475.
- Hoffmann-Biencourt, A., Lockl, K., Schneider, W., Ackerman, R. & Koriat, A. (2010). Self-paced study time as a cue for recall predictions across school age. *British Journal of Developmental Psychology*, *28*, 767-784.

- Howie, P. & Roebbers, C. M. (2007). Developmental progression in the confidence-accuracy relationship in event recall: New insights provided by a calibration perspective. *Applied Cognitive Psychology, 21*, 871-893.
- Hultsch, D. F. & Dixon, R. A. (1983). The role of pre-experimental knowledge in text processing in adulthood. *Experimental Aging Research, 9*, 17-22.
- Hultsch, D. F., Hertzog, C. & Dixon, R. A. (1987). Age differences in metamemory: Resolving the inconsistencies. *Canadian Journal of Psychology/Revue canadienne de psychologie, 41*, 193-208.
- Job, J. M. & Klassen, R. M. (2012). Predicting performance on academic and non-academic tasks: A comparison of adolescents with and without learning disabilities. *Contemporary Educational Psychology, 37*, 162-169.
- Jönsson, F. U. & Lindström, B. R. (2010). Using a multidimensional scaling approach to investigate the underlying basis of ease of learning judgments. *Scandinavian Journal of Psychology, 51*, 103-108.
- Justice, E. M. (1989). Preschoolers' knowledge and use of behaviors varying in strategic effectiveness. *Merrill-Palmer Quarterly, 35*, 363-377.
- Karlen, Y., Maag Merki, K. & Ramseier, E. (2014). The effect of individual differences in the development of metacognitive strategy knowledge. *Instructional Science, 42*, 777-794.
- Karpel, M. E., Hoyer, W. J. & Togli, M. P. (2001). Accuracy and qualities of real and suggested memories: Nonspecific age differences. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 56B*, 103-110.
- Keast, A., Brewer, N. & Wells, G. L. (2007). Children's metacognitive judgments in an eyewitness identification task. *Journal of Experimental Child Psychology, 97*, 286-314.
- Kelemen, W. L., Frost, P. J. & Weaver, C. A., III. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition, 28*, 92-107.
- Kelley, C. M. & Lindsay, D. S. (1993). Remembering mistaken for knowing: ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32*, 1-24.
- Kelley, C. M. & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*, 704-721.
- King, J. F., Zechmeister, E. B. & Shaughnessy, J. J. (1980). Judgments of knowing: The influence of retrieval practice. *The American Journal of Psychology, 93*, 329-343.
- Kintsch, W. (1998). *Comprehension: A paradigm for Cognition*. New York: Cambridge University Press.
- Klassen, R. M. (2007). Using predictions to learn about the self-efficacy of early adolescents with and without learning disabilities. *Contemporary Educational Psychology, 32*, 173-187.
- Kliegel, M., Mackinlay, R. & Jäger, T. (2008). Complex prospective memory: Development across the lifespan and the role of task interruption. *Developmental Psychology, 44*, 612-617.

- Kluwe, R. H. (1982). Cognitive knowledge and executive control: Metacognition. In D. R. Griffin (Ed.), *Animal Mind - Human Mind* (pp. 201-224). Berlin: Springer.
- Kluwe, R. H. & Schiebler, K. (1984). Entwicklung exekutiver Prozesse und kognitiver Leistungen. In F. E. Weinert & R. H. Kluwe (Hrsg.), *Metakognition, Motivation und Lernen* (S. 31-60). Stuttgart: Kohlhammer.
- Kobasigawa, A. & Metcalf-Haggert, A. (1993). Spontaneous allocation of study time by first- and third-grade children in a simple memory task. *The Journal of Genetic Psychology: Research and Theory on Human Development*, 154, 223-235.
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, 100, 609-639.
- Koriat, A. (1997). Monitoring one's own knowledge during study: a cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126, 349-370.
- Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*, 36, 416-428.
- Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, 119, 80-113.
- Koriat, A. & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science*, 13, 441-453.
- Koriat, A., Ackerman, R., Adiv, S., Lockl, K. & Schneider, W. (2014). The effects of goal-driven and data-driven regulation on metacognitive monitoring during learning: A developmental perspective. *Journal of Experimental Psychology: General*, 143, 386-403.
- Koriat, A., Ackermann, R., Lockl, K. & Schneider, W. (2009a). The easily learned, easily remembered heuristic in children. *Cognitive Development*, 24, 169-182.
- Koriat, A., Ackerman, R., Lockl, K. & Schneider, W. (2009b). The memorizing effort heuristic in judgments of learning: A developmental perspective. *Journal of Experimental Child Psychology*, 102, 265-279.
- Koriat, A. & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 187-194.
- Koriat, A. & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, 103, 490-517.
- Koriat, A., Lichtenstein, S. & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology*, 6, 107-118.
- Koriat, A. & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, 52, 478-492.
- Koriat, A., Ma'ayan, H. & Nussinson, R. (2006). The intricate relationship between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology*, 135, 36-69.

- Koriat, A. & Nussinson, R. (2009). Attributing study effort to data-driven and goal-driven effects: Implications for metacognitive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1338-1343.
- Koriat, A., Nussinson, R. & Ackerman, R. (2014). Judgments of learning depend on how learners interpret study effort. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1624-1637.
- Koriat, A. & Shitzer-Reichert, R. (2002). Metacognitive judgments and their accuracy. Insights from the processes underlying judgments of learning in children. In P. Chambres, M. Izaute & P.-J. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 1-18). Boston, MA: Kluwer.
- Krebs, S. S. & Roebers, C. M. (2012). The impact of retrieval processes, age, general achievement level, and test scoring scheme for children's metacognitive monitoring and controlling. *Metacognition and Learning*, 7, 75-90.
- Kreutzer, M. A., Leonard, C. & Flavell, J. H. (1975). An interview study of children's memory about memory. *Monographs of the Society for Research in Child Development*, 40, 1-60.
- Kuhn, D. (1999a). A developmental model of critical thinking. *Educational Researcher*, 28, 16-46.
- Kuhn, D. (1999b). Metacognitive Development. In L. Balter & C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 259-286). Philadelphia, PA: Psychology Press.
- Kuhn, D. (2000). Theory of mind, metacognition and reasoning: A life-span perspective. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 301-326). Hove, UK: Psychology Press.
- Kurtz, B. E. & Schneider, W. (1988). The effects of age, study time, and importance of text units on strategy use and memory for texts. *European Journal of Psychology of Education*, 3, 191-199.
- Labuhn, A. S., Zimmerman, B. J. & Hasselhorn, M. (2010). Enhancing students' self-regulation and mathematics performance: The influence of feedback and self-evaluative standards. *Metacognition and Learning*, 5, 173-194.
- Lachman, J. L., Lachman, R. & Thronesbery, C. (1979). Metamemory through the adult life span. *Developmental Psychology*, 15, 543-551.
- Lachman, M. E. (2006). Perceived control over aging-related declines: Adaptive beliefs and behaviors. *Current Directions in Psychological Science*, 15, 282-286.
- Lachman, M. E. & Jelalian, E. (1984). Self-efficacy and attributions for intellectual performance in young and elderly adults. *Journal of Gerontology*, 39, 577-582.
- Le Ny, J. F., Denhiere, G. & Le Taillanter, D. (1972). Regulation of study-time and interstimulus similarity in self-paced learning conditions. *Acta Psychologica*, 36, 280-289.
- Leonesio, R. J. & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology*, 16, 464-470.
- Lichtenstein, S. & Fischhoff, B. (1980). Training for calibration. *Organizational Behavior & Human Performance*, 26, 149-171.

- Lichtenstein, S., Fischhoff, B. & Phillips, L. D. (1982). Calibration of probabilities: The state of the art. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). New York: Cambridge University Press.
- Lindberg, M. A. (1980). Is knowledge base development a necessary and sufficient condition for memory development? *Journal of Experimental Child Psychology*, 30, 401-410.
- Linden, N. von der & Roebers, C. M. (2006). Developmental changes in uncertainty monitoring during an event recall task. *Metacognition & Learning*, 1, 213-228.
- Linden, N. von der, Schneider, W. & Roebers, C. M. (2011). The effects of summary production and encoding condition on children's metacognitive monitoring. *Metacognition & Learning*, 6, 3-23.
- Lindenberger, U. & Staudinger, U. M. (2012). Höheres Erwachsenenalter. In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie* (S. 283-309). Weinheim: Beltz.
- Lineweaver, T. T. & Hertzog, C. (1998). Adults' efficacy and control beliefs regarding memory and aging: Separating general from personal beliefs. *Aging, Neuropsychology, and Cognition*, 5, 264-296.
- Lingel, K., Götz, L., Artelt, C. & Schneider, W. (2014). *Mathematisches Strategiewissen für 5. und 6. Klassen (MAESTRA 5-6+)*. Göttingen: Hogrefe.
- Lipko, A. R., Dunlosky, J. & Merriman, W. E. (2009). Persistent overconfidence despite practice: The role of task experience in preschoolers' recall predictions. *Journal of Experimental Child Psychology*, 103, 152-166.
- Lipowski, S. L., Merriman, W. E. & Dunlosky, J. (2013). Preschoolers can make highly accurate judgments of learning. *Developmental Psychology*, 49, 1505-1516.
- Lipson, M. Y. (1982). Learning new information from text: The role of prior knowledge and reading ability. *Journal of Reading Behavior*, 14, 243-261.
- Lockl, K. & Schneider, W. (2002a). Developmental trends in children's feeling-of-knowing judgements. *International Journal of Behavioral Development*, 26, 327-333.
- Lockl, K. & Schneider, W. (2002b). Zur Entwicklung des selbstregulierten Lernens im Grundschulalter: Zusammenhänge zwischen Aufgabenschwierigkeit und Lernzeiteinteilung. *Psychologie und Erziehung im Unterricht*, 49, 3-16.
- Lockl, K. & Schneider, W. (2003). Metakognitive Überwachungs- und Selbstkontrollprozesse bei der Lernzeiteinteilung von Kindern. *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology*, 17, 173-183.
- Lockl, K. & Schneider, W. (2004). The effects of incentives and instructions on children's allocation of study time. *European Journal of developmental psychology*, 1, 153-169.
- Lockl, K. & Schneider, W. (2006). Precursors of metamemory in young children: The role of theory of mind and metacognitive vocabulary. *Metacognition & Learning*, 1, 15-31.
- Lovelace, E. A. & Marsh, G. R. (1985). Prediction and evaluation of memory performance by young and old adults. *Journal of Gerontology*, 40, 192-197.
- Lyons, K. E. & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, 82, 1778-1787.

- Maki, R. H. (1998a). Predicting performance on text: Delayed vs. immediate predictions and tests. *Memory and Cognition*, 26, 959-964.
- Maki, R. H. (1998b). Test predictions over text material. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117-144). Mahwah, NJ: Erlbaum.
- Maki, R. H. & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 663-679.
- Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C. & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 609-616.
- Maki, R. H., Shields, M., Wheeler, A. E. & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, 97, 723-731.
- Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, 48, 986-992.
- Marquié, J. C. & Huet, N. (2000). Age differences in feeling-of-knowing and confidence judgments as a function of knowledge domain. *Psychology and Aging*, 15, 451-461.
- Masur, E. F., McIntyre, C. W. & Flavell, J. H. (1973). Developmental changes in apportionment of study time among items in a multitrial free recall task. *Journal of Experimental Child Psychology*, 15, 237-246.
- Mazzoni, G. & Cornoldi, C. (1993). Strategies in study time allocation: why is study time sometimes not effective? *Journal of Experimental Psychology*, 122, 47-60.
- Mazzoni, G., Cornoldi, C. & Marchitelli, G. (1990). Do memorability ratings affect study-time allocation? *Memory and Cognition*, 18, 196-204.
- Mazzoni, G., Cornoldi, C., Tomat, L. & Vecchi, T. (1997). Remembering the grocery shopping list: A study on metacognitive biases. *Applied Cognitive Psychology*, 11, 253-267.
- McCabe, D. P. & Soderstrom, N. C. (2011). Recollection-based prospective metamemory judgments are more accurate than those based on confidence: Judgments of remembering and knowing (JORKs). *Journal of Experimental Psychology: General*, 140, 605-621.
- McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P. & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *The Journal of Neuroscience*, 33, 1897-1906.
- McDonald-Miszczak, L., Hertzog, C. & Hultsch, D. F. (1995). Stability and accuracy of metamemory in adulthood and aging: A longitudinal analysis. *Psychology and Aging*, 10, 553-564.
- Means, M. L. & Voss, J. F. (1985). Star Wars: A developmental study of expert and novice knowledge structures. *Journal of Memory and Language*, 24, 746-757.
- Mengelkamp, C. & Bannert, M. (2010). Accuracy of confidence judgments: Stability and generality in the learning process and predictive validity for learning outcome. *Memory & Cognition*, 38, 441-451.

- Metcalfe, J. & Finn, B. (2008). Familiarity and retrieval processes in delayed judgments of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1084-1097.
- Miles, J. R. & Stine-Morrow, E. A. L. (2004). Adult age differences in self-regulated learning from reading sentences. *Psychology and Aging*, *19*, 626-636.
- Miscione, J. L., Marvin, R. S., O'Brien, R. G. & Greenberg, M. T. (1978). A developmental study of preschool children's understanding of the words 'know' and 'guess'. *Child Development*, *49*, 1107-1113.
- Murphy, M. D., Schmitt, F. A., Caruso, M. J. & Sanders, R. E. (1987). Metamemory in older adults: The role of monitoring in serial recall. *Psychology and Aging*, *2*, 331-339.
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1170-1187.
- Naveh-Benjamin, M., Brav, T. K. & Levy, O. (2007). The associative memory deficit of older adults: The role of strategy utilization. *Psychology and Aging*, *22*, 202-208.
- Neisser, U. (1967). *Cognitive psychology*. East Norwalk, CT: Appleton-Century-Crofts.
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of Feeling-of-Knowing predictions. *Psychological Bulletin*, *95*, 109-133.
- Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item. *Applied Cognitive Psychology*, *10*, 257-260.
- Nelson, T. O. & Dunlosky, J. (1991). When people's judgements of learning (JOLs) are extremely accurate at predicting subsequent recall: The "Delayed-JOL Effect". *Psychological Science*, *2*, 267-270.
- Nelson, T. O., Dunlosky, J., Graf, A. & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, *5*, 207-213.
- Nelson, T. O. & Leonesio, R. J. (1988). Allocation of self-paced study time and the 'labor-in-vain effect'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 676-686.
- Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125-141). New York: Academic Press.
- Nelson, T. O. & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: MIT Press.
- Neuenhaus, N., Artelt, C., Lingel, K. & Schneider, W. (2011). Fifth graders metacognitive knowledge: General or domain-specific? *European Journal of Psychology of Education*, *26*, 163-178.
- Nieding, G. (2006). *Wie verstehen Kinder Texte? Die Entwicklung kognitiver Repräsentationen*. Lengerich: Pabst.

- Nietfeld, J. L., Cao, L. & Osborne, J. W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *Journal of Experimental Education*, 74, 7-28.
- Nietfeld, J. L. & Schraw, G. (2002). The effect of knowledge and strategy training on monitoring accuracy. *The Journal of Educational Research*, 95, 131-142.
- Nolan, J. & Markham, R. (1998). The accuracy-confidence relationship in an eyewitness task: Anxiety as a modifier. *Applied Cognitive Psychology*, 12, 43-54.
- Olin, J. T. & Zelinski, E. M. (1997). Age differences in calibration of comprehension. *Educational Gerontology*, 23, 67-77.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, 29, 261-265.
- O'Sullivan, J. T. (1993). Preschoolers' beliefs about effort, incentives, and recall. *Journal of Experimental Child Psychology*, 55, 396-414.
- Paivio, A. (1971). *Imagery and verbal processes*. Oxford, UK: Holt Rinehart & Winston.
- Pannu, J. K. & Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychology Review*, 15, 105-130.
- Pansky, A., Goldsmith, M., Koriat, A. & Pearlman-Avni, S. (2009). Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants. *European Journal of Cognitive Psychology*, 21, 303-329.
- Paris, S. G. & Lindauer, B. K. (1982). The development of cognitive skills during childhood. In B. B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 333-347). Englewood Cliffs, NJ: Prentice-Hall.
- Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316.
- Paris, S. G. & Oka, E. R. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review*, 6, 25-56.
- Paulus, M., Tsalas, N., Proust, J. & Sodian, B. (2014). Metacognitive monitoring of oneself and others: Developmental changes during childhood and adolescence. *Journal of Experimental Child Psychology*, 122, 153-165.
- Petermann, F. & Petermann, U. (2010). *Hamburg-Wechsler-Intelligenztest für Kinder – IV (HAWIK IV)*. Göttingen: Hogrefe.
- Piaget, J. (2000). *Psychologie der Intelligenz* (10. Aufl.). Stuttgart: Klett-Cotta.
- Pliske, R. M. & Mutter, S. A. (1996). Age differences in the accuracy of confidence judgments. *Experimental Aging Research*, 22, 199-216.
- Posner, M. I. (1988). Introduction: What is it to be an expert? In M. T. H. Chi, R. Glaser & M. J. Farr (Eds.), *The nature of expertise* (pp. xxix-xxxvi). Hillsdale, NJ: Erlbaum.
- Pressley, M., Borkowski, J. G. & O'Sullivan, J. (1985). Children's metamemory and the teaching of memory strategies. In F. Pressley & D. Lynn (Eds.), *Metacognition, cognition and human performance* (Vol. 1, pp. 111-153). Orlando, FL: Academic Press.

- Pressley, M., Borkowski, J. G. & Schneider, W. (1987). Cognitive Strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. 5, pp. 89-129). New York: JAI Press.
- Pressley, M., Borkowski, J. G. & Schneider, W. (1989). Good Information Processing: What it is and how education can promote it. *International Journal of Educational Research*, 13, 857-867.
- Pressley, M. & Ghatala, E. S. (1988). Delusions about performance on multiple-choice comprehension tests. *Reading Research Quarterly*, 23, 454-464.
- Pressley, M. & Levin, J. R. (1977). Developmental differences in subjects' associative-learning strategies and performance: Assessing a hypothesis. *Journal of Experimental Child Psychology*, 24, 431-439.
- Pressley, M., Levin, J. R., Ghatala, E. S. & Ahmad, M. (1987). Test monitoring in young grade school children. *Journal of Experimental Child Psychology*, 43, 96-111.
- Price, J., Hertzog, C. & Dunlosky, J. (2010). Self-regulated learning in younger and older adults: Does aging affect metacognitive control? *Aging, Neuropsychology, and Cognition*, 17, 329-359.
- Rabbitt, P. & Abson, V. (1991). Do older people know how good they are? *British Journal of Psychology*, 82, 137-151.
- Rabbitt, P., Maylor, E., McInnes, L., Bent, N. & Moore, B. (1995). What goods can self-assessment questionnaires deliver for cognitive gerontology? *Applied Cognitive Psychology*, 9, 127-S152.
- Rabinowitz, J. C., Ackerman, B. P., Craik, F. I. M. & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology*, 37, 688-695.
- Rawson, K. A. & Dunlosky, J. (2007). Improving students' self-evaluation of learning of key concepts in textbook materials. *European Journal of Cognitive Psychology*, 19, 559-579.
- Rawson, K. A., Dunlosky, J. & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory and Cognition*, 28, 1004-1010.
- Recht, D. R. & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 80, 16-20.
- Reid, L. M. & MacLulich, A. M. J. (2006). Subjective memory complaints and cognitive impairment in older people. *Dementia and Geriatric Cognitive Disorders*, 22, 471-485.
- Resnick, L. B. (1983). Toward a cognitive theory of instruction. In S. G. Paris, G. Olson & H. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 5-38). Hillsdale, NJ: Erlbaum.
- Rhodes, M. G. & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, 137, 131-148.
- Richardson, J. & Erlebacher, A. (1958). Associative connection between paired verbal items. *Journal of Experimental Psychology*, 56, 62-69.

- Richardson, J. T. E. (1998). The availability and effectiveness of reported mediators in associative learning: A historical review and an experimental investigation. *Psychonomic Bulletin & Review*, 5, 597-614.
- Robinson, A. E., Hertzog, C. & Dunlosky, J. (2006). Aging, encoding fluency, and metacognitive monitoring. *Aging, Neuropsychology, and Cognition*, 13, 458-478.
- Robinson, M. D. & Johnson, J. T. (1996). Recall memory, recognition memory, and the eyewitness confidence-accuracy correlation. *Journal of Applied Psychology*, 81, 587-594.
- Robinson, M. D., Johnson, J. T. & Robertson, D. A. (2000). Process versus content in eyewitness metamemory monitoring. *Journal of Experimental Psychology: Applied*, 6, 207-221.
- Roderer, T. & Roebbers, C. M. (2010). Explicit and implicit confidence judgments and developmental differences in metamemory: An eye-tracking approach. *Metacognition and Learning*, 5, 229-250.
- Roebbers, C. M. (2002). Confidence judgments in children's and adults' event recall and suggestibility. *Developmental Psychology*, 38, 1052-1067.
- Roebbers, C. M., Gelhaar, T. & Schneider, W. (2004). 'It's magic!' The effects of presentation modality on children's event memory, suggestibility, and confidence judgments. *Journal of Experimental Child Psychology*, 87, 320-335.
- Roebbers, C. M. & Howie, P. (2003). Confidence judgments in event recall: Developmental progression in the impact of question format. *Journal of Experimental Child Psychology*, 85, 352-371.
- Roebbers, C. M., Linden, N. von der & Howie, P. (2007). Favourable and unfavourable conditions for children's confidence judgments. *British Journal of Developmental Psychology*, 25, 109-134.
- Roebbers, C. M., Linden, N. von der, Schneider, W. & Howie, P. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology*, 97, 117-137.
- Schlagmüller, M. & Schneider, W. (2002). The development of organizational strategies in children: Evidence from a microgenetic longitudinal study. *Journal of Experimental Child Psychology*, 81, 298-319.
- Schlagmüller, M. & Schneider, W. (2007). Würzburger Lesestrategie-Wissenstest für die Klassen 7-12 (WLST 7-12). Göttingen: Hogrefe.
- Schlagmüller, M., Visé, M. & Schneider, W. (2001). Zur Erfassung des Gedächtniswissens bei Grundschulkindern: Konstruktionsprinzipien und empirische Bewährung der Würzburger Testbatterie zum deklarativen Metagedächtnis. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 33, 91-102.
- Schneider, W. (1985). Developmental trends in the metamemory-memory behavior relationship: An integrative review. In F. Pressley & D. Lynn (Eds.), *Metacognition, cognition and human performance* (Vol. 1, pp. 57-109). Orlando, FL: Academic Press.
- Schneider, W. (1986). The role of conceptual knowledge and metamemory in the development of organizational processes in memory. *Journal of Experimental Child Psychology*, 42, 218-236.

- Schneider, W. (1989). *Zur Entwicklung des Meta-Gedächtnisses bei Kindern*. Bern: Huber.
- Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science*, 1, 291-297.
- Schneider, W. (1999). Introspektion und Metakognition in der Sicht der „Würzburger Schule“ und zeitgenössischer Forschung. In W. Janke & W. Schneider (Hrsg.), *Hundert Jahre Institut für Psychologie und Würzburger Schule der Denkpsychologie* (S. 387-397). Göttingen: Hogrefe.
- Schneider, W. (2000). Giftedness, expertise, and (exceptional) performance: A developmental perspective. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (pp. 165-177). Amsterdam, The Netherlands: Elsevier.
- Schneider, W. (2010). Metacognition and memory development in childhood and adolescence. In H. S. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 54-81). New York: Guilford.
- Schneider, W. (2015). *Memory development from early childhood through emerging adulthood*. Cham, Switzerland: Springer.
- Schneider, W., Gruber, H., Gold, A. & Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. *Journal of Experimental Child Psychology*, 56, 328-349.
- Schneider, W., Körkel, J. & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology*, 81, 306-312.
- Schneider, W. & Lockl, K. (2006). Entwicklung metakognitiver Kompetenzen im Kindes- und Jugendalter. In W. Schneider & B. Sodian (Hrsg.), *Kognitive Entwicklung (Enzyklopädie für Psychologie, Serie Entwicklungspsychologie, Band 2, S. 721-767)*. Göttingen: Hogrefe.
- Schneider, W., Schlagmüller, M. & Visé, M. (1998). The impact of metamemory and domain-specific knowledge on memory performance. *European Journal of Psychology and Education*, 13, 91-103.
- Schneider, W. & Sodian, B. (1988). Metamemory-memory behavior relationships in young children: Evidence from a memory-for-location task. *Journal of Experimental Child Psychology*, 45, 209-233.
- Schneider, W. & Stumpf, E. (2007). Hochbegabung, Expertise und die Erklärung außergewöhnlicher Leistungen. In K. A. Heller & A. Ziegler (Hrsg.), *Begabt sein in Deutschland* (Bd. 1, S. 71-91). Berlin: LIT.
- Schneider, W. & Uhl, C. (1990). Metagedächtnis, Strategienutzung und Gedächtnisleistung: Vergleichende Analysen bei Kindern, jüngeren Erwachsenen und alten Menschen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 22, 22-41.
- Schneider, W., Visé, M., Lockl, K. & Nelson, T. O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning (JOL) task. *Cognitive Development*, 15, 115-134.
- Schraw, G. (1994). The effect of metacognitive knowledge on local and global monitoring. *Contemporary Educational Psychology*, 19, 143-154.

- Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning, 4*, 33-45.
- Schraw, G. & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460-475.
- Schunk, D. H. & Zimmerman, B. J. (Eds.). (1998). *Self-regulated learning: From teaching to self-reflective practice*. New York: Guilford.
- Schwanenflugel, P. J., Fabricius, W. V. & Alexander, J. (1994). Developing theories of mind: Understanding concepts and relations between mental activities. *Child Development, 65*, 1546-1563.
- Schwanenflugel, P. J., Fabricius, W. V. & Noyes, C. R. (1996). Developing organization of mental verbs: evidence for the development of a constructivist theory of mind in middle childhood. *Cognitive Development, 11*, 265-294.
- Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin & Review, 1*, 357-375.
- Schwartz, B. L., Pilot, M. & Bacon, E. (2014). Contextual information influences the feeling of knowing in episodic memory. *Consciousness and Cognition, 29*, 96-104.
- Schwarz, S. & Roebbers, C. M. (2006). Age differences in the effects of social influence on children's eyewitness performance and their metacognitive monitoring. *Journal of Experimental Child Psychology, 94*, 229-248.
- Scott, B. M. & Levy, M. G. (2013). Metacognition: examining the components of a fuzzy concept. *Educational Research Journal, 2*, 120-131.
- Shake, M. C., Noh, S. R. & Stine-Morrow, E. A. L. (2009). Age differences in learning from text: Evidence for functionally distinct text processing systems. *Applied Cognitive Psychology, 23*, 561-578.
- Shaughnessy, J. J. & Zechmeister, E. B. (1992). Memory-monitoring accuracy as influenced by the distribution of retrieval practice. *Bulletin of the Psychonomic Society, 30*, 125-128.
- Shin, H., Bjorklund, D. F. & Beck, E. F. (2007). The adaptive nature of children's overestimation in a strategic memory task. *Cognitive Development, 22*, 197-212.
- Shing, Y. L., Werkle-Bergner, M., Li, S.-C. & Lindenberger, U. (2009). Committing memory errors with high confidence: Older adults do but children don't. *Memory, 17*, 169-179.
- Shiu, L. & Chen, Q. (2013). Self and external monitoring of reading comprehension. *Journal of Educational Psychology, 105*, 78-88.
- Simon, H. A. & Chase, W. G. (1973). Skill in chess. *American Scientist, 61*, 394-403.
- Slater, A., Morison, V. & Rose, D. (1982). Visual memory at birth. *British Journal of Psychology, 73*, 519-525.
- Soederberg Miller, L. M. (2009). Age differences in the effects of domain knowledge on reading efficiency. *Psychology and Aging, 24*, 63-74.
- Son, L. K. & Kornell, N. (2010). The virtues of ignorance. *Behavioural Processes, 83*, 207-212.

- Son, L. K. & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 204-221.
- Souchay, C. & Isingrini, M. (2004). Age related differences in metacognitive control: Role of executive functioning. *Brain and Cognition*, 56, 89-99.
- Souchay, C., Isingrini, M. & Espagnet, L. (2000). Aging, episodic memory feeling-of-knowing, and frontal functioning. *Neuropsychology*, 14, 299-309
- Souchay, C., Moulin, C. J., Clarys, D., Tacconat, L. & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: Evidence from feeling-of-knowing and recollection. *Consciousness and cognition*, 16, 769-784.
- Spellman, B. A. & Bjork, R. A. (1992). When predictions create reality: judgments of learning may alter what they are intended to assess. *Psychological Science*, 3, 315-316.
- Sperling, R. A., Howard, B. C., Miller, L. A. & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, 27, 51-79.
- Statistisches Bundesamt (2016). 13. *Koordinierte Bevölkerungsvorausberechnung für Deutschland*. Zugriff am 25.01.2016. <https://www.destatis.de/bevoelkerungspyramide>.
- Stine-Morrow, E. A. L., Shake, M. C., Miles, J. R. & Noh, S. R. (2006). Adult age differences in the effects of goals on self-regulated sentence processing. *Psychology and Aging*, 21, 790-803.
- Stipek, D. J. (1984). Young children's performance expectations: Logical analysis or wishful thinking? In J. G. Nicholls (Ed.), *The development of achievement motivation* (pp. 121-142). Greenwich, CT: JAI Press.
- Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review*, 12, 437-475.
- Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, 6, 662-667.
- Thiede, K. W., Anderson, M. C. M. & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, 95, 66-73.
- Thiede, K. W. & Dunlosky, J. (1994). Delaying students' metacognitive monitoring improves their accuracy in predicting their recognition performance. *Journal of Educational Psychology*, 86, 290-302.
- Thiede, K. W. & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1024-1037.
- Thiede, K. W., Dunlosky, J., Griffin, T. D. & Wiley, J. (2005). Understanding the delayed-keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1267-1280.
- Thiede, K. W., Griffin, T. D., Wiley, J. & Anderson, M. C. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes*, 47, 331-362.

- Thiede, K. W., Griffin, T. D., Wiley, J. & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 85-106). New York: Routledge.
- Thomas, A. K., Bulevich, J. B. & Dubois, S. J. (2012). An analysis of the determinants of the feeling of knowing. *Consciousness and Cognition: An International Journal*, *21*, 1681-1694.
- Toth, J. P., Daniels, K. A. & Solinger, L. A. (2011). What you know can hurt you: Effects of age and prior knowledge on the accuracy of judgments of learning. *Psychology and Aging*, *26*, 919-931.
- Touron, D. R. & Hertzog, C. (2004). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology and Aging*, *19*, 452-466.
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.
- Umanath, S. & Marsh, E. J. (2014). Understanding how prior knowledge influences memory in older adults. *Perspectives on Psychological Science*, *9*, 408-426.
- Underwood, B. J. (1966). Individual and group predictions of item difficulty for free learning. *Journal of Experimental Psychology*, *71*, 673-679.
- Undorf, M. & Erdfelder, E. (2011). Judgments of learning reflect encoding fluency: Conclusive evidence for the ease-of-processing hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1264-1269.
- Valentijn, S. A. M., Hill, R. D., van Hooren, S. A. H., Bosma, H., van Boxtel, M. P. J., Jolles, J. & Ponds, R. W. H. M. (2006). Memory self-efficacy predicts memory performance: Results from a 6-year follow-up study. *Psychology and Aging*, *21*, 165-172.
- van der Stel, M. & Veenman, M. V. (2014). Metacognitive skills and intellectual ability of young adolescents: A longitudinal study from a developmental perspective. *European Journal of Psychology of Education*, *29*, 117-137.
- van Loon, M. H., de Bruin, A. B. H., van Gog, T. & van Merriënboer, J. J. G. (2013). The effect of delayed-JOLs and sentence generation on children's monitoring accuracy and regulation of idiom study. *Metacognition and Learning*, *8*, 173-191.
- van Overschelde, J. P. & Nelson, T. O. (2006). Delayed judgments of learning cause both a decrease in absolute accuracy (calibration) and an increase in relative accuracy (resolution). *Memory & Cognition*, *34*, 1527-1538.
- Veenman, M. & Elshout, J. J. (1999). Changes in the relation between cognitive and metacognitive skills during the acquisition of expertise. *European Journal of Psychology of Education*, *14*, 509-523.
- Verhaeghen, P., Marcoen, A. & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology and Aging*, *7*, 242-251.
- Vise, M. & Schneider, W. (2000). Determinanten der Leistungsvorhersage bei Kindergarten- und Grundschulkindern: Zur Bedeutung metakognitiver und motivationaler Einflussfaktoren. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *32*, 51-58.

- Wahlheim, C. N., Dunlosky, J. & Jacoby, L. L. (2011). Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & Cognition*, *39*, 750-763.
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*, 158-177.
- Weaver, C. A., III (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 214-222.
- Weaver, C. A., III & Kelemen, W. L. (1997). Judgments of learning at delays: Shifts in response patterns or increased metamemory accuracy? *Psychological Science*, *8*, 318-321.
- Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G., Dolan, R. J. & Blakemore, S.-J. (2013). The development of metacognitive ability in adolescence. *Consciousness and Cognition: An International Journal*, *22*, 264-271.
- Wellman, H. M. (1977a). Preschoolers' understanding of memory-relevant variables. *Child Development*, *48*, 1720-1723.
- Wellmann, H. M. (1977b). Tip of the tongue and feeling of knowing experiences: a developmental study of memory monitoring. *Child Development*, *48*, 13-21.
- Wellman, H. M. (1978). Knowledge of the interaction of memory variables: A developmental study of metamemory. *Developmental Psychology*, *14*, 24-29.
- Wellman, H. M., Collins, J. & Gliberman, J. (1981). Understanding the combination of memory variables: Developing conceptions of memory limitations. *Child Development*, *52*, 1313-1317.
- Wellman, H. M. & Johnson, C. N. (1979). Understanding of mental processes: A developmental study of 'remember' and 'forget'. *Child Development*, *50*, 79-88.
- Wiley, J., Griffin, T. D. & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology*, *132*, 408-428.
- Wong, J. T., Cramer, S. J. & Gallo, D. A. (2012). Age-related reduction of the confidence-accuracy relationship in episodic memory: Effects of recollection quality and retrieval monitoring. *Psychology and Aging*, *27*, 1053-1065.
- Woo, E., Schmitter-Edgecombe, M. & Fancher, J. B. (2008). Memory prediction accuracy in younger and older adults: A cross-sectional and longitudinal analysis. *Aging, Neuropsychology, and Cognition*, *15*, 1-27.
- Worden, P. E. & Sladewski-Awig, L. J. (1982). Children's awareness of memorability. *Journal of Educational Psychology*, *74*, 341-350.
- Yussen, S. R. (1985). The role of metacognition in contemporary theories of cognitive development. In F. Pressley & D. Lynn (Eds.), *Metacognition, cognition and human performance* (Vol. 1, pp. 253-283). Orlando, FL: Academic Press.
- Yussen, S. R. & Levy, V. M. (1975). Developmental changes in predicting one's own span of short-term memory. *Journal of Experimental Child Psychology*, *19*, 502-508.
- Zabucky, K. M. & Ratner, H. H. (1986). Children's comprehension monitoring and recall of inconsistent stories. *Child Development*, *57*, 1401-1418.

- Zacks, R. T. (1969). Invariance of total learning time under different conditions of practice. *Journal of Experimental Psychology*, 82, 441-447.
- Zhao, Q. & Linderholm, T. (2008). Adult metacomprehension: Judgment processes and accuracy constraints. *Educational Psychology Review*, 20, 191-206.

ANHANG

Anhang A

A 1 Studie 1: Bilderliste

A 2 Studie 1: Gedächtnisfragen zum Film

A 3 Studie 2: Wortpaare

A 4 Studie 3: Fußballtexte

A 5 Studie 3: Fragen zum Text

A 6 Studie 3: Klassifikation von Sätzen

A 7 Studie 4: Wortpaare

A 8 Studie 5: Wortpaare

A 9 Studie 6: Wortpaare

Anhang B

B 1 Studie 1: Instruktion

B 2 Fragebogen zu demografischen Angaben

Anhang C








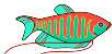





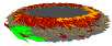














C 1 Studien 2 und 3: Fußball-Wissenstest





















C 2 Studie 4: Strategietraining

C 3 Studie 4: Fragebogen zum Strategietraining

C 4 Studie 6: Fragebogen zum **Lernanreiz**

A 1 Studie 1: Bilderliste

Nr.	Bild 1	Bild 2	Bezeichnung Bild 1	Bezeichnung Bild 2
Übung			Tennisschläger	Tennisball
Übung			Jacke	Hose
Übung			Haus	Luftballons
Übung			Auto	Fisch
1			Katze	Hund
2			Fuß	Hand
3			Vogel	Nest
4			Block	Stift
5			Socke	Schuh
6			Henne	Eier
7			Taschenlampe	Sonne
8			Geige	Trompete
9			Zitrone	Pflaume
10			Adventskranz	Schlitten

11			Schmetterling	Bett
12			Eis	Fahrrad
13			Kirche	Brille
14			Lineal	Elefant
15			Klavier	Banane
16			Schiff	Hut
17			Stuhl	Uhr
18			Giraffe	Eimer
19			Wolke	Schrank
20			Kuh	Schere

A 2 Studie 1: Gedächtnisfragen zum Film**Grundschüler**

1. Welche Farbe hat der Hase?
2. Welche zwei Dinge braucht jeder Zauberer zum Zaubern außer einem Buch?
3. Wo holt der Zauberer das Zauberbuch hervor?
4. Was zaubert der Zauberer als erstes in sein Zauberbuch?
5. Woher holt der Zauberer die Farbe Rot wieder in sein Buch?
6. Was hat der Zauberer die Zuschauer vor der Vorstellung gefragt?
7. Was fängt der Zauberer aus der Luft?
8. Welche Farbe hat der Schirm des Zauberers?
9. Wo hat der Zauberer die Verpackung für den Schirm gekauft?
10. Worüber ist der Zauberer ganz am Ende der Vorführung froh?
11. Wie viele Kunststücke führt der Zauberer insgesamt vor?
12. Wie heißt der Zauberer?

Jugendliche

1. Welche Farbe hat der Hase?
2. Was steht zu Beginn der Vorführung auf dem mittleren Tisch?
3. Welche zwei Dinge braucht jeder Zauberer zum Zaubern außer einem Buch?
4. Wo holt der Zauberer das Zauberbuch hervor?
5. Wo hat der Zauberer den kleinen schwarzen Beutel gefunden?
6. Was hat der Zauberer die Zuschauer vor der Vorstellung gefragt?
7. Warum ist es für den Zauberer schwierig Eier aus der Luft zu fangen?
8. Welche Farbe hat der Schirm des Zauberers?
9. Wo genau, sagt der Zauberer, müsse er manchmal auftreten, wo er nass werden könnte?
10. Welche Farbe hat der Spendenbeutel innen?
11. Wie viele Kunststücke führt der Zauberer insgesamt vor?
12. Wie heißt der Zauberer?

Jüngere Erwachsene

1. Was steht zu Beginn der Vorführung auf dem mittleren Tisch?
2. Welche zwei Dinge braucht jeder Zauberer zum Zaubern außer einem Buch?
3. Wo holt der Zauberer das Zauberbuch hervor?
4. Woher holt der Zauberer die Farbe Rot wieder in sein Buch?
5. Was passiert am Ende mit den Tüchern?
6. Wo hat der Zauberer den kleinen schwarzen Beutel gefunden?
7. Was hat der Zauberer die Zuschauer vor der Vorstellung gefragt?
8. Warum ist es für den Zauberer schwierig Eier aus der Luft zu fangen?
9. Wo hat der Zauberer die Verpackung für den Schirm gekauft?
10. Worüber ist der Zauberer am Ende der Vorstellung froh?
11. Welche Farbe hat der Tisch, auf dem der Hase steht?

12. Wie heißt der Zauberer?

Ältere Erwachsene

1. Welche Farbe hat der Hase?
2. Was steht zu Beginn der Vorführung auf dem mittleren Tisch?
3. Wo holt der Zauberer das Zauberbuch hervor?
4. Was zaubert der Zauberer als erstes in sein Zauberbuch?
5. Woher holt der Zauberer die Farbe Rot wieder in sein Buch?
6. Wo hat der Zauberer den kleinen schwarzen Beutel gefunden?
7. Was fängt der Zauberer aus der Luft?
8. Wo hat der Zauberer die Verpackung für den Schirm gekauft?
9. Welche Farbe hat der Spendenbeutel innen?
10. Was passiert am Ende mit den Tüchern?
11. Worüber ist der Zauberer ganz am Ende der Vorführung froh?
12. Wie viele Kunststücke führt der Zauberer insgesamt vor?

A 3 Studie 2: Wortpaare

Haus – Garten (Übungspaar)

Kind – Trompete (Übungspaar)

1. Trikot - Schuhe
2. Trainer – Kapitän
3. Klasse – Liga
4. Fan – Zuschauer
5. Feld – Stadion
6. Freistoß – Mauer
7. Abseits – Stürmer
8. Torwart – Parade
9. Sechser – Mittelfeld
10. Tunnel – Zweikampf
11. Grätsche – Gelbe Karte
12. Schiedsrichter – Flanke
13. Einwurf – Notbremse
14. Eckfahne – Libero
15. Sieger – Pass
16. Pfosten – Verlängerung
17. Spielbeginn – Kopfball
18. Hand – Rasen
19. Strafraum – Verein
20. Abstieg – Taktik
21. Pokal – Anpfiff
22. Spieler – Linie

A 4 Studie 3: Fußballtexte

Grundschüler

Max hat heute Geburtstag. Er wird 11 Jahre alt. Max interessiert sich sehr für Fußball und spielt auch selbst in der D-Jugend. Ausgerechnet heute, an seinem Geburtstag, findet das K.O.-Duell seiner Lieblingsmannschaft Borussia Dortmund gegen Bayern München statt. Max ist mit seinem Vater nach München gefahren, um sich das Spiel anzuschauen. Die Eintrittskarte ist nämlich das Geburtstagsgeschenk von seinem Vater. Er freut sich sehr auf dieses Spiel in der Champions League und hofft, dass seine Mannschaft es schaffen und weiter kommen wird.

Gleich geht es los, die Mannschaften sind schon im Spielertunnel. Im Stadion ist es höllisch laut, die Fans feuern schon vor dem Spiel ihre Teams unermüdlich an. Bayerns Kapitän gewinnt den Münzwurf. Der Schiedsrichter pfeift die 1. Halbzeit an. Dortmund stößt an und versucht sofort mit einem langen Ball nach vorne einen ersten Angriff aufzubauen, Bayern geht jedoch dazwischen und hat den Ball. Max ist sehr nervös, da es heute um sehr viel für seine Mannschaft Borussia Dortmund geht. Auch die Spieler sind nicht frei im Kopf. Es gibt viele Fehlpässe und Ballverluste zu sehen. Nun gibt es den ersten Vorstoß der Gastgeber in den Strafraum der Dortmunder. Dortmunds Torwart hält bei der harmlosen Flanke von rechts sicher und sucht sofort einen Mitspieler, den er anspielen kann. Bisher gab es deutlich mehr Ballbesitz für Dortmund. Bayern wartet dagegen ab und schaut sich an, was die Gäste zu bieten haben. Der Dortmunder Stürmer steht als einziger Feldspieler in der Hälfte der bayerischen Mannschaft und kann deshalb in Ruhe den Ball annehmen und unbehindert auf das Tor zulaufen. Der Linienrichter hebt seine Fahne. Es gibt einen Freistoß für Bayern. Dortmunds Hintermannschaft scheint zu schlafen; sie sichert den Ball nicht und macht so einem Gegenspieler die Bahn zum Tor frei. Dieser geht durch und lässt dem Torwart keine Chance. Max ist verärgert über das Abwehrverhalten seiner Mannschaft. Die Dortmunder bemühen sich sehr, doch der Erfolg bleibt aus. Ein Dortmunder Spieler ist angeschlagen liegen geblieben, er hat einen Tritt auf das Bein abbekommen. Es gibt eine Behandlungspause. Er kann aber schnell wieder weiter spielen. Der Verteidiger von Bayern schießt scharf auf das Tor, der Torwart lässt den Ball zur Seite prallen. Ein Spieler von Bayern reagiert am schnellsten und schießt wieder auf das Tor. Der Torwart streckt sich vergeblich. Jetzt sieht es aber ganz schlecht aus für Dortmund. Max ist schockiert. So hatte er sich das Spiel nicht vorgestellt! Jetzt heißt es weitermachen, die Mannschaft anfeuern und vielleicht geht da ja doch noch was! Doch es greifen weiterhin nur die Gastgeber an.

Abpfiff 1. Hälfte. In der Halbzeit holt Max sich jetzt erst einmal etwas zu Essen. Sein Vater spendiert ihm eine Bratwurst. Danach essen sie noch von dem Geburtstagskuchen, den sie von zu Hause mitgebracht haben.

Anpfiff 2. Hälfte. Gerade rechtzeitig zum Wiederanpfiff kommt auch Max auf seinen Platz zurück. Bayern stößt an. Ein Dortmunder bekommt den Ball, dribbelt an zwei Bayern vorbei und schießt dann aus sehr spitzem Winkel an das Außennetz. Immerhin ein Torschuss! Dann gibt es eine rote Karte für Bayern! Nach einem Foul eines Dortmunders an der Außenlinie springt sein gefoulter Gegner auf und versetzt dem Dortmunder einen Schlag vor die Brust –

und das vor den Augen des Linienrichters. Der Schiedsrichter hat keine Wahl: Platzverweis. Das bedeutet Überzahl für die Gäste und neue Hoffnung! Auch Max freut sich und schwenkt seinen schwarz-gelben Schal. Spielerwechsel bei Dortmund: Ein frischer Stürmer kommt und soll noch einmal für etwas Leben im Angriff sorgen. Doch die Bayern spielen die Zeit nun runter. Das Spiel ist schon längst entschieden! Der Schlusspfiff ertönt, das Spiel ist aus. Max hat nun noch die Heimreise vor sich. Mit dem Zug geht es zurück nach Hause! Er wird daheim viel zu erzählen haben – denn so einen aufregenden Geburtstag hatte er noch nie!

Jugendliche und ältere Erwachsene

Max hat heute Geburtstag. Er wird 11 Jahre alt. Max interessiert sich sehr für Fußball und spielt auch selbst in der D-Jugend. Ausgerechnet heute, an seinem Geburtstag, findet das K.O.-Duell seiner Lieblingsmannschaft Borussia Dortmund gegen Bayern München statt. Max ist mit seinem Vater nach München gefahren, um sich das Spiel anzuschauen. Die Eintrittskarte ist nämlich das Geburtstagsgeschenk von seinem Vater. Er freut sich sehr auf dieses Spiel in der Champions League und hofft, dass seine Mannschaft es schaffen und weiter kommen wird.

Gleich geht es los, die Mannschaften sind schon im Spielertunnel. Im Stadion ist es höllisch laut, die Fans feuern schon vor dem Spiel ihre Teams unermüdlich an. Bayerns Kapitän gewinnt den Münzwurf. Der Schiedsrichter pfeift die 1. Halbzeit an. Dortmund stößt an und versucht sofort mit einem langen Ball auf den Flügel einen ersten Angriff aufzubauen, Bayern geht jedoch dazwischen und hat den Ball. Max ist sehr nervös, da es heute um sehr viel für seine Mannschaft geht. Auch die Spieler sind nicht frei im Kopf. Es gibt viele Fehlpässe und Ballverluste zu sehen. Max gefällt die Leistung der Dortmunder nicht, aber er hofft, dass sie sich noch verbessern. Nun gibt es den ersten Vorstoß der Gastgeber in den Strafraum der Dortmunder. Dortmunds Torwart hält bei der harmlosen Flanke von rechts sicher und sucht sofort einen Mitspieler, den er anspielen kann. Bislang gab es deutlich mehr Ballbesitz für Dortmund. Bayern wartet dagegen ab und schaut sich an, was die Gäste zu bieten haben. Der Dortmunder Stürmer ist schon zu weit nach vorn gesprintet, ehe dann sein Mitspieler zu ihm spielt, deshalb hebt der Linienrichter seine Fahne. Es gibt einen Freistoß für Bayern. Dortmunds Hintermannschaft scheint zu schlafen; sie sichert den Ball nicht und macht so einem Gegenspieler die Bahn frei. Dieser geht durch und lässt dem Torwart keine Chance. Max ist verärgert über das Abwehrverhalten seiner Mannschaft. In der Mitte der ersten Halbzeit mühen sich die Dortmunder, doch der Ertrag bleibt aus. Sie versuchen es mit Schüssen aus der Distanz, doch diese gehen meist weit vorbei, wie auch dieser. Es gibt mal wieder einen Abstoß für Bayern. Ein Dortmunder Spieler ist angeschlagen liegen geblieben, er hat einen Tritt auf das Bein abbekommen. Es gibt eine Behandlungspause. Er kann aber schnell wieder weitermachen. Und erneut greifen die Gastgeber an. Der aufgerückte Außenverteidiger Bayerns schießt scharf auf das Tor, der Torwart lässt den Ball zur Seite prallen und dort reagiert erneut ein Bayer am schnellsten. Der Torwart streckt sich vergeblich. Jetzt sieht es aber ganz schlecht aus für Dortmund. Max ist schockiert. So hatte er sich das Spiel nicht vorgestellt! Jetzt heißt es weitermachen, die Mannschaft anfeuern und vielleicht geht da ja noch was! Ein Bayer schießt aus der Distanz, Dortmunds Torhüter kann den Ball erneut nicht festhalten, diesmal rettet aber ein Verteidiger zur Ecke.

Abpfiff 1. Hälfte. In der Halbzeit holt Max sich jetzt erst einmal etwas zu Essen. Sein Vater spendiert ihm eine Bratwurst. Danach essen sie noch von dem Geburtstagskuchen, den sie von zu Hause mitgebracht haben.

Anpfiff 2. Hälfte. Bayern stößt an. Und es geht weiter, wie die erste Hälfte aufgehört hat: Eine Flanke von Bayern - und der Torhüter der Dortmunder fängt den Ball sicher ab. Auf einmal sieht Max, dass in der Südkurve Rauch aufsteigt. Fans der Bayern müssen ein bengalisches Feuer gezündet haben. Doch bevor der Nebel das Spiel stören kann, hat auch der Schiedsrichter das Feuer bemerkt. Er pfeift das Spiel ab. Nach zehn Minuten haben die Stadionordner die Lage im Griff: Das Feuer wurde gelöscht, die verantwortlichen Fans haben einen Stadionverweis bekommen und das Spiel kann weiter gehen. Nun gibt es doch einen Spielerwechsel bei Dortmund. Der neue Mann wirbelt an der Grundlinie, dribbelt an zwei Bayern vorbei und schießt dann aus sehr spitzem Winkel an das Außennetz. Max und die anderen Fans freuen sich und sofort werden die Anfeuerungen wieder lauter. Der Schuss bleibt jedoch lange das einzige Lebenszeichen. Inzwischen ist die Enttäuschung bei Max schon recht groß. Dann gibt es Rot für Bayern! Nach einem Foul eines Dortmunders an der Außenlinie springt sein gefoulter Gegner auf und versetzt dem Gästespieler einen Schlag vor die Brust – und das vor den Augen des Linienrichters. Der Schiedsrichter hat keine Wahl: Platzverweis. Das bedeutet Überzahl für die Gäste und noch einmal neue Hoffnung! Auch Max freut sich und schwenkt seinen schwarz-gelben Schal. Es bleibt nur noch eine knappe halbe Stunde Zeit für Dortmund. Spielerwechsel bei Dortmund: Ein frischer Stürmer kommt und soll noch einmal für etwas Leben im Angriff sorgen. Doch die Bayern spielen die Zeit nun runter. Der Schlusspfiff ertönt, das Spiel ist aus. Max hat nun noch die Heimreise vor sich. Mit dem Zug geht es zurück nach Hause! Er wird daheim viel zu erzählen haben – denn so einen aufregenden Geburtstag hatte er noch nie!

Jüngere Erwachsene

Max hat heute Geburtstag. Er wird 11 Jahre alt. Max interessiert sich sehr für Fußball und spielt auch selbst in der D-Jugend. Ausgerechnet heute, an seinem Geburtstag, findet das K.O.-Duell seiner Lieblingsmannschaft Borussia Dortmund gegen Bayern München statt. Max ist mit seinem Vater nach München gefahren, um sich das Spiel anzuschauen. Die Eintrittskarte ist nämlich das Geburtstagsgeschenk von seinem Vater. Er freut sich sehr auf dieses Spiel in der Champions League und hofft, dass seine Mannschaft es schaffen und weiter kommen wird.

Gleich geht es los, die Mannschaften sind schon im Spielertunnel. Im Stadion ist es höllisch laut, die Fans feuern schon vor dem Spiel ihre Teams unermüdlich an. Bayerns Kapitän gewinnt den Münzwurf. Der Schiedsrichter pfeift die 1. Halbzeit an. Dortmund stößt an und versucht sofort mit einem langen Ball auf den Flügel einen ersten Angriff aufzubauen, Bayern klärt jedoch souverän und hat den Ball. Max ist sehr nervös, da es heute um sehr viel für seine Mannschaft geht. Auch die Spieler sind nicht frei im Kopf. Es gibt viele Fehlpässe und Ballverluste zu sehen. Max gefällt die Leistung der Dortmunder nicht, aber er hofft, dass sie sich noch verbessern. Nun gibt es den ersten Vorstoß der Gastgeber in den Strafraum der Dortmunder. Dortmunds Torwart packt bei der harmlosen Flanke von rechts sicher zu und sucht sofort einen Mitspieler, den er anspielen kann. Bislang gab es deutlich mehr Ballbesitz für

Dortmund. Bayern wartet dagegen ab und schaut sich an, was die Gäste zu bieten haben. Der Dortmunder Stürmer ist schon zu weit nach vorn gesprintet, ehe dann sein Mitspieler zu ihm spielt, deshalb hebt der Linienrichter seine Fahne. Es gibt einen indirekten Freistoß für Bayern. Dortmunds Hintermannschaft scheint zu schlafen; sie sichert den Ball nicht und macht so einem Gegenspieler die Bahn frei. Dieser geht durch und lässt dem Torwart keine Chance. Max ist verärgert über das Abwehrverhalten seiner Mannschaft. Wenig später gibt es wieder eine kleine Unsicherheit in der Ballbehandlung bei Dortmunds Verteidigern. Die Gastgeber kommen zu einer weiteren guten Tormöglichkeit. Ein Dortmunder Verteidiger holt seinen Gegenspieler an der Seitenauslinie von den Beinen und sieht dafür gelb. Noch einmal darf er sich so etwas nicht erlauben!

Der Torhüter der Dortmunder ist heute viel beschäftigt. Schon wieder muss er ran und entschärft einen Distanzschuss zur Ecke. Ein bayerischer Spieler bringt den Ball mit Schnitt herein, ein Verteidiger schlägt die Kugel hoch und weit aus dem Sechszehner. Max beginnt so langsam, den Glauben an einen Sieg zu verlieren. In der Mitte der ersten Halbzeit mühen sich die Dortmunder, doch der Ertrag bleibt aus. Sie versuchen es mit Schüssen aus der Distanz, doch diese gehen meist weit vorbei, wie auch dieser. Es gibt mal wieder einen Abstoß für Bayern. Ein Dortmunder Spieler ist angeschlagen liegen geblieben, er hat einen Tritt auf das Bein abbekommen. Es gibt eine Behandlungspause. Er kann aber schnell wieder weitermachen. Und erneut greifen die Gastgeber an. Der aufgerückte Außenverteidiger Bayerns schießt scharf auf das Tor, der Torwart lässt den Ball zur Seite prallen und dort reagiert erneut ein Bayer am schnellsten. Der Torwart streckt sich vergeblich. Doch kurz nach dem Anstoß gibt es ein Foul der Bayern – Freistoß für Dortmund in halblinker Position etwa 35 Meter vor dem Tor – jetzt ein guter Schuss und alles ist wieder möglich! Ein Dortmunder Angreifer schaut sich aus dem Halbfeld nach Abnehmern seiner Flanke vor dem Tor um. Doch ein Griff ans Trikot des Gegenspielers in der Mitte sorgt für einen Freistoß für Bayern. Ein Bayer schießt aus der Distanz, Dortmunds Torhüter kann den Ball erneut nicht festhalten, diesmal rettet aber ein Verteidiger zur Ecke.

Abpfiff 1. Hälfte. In der Halbzeit holt Max sich jetzt erst einmal etwas zu Essen. Sein Vater spendiert ihm eine Bratwurst. Danach essen sie noch von dem Geburtstagskuchen, den sie von zu Hause mitgebracht haben.

Anpfiff 2. Hälfte. Ohne Spielerwechsel bei beiden Mannschaften geht es weiter. Dortmund hofft nun auf ein Wunder, allen voran die Fans auf der Tribüne. Gerade rechtzeitig zum Wiederanpfiff kommt auch Max auf seinen Platz zurück. Bayern stößt an. Und es geht weiter, wie die erste Hälfte aufgehört hat: Eine Flanke von Bayern - und der Torhüter der Dortmunder fängt den Ball sicher ab. Auf einmal sieht Max, dass in der Südkurve Rauch aufsteigt. Fans der Bayern müssen ein bengalisches Feuer gezündet haben. Doch bevor der Nebel das Spiel stören kann, hat auch der Schiedsrichter das Feuer bemerkt. Er pfeift das Spiel ab. Nach zehn Minuten haben die Stadionordner die Lage im Griff: Das Feuer wurde gelöscht, die verantwortlichen Fans haben einen Stadionverweis bekommen und das Spiel kann weiter gehen. Nun gibt es doch einen Spielerwechsel bei Dortmund. Der neue Mann wirbelt an der Grundlinie, dribbelt an zwei Bayern vorbei und schießt dann aus sehr spitzem Winkel an das Außennetz. Max und die anderen Fans freuen sich und sofort werden die Anfeuerungen wieder lauter. Der Schuss bleibt jedoch lange das einzige Lebenszeichen. Inzwischen ist die Enttäuschung bei Max schon

recht groß. Dann gibt es Rot für Bayern! Nach einem Foul eines Dortmunders an der Außenlinie springt sein gefoulter Gegner auf und versetzt dem Gästespieler einen Schlag vor die Brust – und das vor den Augen des Linienrichters. Der Schiedsrichter hat keine Wahl: Platzverweis. Das bedeutet Überzahl für die Gäste und noch einmal neue Hoffnung! Auch Max freut sich und schwenkt seinen schwarz-gelben Schal. Es bleibt nur noch eine knappe halbe Stunde Zeit für Dortmund. Wieder einmal greifen die Gäste an: Auf der Außenbahn umkurvt der eingewechselte Angreifer seinen Gegenspieler jetzt zum zweiten Mal, spielt flach nach innen, sucht seinen Mitspieler. Der ist jedoch zugestellt, aber es gibt zumindest eine Ecke. Spielerwechsel bei Dortmund: Ein frischer Stürmer kommt und soll noch einmal für etwas Leben im Angriff sorgen. Doch die Bayern spielen die Zeit nun runter. Der Schlusspfiff ertönt, das Spiel ist aus. Max hat nun noch die Heimreise vor sich. Mit dem Zug geht es zurück nach Hause! Er wird daheim viel zu erzählen haben – denn so einen aufregenden Geburtstag hatte er noch nie!

A 5 Studie 3: Fragen zum Text**Grundschüler**

1. Welche Mannschaft führt jeweils zu Beginn der Halbzeit den Anstoß aus?
2. Wie viele Tore sind in dem Spiel gefallen?
3. Wie alt ist Max?
4. Welche Mannschaft feuert Max an?
5. Wie viele rote Karten gab es?
6. In welcher Stadt findet das Spiel statt?
7. Welche Mannschaft sind die Gäste und welche Mannschaft sind die Gastgeber?
8. Was isst Max in der Halbzeitpause?
9. Welche Mannschaft gewinnt den Münzwurf?
10. In welchem Wettbewerb findet das Spiel statt?
11. Wie verändert sich der Ballbesitz im Spielverlauf?
12. Auf welcher Position spielt der eingewechselte Spieler bei Dortmund?
13. Warum muss ein Spieler medizinisch behandelt werden?
14. Was für Auswirkungen hat eine Niederlage in diesem Spiel für die Mannschaft, die verliert?
15. Wo befindet sich die Mannschaft, bevor das Spiel losgeht?
16. Was passiert, nachdem der Linienrichter die Fahne erhoben hat?
17. In welcher Fußballjugend spielt Max?
18. Welche Farbe hat Max' Schal?

Jugendliche und ältere Erwachsene

1. Wie verändert sich der Ballbesitz im Spielverlauf?
2. Welche Mannschaft führt jeweils zu Beginn der Halbzeiten den Anstoß aus?
3. Welche Auswirkung hat ein Sieg in diesem Spiel für die Siegermannschaft?
4. Wie viele Tore fallen in diesem Spiel?
5. Wie oft wird der Anstoß ausgeführt?
6. In welcher Stadt findet das Spiel statt?
7. Wie viele Torchancen hat Borussia Dortmund?
8. Warum erhält Bayern in der zweiten Halbzeit eine rote Karte?
9. Wie viele Spieler spielen nur einen Teil des Spiels mit?
10. Aus welcher Mannschaft stammt der Spieler, der im Abseits steht?
11. Welche Mannschaft gewinnt den Münzwurf?
12. In welchem Wettbewerb findet das Spiel statt?
13. Warum muss ein Spieler medizinisch behandelt werden?
14. Auf welcher Position spielt der eingewechselte Spieler bei Dortmund?
15. Woran kann man erkennen, dass die Dortmunder Spieler nervös sind?
16. Warum erhält Dortmund in der ersten Halbzeit eine gelbe Karte?
17. In welcher Fußballjugend spielt Max?
18. Welche Farbe hat Max' Schal?

Jüngere Erwachsene

1. In welchem Wettbewerb findet das Spiel statt?
2. Welche Mannschaft gewinnt den Münzwurf?
3. Welche Mannschaft führt jeweils zu Beginn der Halbzeiten den Anstoß aus?
4. Wie fällt das erste Tor?
5. Wie verändert sich der Ballbesitz im Spielverlauf?
6. Welche Auswirkung hat ein Sieg in diesem Spiel für die Siegermannschaft?
7. Wie viele Tore fallen in dem Spiel?
8. Wie oft wird der Anstoß ausgeführt?
9. In welcher Stadt findet das Spiel statt?
10. Wie viele Torchancen hat Borussia Dortmund?
11. Wie foult Dortmund kurz vor der Halbzeit?
12. Warum erhält Bayern in der zweiten Halbzeit eine rote Karte?
13. Wie viele Spieler spielen nur einen Teil des Spiels mit?
14. Aus welcher Mannschaft stammt der Spieler, der im Abseits steht?
15. Warum muss ein Spieler medizinisch behandelt werden?
16. Auf welcher Position spielt der eingewechselte Spieler bei Dortmund?
17. Woran kann man merken, dass die Dortmunder Spieler nervös sind?
18. Wie viele Torchancen hat Dortmund?
19. Warum erhält Dortmund in der ersten Halbzeit eine gelbe Karte?
20. In welcher Fußballjugend spielt Max?

A 6 Studie 3: Klassifikation von Sätzen

Grundschüler

1. Bayerns Kapitän gewinnt den Münzwurf.
2. Ein Dortmunder Spieler ist angeschlagen liegengeblieben, er ist von seinem Gegenspieler gefoult worden.
3. Gleich geht es los, die Mannschaften sind schon im Stadion.
4. Dortmunds Torwart hält bei der harmlosen Flanke von rechts sicher und sucht sofort einen Mitspieler, den er anspielen kann.
5. Ein Spieler von Bayern reagiert am schnellsten und schießt wieder auf das Tor.
6. Bislang gab es deutlich mehr Ballbesitz für Bayern.
7. Der Linienrichter hat keine Wahl: Platzverweis.
8. Auch die Spieler sind nicht frei im Kopf.
9. Max ist sehr nervös, da es heute um sehr viel für seine Mannschaft Borussia Dortmund geht.
10. Dortmund stößt an und versucht sofort mit einem langen Ball nach vorne einen ersten Angriff aufzubauen.
11. Der Linienrichter hebt seine Fahne, um das Abseits anzuzeigen.
12. Doch es greifen weiterhin nur die Gäste an.
13. Max ist verärgert über das Abwehrverhalten seiner Mannschaft.
14. Dieser geht durch und lässt dem Torwart keine Chance.
15. Nach einem Foul eines Dortmunders an der Außenlinie springt sein gefoulter Gegner auf und versetzt dem Dortmunder einen Schlag vor die Brust – und das vor den Augen des Linienrichters.
16. Der Torwart streckt sich nach dem Ball.
17. Ein Dortmunder bekommt den Ball, dribbelt an zwei Bayern vorbei und schießt dann aus sehr spitzem Winkel ins Tor.
18. Der Verteidiger von Bayern schießt scharf auf das Tor, der Torwart lässt den Ball zur Seite prallen.
19. Es gibt viele Fehlpässe zu sehen.
20. Max interessiert sich sehr für Fußball und spielt auch selbst in der C-Jugend.

Jugendliche und ältere Erwachsene

1. Und erneut greifen die Gäste an.
2. Dortmunds Torwart packt bei der harmlosen Flanke von rechts sicher zu und sucht sofort einen Mitspieler, den er anspielen kann.
3. Gleich geht es los, die Mannschaften sind schon im Stadion.
4. Ein frischer Stürmer kommt und soll noch einmal für etwas Leben im Angriff sorgen.
5. Ein Dortmunder Spieler ist angeschlagen liegen geblieben, er ist von seinem Gegenspieler gefoult worden.
6. Bislang gab es deutlich mehr Ballbesitz für Bayern.

7. Sie sichert den Ball nicht und macht so einem Gegenspieler die Bahn frei.
8. Nun gibt es den ersten Vorstoß der Gastgeber in den Strafraum der Dortmunder.
9. Sie versuchen es mit Schüssen aus der Distanz, doch diese gehen meist weit vorbei, wie auch dieser.
10. Der Schuss bleibt jedoch lange das einzige Lebenszeichen.
11. Nach einem Foul eines Dortmunders im Strafraum springt sein gefoulter Gegner auf und versetzt dem Gästespieler einen Schlag vor die Brust – und das vor den Augen des Linienrichters.
12. Ein Bayer schießt aus der Distanz, Dortmunds Torhüter kann den Ball erneut nicht festhalten; der Torwart streckt sich vergeblich.
13. In der Mitte der ersten Halbzeit mühen sich die Dortmunder, doch der Ertrag bleibt aus.
14. Der Linienrichter hebt seine Fahne, um das Abseits anzuzeigen.
15. Dortmund stößt an und versucht sofort mit einem langen Ball auf das Tor einen ersten Angriff aufzubauen.
16. Eine Flanke von Bayern - und der Torhüter der Dortmunder fängt den Ball sicher ab.
17. Der neue Mann wirbelt an der Grundlinie, dribbelt an zwei Bayern vorbei und schießt dann aus sehr spitzem Winkel an das Außennetz.
18. Max ist verärgert über den Torwart seiner Mannschaft.
19. Es gibt mal wieder einen Freistoß für Bayern.
20. Das bedeutet Überzahl für die Gäste und noch einmal neue Hoffnung!

Jüngere Erwachsene

1. Und erneut greifen die Gäste an.
2. Dortmunds Torwart packt bei der harmlosen Flanke von rechts sicher zu und sucht sofort einen Mitspieler, den er anspielen kann.
3. Ein Dortmunder Verteidiger holt seinen Gegenspieler an der Seitenauslinie von den Beinen und sieht dafür gelb.
4. Wenig später gibt es wieder eine kleine Unsicherheit in der Ballbehandlung bei Dortmunds Verteidigern.
5. Ein Dortmunder Spieler ist angeschlagen liegen geblieben, er ist von seinem Gegenspieler gefoult worden.
6. Bislang gab es deutlich mehr Ballbesitz für Bayern.
7. Ein bayerischer Spieler bringt den Ball mit Schnitt herein, ein Verteidiger schlägt die Kugel hoch und weit aus dem Sechszehner.
8. Nun gibt es den ersten Vorstoß der Gastgeber in den Strafraum der Dortmunder.
9. Ein Dortmunder Angreifer schaut sich nach Abnehmern seiner Flanke vor dem Tor um.
10. Der ist zugestellt, aber es gibt zumindest einen Freistoß.
11. Nach einem Foul eines Dortmunders im Strafraum springt sein gefoulter Gegner auf und versetzt dem Gästespieler einen Schlag vor die Brust – und das vor den Augen des Linienrichters.

12. Ein Bayer schießt aus der Distanz, Dortmunds Torhüter kann den Ball erneut nicht festhalten; der Torwart streckt sich vergeblich.
13. Doch ein Griff ans Trikot des Gegenspielers in der Mitte sorgt für einen Freistoß für Bayern.
14. Der Linienrichter hebt seine Fahne, um das Abseits anzuzeigen.
15. Dortmund stößt an und versucht sofort mit einem langen Ball auf das Tor einen ersten Angriff aufzubauen.
16. Eine Flanke von Bayern - und der Torhüter der Dortmunder fängt den Ball sicher ab.
17. Der neue Mann wirbelt an der Grundlinie, dribbelt an zwei Bayern vorbei und schießt dann aus sehr spitzem Winkel an das Außennetz.
18. In der Mitte der ersten Halbzeit mühen sich die Dortmunder, doch der Ertrag bleibt aus.
19. Max ist verärgert über den Torwart seiner Mannschaft.
20. Das bedeutet Überzahl für die Gäste und noch einmal neue Hoffnung!

A 7 Studie 4: Wortpaare**Grundschüler und ältere Erwachsene**

	Wort 1	Kategorie	Wort 2	Kategorie
Übungswortpaare				
1	Tür		Fenster	
2	Fußball		Krone	
3	Sonne		Mond	
4	Nadel		Buch	
Wortpaare in der Lernphase				
1	Giraffe	1	Nilpferd	1
2	Leopard	1	Nashorn	1
3	Fahrrad	2	Schiff	2
4	Zug	2	Traktor	2
5	Puppe	3	Rutsche	3
6	Teddy	3	Ball	3
7	Tisch	4	Sofa	4
8	Schrank	4	Sessel	4
9	Buche	5	Palme	5
10	Eiche	5	Pappel	5
11	Ahorn	5	Linde	5
12	Zwiebel	6	Kürbis	6
13	Kohlrabi	6	Gurke	6
14	Brokkoli	6	Spargel	6
15	Auge	7	Schulter	7
16	Rücken	7	Zahn	7
17	Klavier	8	Trompete	8
18	Posaune	8	Orgel	8
19	Socken	9	Krawatte	9
20	Pullover	9	Stiefel	9
21	Mantel	9	Kleid	9
22	Bohrer	10	Axt	10
23	Säge	10	Feile	10
24	Gepard	1	Knete	3
25	Elefant	1	Jacke	9
26	Taxi	2	Ohr	7
27	Kutsche	2	Hocker	4
28	Buntstift	3	Flöte	8
29	Bauklötze	3	Kohl	6
30	Teppich	4	Schraube	10
31	Gardine	4	Flugzeug	2

32	Fichte	5	Puzzle	3
33	Birke	5	Karotte	6
34	Tomate	6	Kamel	1
35	Sellerie	6	Motorrad	2
36	Haare	7	Bluse	9
37	Nase	7	Kastanie	5
38	Klarinette	8	Beil	10
39	Glockenspiel	8	Zebra	1
40	Rock	9	Schaukel	3
41	Sandalen	9	Harfe	8
42	Hammer	10	Stirn	7
43	Hacke	10	Stuhl	4
44	Bagger	2	Trommel	8
45	Gitarre	8	Eisbär	1
Neue Wortpaare				
1	Löwe	1	Gorilla	1
2	Hubschrauber	2	Auto	2
3	Kreisel	3	Kugelbahn	3
4	Regal	4	Lampe	4
5	Schenkel	7	Mund	7
6	Geige	8	Pauke	8
7	Nagel	10	Zange	10
8	Tiger	1	Dübel	10
9	Murmel	3	Erbse	6
10	Bett	4	Eberesche	5
11	Tanne	5	Hemd	9
12	Paprika	6	Hals	7
13	Knie	7	Herd	4
14	Anzug	9	Lastwagen	2
15	Schaufel	10	Rübe	6

Jugendliche

	Wort 1	Kategorie	Wort 2	Kategorie
Übungswortpaare				
1	Tür	0	Fenster	0
2	Fußball	0	Krone	0
3	Sonne	0	Mond	0
4	Nadel	0	Buch	0
Wortpaare in der Lernphase				
1	Giraffe	1	Nilpferd	1
2	Leopard	1	Delfin	1
3	Krokodil	1	Nashorn	1

4	Fahrrad	2	Bus	2
5	Zug	2	Traktor	2
6	Schiff	2	Straßenbahn	2
7	Puppe	3	Rutsche	3
8	Teddy	3	Ball	3
9	Tisch	4	Sofa	4
10	Schrank	4	Sessel	4
11	Buche	5	Palme	5
12	Eiche	5	Pappel	5
13	Ahorn	5	Linde	5
14	Zwiebel	6	Kürbis	6
15	Kohlrabi	6	Gurke	6
16	Brokkoli	6	Spargel	6
17	Auge	7	Schulter	7
18	Rücken	7	Zahn	7
19	Klavier	8	Trompete	8
20	Posaune	8	Orgel	8
21	Socken	9	Krawatte	9
22	Pullover	9	Stiefel	9
23	Mantel	9	Kleid	9
24	Bohrer	10	Axt	10
25	Säge	10	Feile	10
26	Marzipan	11	Brause	11
27	Lutscher	11	Kaugummi	11
28	Gepard	1	Knete	3
29	Elefant	1	Jacke	9
30	Taxi	2	Ohr	7
31	Kutsche	2	Hocker	4
32	Buntstift	3	Flöte	8
33	Bauklötze	3	Kohl	6
34	Teppich	4	Schraube	10
35	Gardine	4	Flugzeug	2
36	Fichte	5	Bonbon	11
37	Birke	5	Karotte	6
38	Sellerie	6	Motorrad	2
39	Haare	7	Bluse	9
40	Nase	7	Kastanie	5
41	Klarinette	8	Eis	11
42	Glockenspiel	8	Zebra	1
43	Rock	9	Schaukel	3
44	Mütze	9	Harfe	9
45	Hammer	10	Lakritz	11
46	Hacke	10	Gitarre	8

47	Bagger	2	Tomate	6
48	Känguru	1	Trommel	8
49	Schokolade	11	Stirn	7
50	Beil	10	Kamel	1
51	Gummibärchen	11	Puzzle	3
52	Bratsche	8	Stuhl	4
53	Bauch	7	Ulme	5
54	Meißel	10	Kuchen	11
Neue Wortpaare				
1	Löwe	1	Gorilla	1
2	Hubschrauber	2	Auto	2
3	Kreisel	3	Kugelbahn	3
4	Regal	4	Lampe	4
5	Schenkel	7	Mund	7
6	Geige	8	Pauke	8
7	Nagel	10	Zange	10
8	Keks	11	Praline	11
9	Gürtel	9	Sandalen	9
10	Tiger	1	Dübel	10
11	Murmel	3	Erbse	6
12	Bett	4	Eberesche	5
13	Tanne	5	Hemd	9
14	Paprika	6	Hals	7
15	Knie	7	Herd	4
16	Anzug	9	Lastwagen	2
17	Schaufel	10	Rübe	6
18	Schuhe	9	Rettich	6

Jüngere Erwachsene

	Wort 1	Kategorie	Wort 2	Kategorie
Übungswortpaare				
1	Tür	0	Fenster	0
2	Fußball	0	Krone	0
3	Sonne	0	Mond	0
4	Nadel	0	Buch	0
Wortpaare in der Lernphase				
1	Giraffe	1	Nilpferd	1
2	Leopard	1	Nashorn	1
3	Fahrrad	2	Bus	2
4	Zug	2	Traktor	2
5	Schiff	2	Straßenbahn	2
6	Puppe	3	Rutsche	3

7	Teddy	3	Ball	3
8	Tisch	4	Sofa	4
9	Schrank	4	Sessel	4
10	Buche	5	Palme	5
11	Eiche	5	Pappel	5
12	Ahorn	5	Linde	5
13	Zwiebel	6	Kürbis	6
14	Kohlrabi	6	Gurke	6
15	Brokkoli	6	Spargel	6
16	Auge	7	Schulter	7
17	Rücken	7	Zahn	7
18	Klavier	8	Trompete	8
19	Posaune	8	Orgel	8
20	Socken	9	Krawatte	9
21	Pullover	9	Stiefel	9
22	Mantel	9	Kleid	9
23	Bohrer	10	Axt	10
24	Säge	10	Feile	10
25	Schloss	11	Hochhaus	11
26	Theater	11	Bauernhof	11
27	Ordner	12	Radierer	12
28	Tacker	12	Spitzer	12
29	Marzipan	13	Brause	13
30	Lutscher	13	Kaugummi	13
31	Gepard	1	Knete	3
32	Elefant	1	Jacke	9
33	Taxi	2	Ohr	7
34	Kutsche	2	Hocker	4
35	Buntstift	3	Flöte	8
36	Bauklötze	3	Kohl	6
37	Teppich	4	Schraube	10
38	Gardine	4	Flugzeug	2
39	Fichte	5	Bonbon	13
40	Birke	5	Karotte	6
41	Tomate	6	Locher	12
42	Sellerie	6	Motorrad	2
43	Haare	7	Bluse	9
44	Nase	7	Kastanie	5
45	Klarinette	8	Burg	11
46	Glockenspiel	8	Zebra	1
47	Rock	9	Schaukel	3
48	Sandalen	9	Harfe	8
49	Hammer	10	Lakritz	13

50	Hacke	10	Stuhl	4
51	Bagger	2	Papier	12
52	Gitarre	8	Hotel	11
53	Kirche	11	Känguru	1
54	Kleber	12	Trommel	8
55	Schokolade	13	Stirn	7
56	Rathaus	11	Beil	10
57	Akte	12	Kamel	1
58	Gummibärchen	13	Puzzle	3
59	Museum	11	Bratsche	8
60	Bauch	7	Ulme	5
Neue Wortpaare				
1	Löwe	1	Gorilla	1
2	Hubschrauber	2	Auto	2
3	Kreisel	3	Kugelbahn	3
4	Regal	4	Lampe	4
5	Schenkel	7	Mund	7
6	Geige	8	Pauke	8
7	Nagel	10	Zange	10
8	Hütte	11	Schule	11
9	Lineal	12	Schere	12
10	Keks	13	Praline	13
11	Tiger	1	Dübel	10
12	Murmel	3	Erbse	6
13	Bett	4	Eberesche	5
14	Tanne	5	Hemd	9
15	Paprika	6	Hals	7
16	Knie	7	Herd	4
17	Anzug	9	Lastwagen	2
18	Schaufel	10	Rübe	6
19	Kloster	11	Drucker	12
20	Schuhe	9	Rettich	6

A 8 Studie 5: Wortpaare**Grundschüler und ältere Erwachsene**

Blume – Garten (Übungspaar)

Ratte – Faden (Übungspaar)

Junge – Mädchen (Übungspaar)

Hund – Spinne (Übungspaar)

1. Kirche – Pfarrer
2. Möhre – Erbse
3. Hammer – Zange
4. Apfel – Birne
5. Mutter – Vater
6. Hexe – Rabe
7. Igel – Hecke
8. Fußball – Torwart
9. König – Krone
10. Biene – Honig
11. Sofa – Sessel
12. Jacke – Mantel
13. Tasche – Rucksack
14. Finger – Zehen
15. Wolke – Regen
16. Rutsche – Schaukel
17. Schere – Kleber
18. Butter – Käse
19. Käfer – Spinne
20. Besen – Schaufel
21. Wecker – Zeiger
22. Kerze – Flamme
23. Zucker – Pfeffer
24. Kissen – Decke
25. Eimer – Henkel
26. Kuchen – Torte
27. Henne – Küken
55. Pirat – Tafel
56. Wiese – Dose
57. Tube – Regal
28. Rose – Tulpe
29. Sommer – Winter
30. Schule – Lehrer
31. Zwiebel – Feuer
32. Auge – Lutscher
33. Knochen – Papier
34. Socke – Bagger
35. Seife – Kamel
36. Hotel – Knete
37. Tanne – Foto
38. Straße – Säge
39. Spinat – Kamin
40. Koffer – Lampe
41. Löwe – Flasche
42. Kiste – Treppe
43. Pinsel – Brille
44. Hase – Krücke
45. Raupe – Trommel
46. Würfel – Erde
47. Zebra – Hose
48. Katze – Plakat
49. Auto – Pfanne
50. Pflaster – Karte
51. Fahne – Esel
52. Tinte – Mauer
53. Suppe – Fenster
54. Kugel – Zimmer
58. Gesicht – Leiter
59. Eule – Flöte
60. Nebel – Zeitung

Jugendliche

Blume – Garten (Übungspaar)

Ratte – Faden (Übungspaar)

Junge – Mädchen (Übungspaar)

Hund – Spinne (Übungspaar)

1. Kirche – Pfarrer
2. Möhre – Erbse
3. Salat – Gurke
4. Hammer – Zange
5. Apfel – Birne
6. Mutter – Vater
7. Knospe – Blüte
8. Welle – Anker
9. Gitter – Käfig
10. Hexe – Rabe
11. Keller – Höhle
12. Igel – Hecke
13. Fußball – Torwart
14. Schule – Lehrer
15. Bohne – Kaffee
16. Soldat – Gewehr
17. Dusche – Wasser
18. König – Krone
19. Biene – Honig
20. Sommer – Winter
21. Tiger – Nashorn
22. Pfeife – Tabak
23. Spüle – Ofen
24. Löffel – Gabel
25. Schüssel – Tasse
26. Sofa – Sessel
27. Jacke – Mantel
28. Tasche – Rucksack
29. Stiefel – Anzug
30. Schulter – Rücken
61. Koffer – Lampe
62. Bürste – Drucker
63. Löwe – Flasche
64. Kiste – Treppe
65. Pinsel – Brille
66. Wolle – Krücke
67. Raupe – Trommel
68. Würfel – Erde
31. Finger – Zehen
32. Wolke – Regen
33. Kreisel – Puppe
34. Rutsche – Schaukel
35. Schere – Kleber
36. Butter – Käse
37. Käfer – Spinne
38. Besen – Schaufel
39. Wecker – Zeiger
40. Palme – Kaktus
41. Schlitten – Kutsche
42. Wespe – Hummel
43. Kerze – Flamme
44. Zucker – Pfeffer
45. Kissen – Decke
46. Teppich – Himmel
47. Tube – Regal
48. Nadel – Murmel
49. Zwiebel – Feuer
50. Auge – Lutscher
51. Nase – Locher
52. Knochen – Papier
53. Socke – Bagger
54. Seife – Kamel
55. Hotel – Knete
56. Tanne – Foto
57. Hocker – Schraube
58. Straße – Säge
59. Adler – Spitzer
60. Spinat – Kamin
69. Leiter – Krippe
70. Pilot – Lappen
71. Kanal – Feder
72. Zebra – Hose
73. Katze – Plakat
74. Auto – Pfanne
75. Pflaster – Karte
76. Fahne – Esel

- 77. Tinte – Mauer
- 78. Suppe – Fenster
- 79. Kugel – Zimmer
- 80. Eule – Flöte
- 81. Bahnhof – Geige
- 82. Hase – Spritze
- 83. Brücke – Muskel

- 84. Essig – Wippe
- 85. Kino – Apfel
- 86. Fabrik – Kanu
- 87. Pirat – Tafel
- 88. Magen – Lupe
- 89. Pokal – Schatten
- 90. Wiese – Dose

Jüngere Erwachsene

Blume – Garten (Übungspaar)

Ratte – Faden (Übungspaar)

Junge – Mädchen (Übungspaar)

Hund – Spinne (Übungspaar)

- | | |
|-----------------------|-------------------------|
| 1. Kirche – Pfarrer | 31. Finger – Zehen |
| 2. Möhre – Erbse | 32. Wolke – Regen |
| 3. Salat – Gurke | 33. Kreisel – Puppe |
| 4. Hammer – Zange | 34. Rutsche – Schaukel |
| 5. Apfel – Birne | 35. Schere – Kleber |
| 6. Mutter – Vater | 36. Butter – Käse |
| 7. Knospe – Blüte | 37. Käfer – Spinne |
| 8. Welle – Anker | 38. Besen – Schaufel |
| 9. Gitter – Käfig | 39. Wecker – Zeiger |
| 10. Hexe – Rabe | 40. Palme – Kaktus |
| 11. Keller – Höhle | 41. Schlitten – Kutsche |
| 12. Igel – Hecke | 42. Wespe – Hummel |
| 13. Fußball – Torwart | 43. Kerze – Flamme |
| 14. Schule – Lehrer | 44. Zucker – Pfeffer |
| 15. Bohne – Kaffee | 45. Kissen – Decke |
| 16. Soldat – Gewehr | 46. Eimer – Henkel |
| 17. Dusche – Wasser | 47. Kuchen – Torte |
| 18. König – Krone | 48. Henne – Küken |
| 19. Biene – Honig | 49. Rose – Tulpe |
| 20. Sommer – Winter | 50. Traube – Pfirsich |
| 21. Tiger – Nashorn | 51. Armband – Kette |
| 22. Pfeife – Tabak | 52. Wimper – Braue |
| 23. Spüle – Ofen | 53. Drache – Ritter |
| 24. Löffel – Gabel | 54. Wunde – Narbe |
| 25. Schüssel – Tasse | 55. Fahrrad – Klingel |
| 26. Sofa – Sessel | 56. Reiter – Sattel |
| 27. Jacke – Mantel | 57. Karton – Pappe |
| 28. Tasche – Rucksack | 58. Tempel – Kuppel |
| 29. Stiefel – Anzug | 59. Muskat – Nelke |
| 30. Schulter – Rücken | 60. Maler – Leinwand |

61. Teppich – Himmel
62. Tube – Regal
63. Nadel – Murrel
64. Zwiebel – Feuer
65. Auge – Lutscher
66. Nase – Locher
67. Knochen – Papier
68. Socke – Bagger
69. Seife – Kamel
70. Hotel – Knete
71. Tanne – Foto
72. Hocker – Schraube
73. Straße – Säge
74. Adler – Spitzer
75. Spinat – Kamin
76. Koffer – Lampe
77. Bürste – Drucker
78. Löwe – Flasche
79. Kiste – Treppe
80. Pinsel – Brille
81. Wolle – Krücke
82. Raupe – Trommel
83. Würfel – Erde
84. Leiter – Krippe
85. Pilot – Lappen
86. Kanal – Feder
87. Zebra - Hose
88. Katze – Plakat
89. Auto – Pfanne
90. Pflaster – Karte
91. Fahne – Esel
92. Tinte – Mauer
93. Suppe – Fenster
94. Kugel – Zimmer
95. Eule – Flöte
96. Bahnhof – Geige
97. Hase – Spritze
98. Brücke – Muskel
99. Essig – Wippe
100. Kino – Ampel
101. Fabrik – Kanu
102. Pirat – Tafel
103. Magen – Lupe
104. Pokal – Schatten
105. Wiese – Dose
106. Felsen – Leine
107. Gesicht – Donner
108. Spiegel – Möwe
109. Nebel – Zeitung
110. Samen – Villa
111. Seide – Heizung
112. Reifen – Tacker
113. Liste – Hamster
114. Tüte – Zecke
115. Kasse – Orden
116. Boden – Vase
117. Aufzug – Rechner
118. Gelenk – Futter
119. Hütte – Gummi
120. Palast - Fichte

A 9 Studie 6: Wortpaare**Grundschüler und ältere Erwachsene**

Blume – Garten (Übungspaar)

Ratte – Faden (Übungspaar)

Junge – Mädchen (Übungspaar)

Hund – Spinne (Übungspaar)

- | | |
|------------------------|----------------------|
| 1. Kirche – Pfarrer | 25. Auge – Lutscher |
| 2. Möhre – Erbse | 26. Knochen – Papier |
| 3. Apfel – Birne | 27. Socke – Bagger |
| 4. Mutter – Vater | 28. Seife – Kamel |
| 5. Igel – Hecke | 29. Hotel – Knete |
| 6. Fußball – Torwart | 30. Tanne – Foto |
| 7. König – Krone | 31. Straße – Säge |
| 8. Biene – Honig | 32. Spinat – Kamin |
| 9. Sofa – Sessel | 33. Koffer – Lampe |
| 10. Jacke – Mantel | 34. Löwe – Flasche |
| 11. Tasche – Rucksack | 35. Pinsel – Brille |
| 12. Finger – Zehen | 36. Hase – Krücke |
| 13. Wolke – Regen | 37. Zebra – Hose |
| 14. Rutsche – Schaukel | 38. Katze – Plakat |
| 15. Schere – Kleber | 39. Auto – Pfanne |
| 16. Käfer - Spinne | 40. Pflaster – Karte |
| 17. Besen – Schaufel | 41. Fahne – Esel |
| 18. Wecker – Zeiger | 42. Tinte – Mauer |
| 19. Kerze – Flamme | 43. Suppe – Fenster |
| 20. Zucker – Pfeffer | 44. Kugel – Zimmer |
| 21. Kissen – Decke | 45. Pirat – Tafel |
| 22. Eimer – Henkel | 46. Wiese – Dose |
| 23. Kuchen – Torte | 47. Tube – Regal |
| 24. Henne – Küken | 48. Eule – Flöte |

Jugendliche

Blume – Garten (Übungspaar)

Ratte – Faden (Übungspaar)

Junge – Mädchen (Übungspaar)

Hund – Spinne (Übungspaar)

- | | |
|---------------------|-------------------|
| 1. Kirche – Pfarrer | 3. Apfel – Birne |
| 2. Möhre – Erbse | 4. Mutter – Vater |

5. Knospe – Blüte
6. Welle – Anker
7. Gitter – Käfig
8. Hexe – Rabe
9. Keller – Höhle
10. Igel – Hecke
11. Fußball – Torwart
12. Schule – Lehrer
13. Bohne – Kaffee
14. Soldat – Gewehr
15. Dusche – Wasser
16. König – Krone
17. Biene – Honig
18. Pfeife – Tabak
19. Spüle – Ofen
20. Löffel – Gabel
21. Schüssel – Tasse
22. Sofa – Sessel
23. Jacke – Mantel
24. Tasche – Rucksack
25. Stiefel – Anzug
26. Schulter – Rücken
27. Finger – Zehen
28. Wolke – Regen
29. Kreisel – Puppe
30. Rutsche – Schaukel
31. Schere – Kleber
32. Käfer – Spinne
33. Besen – Schaufel
34. Wecker – Zeiger
35. Drache – Ritter
36. Muskat – Nelke
37. Teppich – Himmel
38. Tube – Regal
39. Nadel – Marmor
40. Zwiebel – Feuer
41. Auge – Lutscher
42. Knochen – Papier
43. Socke – Bagger
44. Seife – Kamel
45. Hotel – Knete
46. Tanne – Foto
47. Hocker – Schraube
48. Straße – Säge
49. Adler – Spitzer
50. Spinat – Kamin
51. Koffer – Lampe
52. Löwe – Flasche
53. Pinsel – Brille
54. Wolle – Krücke
55. Raupe – Trommel
56. Würfel – Erde
57. Leiter – Krippe
58. Zebra – Hose
59. Katze – Plakat
60. Auto – Pfanne
61. Pflaster – Karte
62. Fahne – Esel
63. Tinte – Mauer
64. Suppe – Fenster
65. Kugel – Zimmer
66. Eule – Flöte
67. Bahnhof – Geige
68. Kino – Ampel
69. Fabrik – Kanu
70. Pirat – Tafel
71. Boden – Vase
72. Palast – Fichte

Jüngere Erwachsene

Blume – Garten (Übungspaar)

Ratte – Faden (Übungspaar)

Junge – Mädchen (Übungspaar)

Hund – Spinne (Übungspaar)

1. Kirche – Pfarrer
2. Möhre – Erbse
3. Salat – Gurke
4. Hammer – Zange
5. Apfel – Birne
6. Mutter – Vater
7. Knospe – Blüte
8. Welle – Anker
9. Gitter – Käfig
10. Hexe – Rabe
11. Keller – Höhle
12. Igel – Hecke
13. Fußball – Torwart
14. Schule – Lehrer
15. Bohne – Kaffee
16. Soldat – Gewehr
17. Dusche – Wasser
18. König – Krone
19. Biene – Honig
20. Sommer – Winter
21. Tiger – Nashorn
22. Pfeife – Tabak
23. Spüle – Ofen
24. Löffel – Gabel
25. Schüssel – Tasse
26. Sofa – Sessel
27. Jacke – Mantel
28. Tasche – Rucksack
29. Stiefel – Anzug
30. Schulter – Rücken
31. Finger – Zehen
32. Wolke – Regen
33. Kreisel – Puppe
34. Rutsche – Schaukel
35. Schere – Kleber
36. Butter – Käse
37. Käfer – Spinne
38. Besen – Schaufel
39. Wecker – Zeiger
40. Palme – Kaktus
41. Schlitten – Kutsche
42. Wespe – Hummel
43. Kerze – Flamme
44. Zucker – Pfeffer
45. Kissen – Decke
46. Eimer – Henkel
47. Kuchen – Torte
48. Henne – Küken
49. Rose – Tulpe
50. Traube – Pfirsich
51. Armband – Kette
52. Wimper – Braue
53. Drache – Ritter
54. Wunde – Narbe
55. Fahrrad – Klingel
56. Reiter – Sattel
57. Karton – Pappe
58. Tempel – Kuppel
59. Muskat – Nelke
60. Maler – Leinwand
61. Teppich – Himmel
62. Tube – Regal
63. Nadel – Murmel
64. Zwiebel – Feuer
65. Auge – Lutscher
66. Nase – Locher
67. Knochen – Papier
68. Socke – Bagger
69. Seife – Kamel
70. Hotel – Knete
71. Tanne – Foto
72. Hocker – Schraube
73. Straße – Säge
74. Adler – Spitzer

75. Spinat – Kamin
76. Koffer – Lampe
77. Bürste – Drucker
78. Löwe – Flasche
79. Kiste – Treppe
80. Pinsel – Brille
81. Wolle – Krücke
82. Raupe – Trommel
83. Würfel – Erde
84. Leiter – Krippe
85. Pilot – Lappen
86. Kanal – Feder
87. Zebra - Hose
88. Katze – Plakat
89. Auto – Pfanne
90. Pflaster – Karte
91. Fahne – Esel
92. Tinte – Mauer
93. Suppe – Fenster
94. Kugel – Zimmer
95. Eule – Flöte
96. Bahnhof – Geige
97. Hase – Spritze
98. Brücke – Muskel
99. Essig – Wippe
100. Kino – Ampel
101. Fabrik – Kanu
102. Pirat – Tafel
103. Magen – Lupe
104. Pokal – Schatten
105. Wiese – Dose
106. Felsen – Leine
107. Gesicht – Donner
108. Spiegel – Möwe
109. Nebel – Zeitung
110. Samen – Villa
111. Seide – Heizung
112. Reifen – Tacker
113. Liste – Hamster
114. Tüte – Zecke
115. Kasse – Orden
116. Boden – Vase
117. Aufzug – Rechner
118. Gelenk – Futter
119. Hütte – Gummi
120. Palast - Fichte

B 1 Studie 1: Instruktion

Hallo,

ich freue mich sehr, dass du bei unserem Experiment mitmachen möchtest. Vielen Dank!

Bei Kindern und Jugendlichen: Es ist ganz wichtig, dass du weißt, dass es bei unserem Versuch nicht um deine schulische Leistung oder eine Schulnote geht.

Bei Erwachsenen: Die Aufgaben sind zum Teil extra so gemacht, dass man nicht alles können kann. Wenn alle alles könnten, müsste man ja keine Studie machen...

Film

Du wirst nun einen 6-minütigen Film sehen. Am Ende des Experiments werden dir zum Film Fragen gestellt. Dabei ist es wichtig, dass du dich an möglichst viele Einzelheiten aus dem Film erinnerst. Aber es ist auch wichtig, dass du dir alles richtig merkst, so wie es im Film vorkommt. Also passe wirklich gut auf und versuche dir möglichst viel zu merken!

(Film am PC zeigen)

Abgabe JOLs zum Film

Hier siehst du nun einige Fragen zum Film.

Du sollst diese Fragen nun noch nicht beantworten, sondern einschätzen, wie gut du diese Fragen wohl in einer halben Stunde beantworten kannst. Wenn du dir ganz sicher bist, dass du diese Frage später richtig beantworten kannst, dann kreuze den lachenden Smiley an. Wenn du denkst, dass du diese Frage später nicht beantworten kannst, kreuze den unglücklichen Smiley an. Du kannst jeden Punkt bzw. jeden Smiley ankreuzen; du kannst also die ganze Skala verwenden.

Bildpaare

Lernen der Bildpaare

Jetzt folgt eine weitere Aufgabe:

Du wirst gleich nacheinander 20 Bildpaare am Laptop gezeigt bekommen, die du dir genau anschauen und gut merken sollst. Das kann zum Beispiel so aussehen. (*Erstes Übungsbildpaar wird gezeigt, auf „weiter“ klicken für zweites Übungsbildpaar*). Versuche, dir am besten so viele Bildpaare wie möglich einzuprägen. Denn danach wird dir nur das linke Bild gezeigt und du sollst einschätzen, wie gut du dich an das dazugehörige Bild erinnern wirst und auch das fehlende Bild benennen.

Jetzt zeige ich dir die Bildpaare und du sollst sie dir gut merken!

(Bildpaarpräsentation starten)

Abgabe JOLs zu den Bildpaaren

Nun wird dir von den Bildpaaren immer das linke Bild gezeigt. Du sollst aber noch nicht das dazugehörige Bild benennen, sondern erst einmal auf einer Skala einschätzen, wie gut du dich wohl in einer halben Stunde an das Bild, das dazu gehört, erinnern wirst. Die Skala geht von „ganz kalt“ bis „ganz heiß“. Wobei „ganz kalt“ dafür steht, dass du dich gar nicht an das Bild erinnern wirst und „ganz heiß“, dass du ganz sicher bist, dass du dich an die richtige Antwort erinnern wirst. Je länger du überlegen müsstest, um das richtige Bild zu finden, desto „kälter“ sollst du deine Einschätzung bewerten. Je sicherer du allerdings bist, dass du das richtige Bild später weißt, desto „wärmer“ sollte deine Bewertung sein.

Du sollst die Maus selbst bedienen und kannst dir die Skala als Thermometer vorstellen. Wichtig ist noch, dass du auch wirklich die ganze Breite der Skala ausnutzen kannst.

Jetzt können wir uns noch die beiden Übungsbeispiele anschauen. Bediene die Maus am besten selbst. Mit Klick auf „Weiter“ gelangst du zum nächsten Bildpaar.

(Abgabe der JOLs starten)

Zwischenaufgaben: Zahlennachsprechen, Wortschatz-Test, demografischer Fragebogen bei jüngeren und älteren Erwachsenen

Fragen zum Film

Nun sollst du die Fragen zum Film beantworten. Schreibe unter jede Frage deine Antwort. Wenn du dir nicht ganz sicher bist bei deiner Antwort, dann macht das nichts. Du darfst auch raten.

Nach jeder Antwort sollst du noch ankreuzen, wie sicher du dir bist, dass du richtig geantwortet hast. Wenn du dir ganz sicher bist, kreuze den lachenden Smiley an, wenn du dir überhaupt nicht sicher bist, ob du richtig geantwortet hast, kreuze den unglücklichen Smiley an. Du kannst wieder jeden Punkt bzw. jeden Smiley ankreuzen, also die ganze Skala ausnützen.

Nach jeder Antwort müssen die Sicherheitsurteile erfasst werden, außer wenn eine Antwort nicht gewusst wird!

Abfrage der Bildpaare

Jetzt kommen wir wieder zu dem Teil mit den Bildpaaren zurück. Dir wird nun immer nur das linke Bild eines Paares gezeigt und du sollst dich an das dazugehörige Bild erinnern. Du sagst mir dann deine Antwort und ich werde sie eingeben. Wenn du dich an ein Bildpaar nicht mehr erinnern kannst, ist das nicht weiter schlimm und wir gehen einfach zum nächsten weiter. Versuche dich aber an so viele Bildpaare wie möglich zu erinnern.

(Abfrage der Bildpaare starten)

Auch hier sollst du mir sagen, wie sicher du dir bist, dass deine Antwort stimmt. Wenn du dir ganz sicher bist, zeige auf diesem Bogen (*Vorlage mit den Smileys zeigen*) auf den lachenden Smiley, wenn du dir sehr unsicher bist, zeige auf den unglücklichen Smiley; du kannst wieder auf jeden Punkt bzw. jeden Smiley deuten, also die ganze Skala ausnutzen.

(Nach jeder Antwort müssen die Sicherheitsurteile erfasst werden! Der Versuchsleiter trägt die Antwort, die die VP auf der Vorlage zeigt, in die entsprechende Zeile des Antwortbogens ein.)

Vielen Dank für deine Teilnahme!

B 2 Fragebogen zu demografischen Angaben

1. Geburtsjahr: _____
2. Alter: _____
3. Geschlecht: _____
4. Was ist Ihre derzeitige Nationalität? _____
5. Was ist Ihr Familienstand? *(Bitte kreuzen Sie die für Sie zutreffende Antwortmöglichkeit an.)*

Ledig	<input type="checkbox"/>
Verheiratet	<input type="checkbox"/>
Geschieden	<input type="checkbox"/>
Verwitwet	<input type="checkbox"/>

6. Haben Sie einen Partner / eine Partnerin? Ja Nein
7. Leben Sie alleine mit anderen Personen zusammen?
8. Was ist der höchste Bildungsabschluss, den Sie erreicht haben? *(Bitte kreuzen Sie eine Antwortmöglichkeit an.)*

Ohne Abschluss	<input type="checkbox"/>
Hauptschulabschluss oder Ähnliches	<input type="checkbox"/>
Realschul- / Mittelschulabschluss	<input type="checkbox"/>
Berufsschulabschluss oder Ähnliches	<input type="checkbox"/>
Abitur	<input type="checkbox"/>
Universitätsabschluss	<input type="checkbox"/>
Promotion	<input type="checkbox"/>
Andere, bitte angeben _____	<input type="checkbox"/>

9. Wie viele Jahre haben Sie insgesamt die Schule besucht (Universität / Berufsschule eingeschlossen)?
_____ Jahre
10. Was war / ist Ihre hauptsächliche Berufstätigkeit? _____
11. Gehen Sie derzeit einer bezahlten Tätigkeit nach? Ja Nein
12. Falls Sie im Ruhestand sind, seit wie lange? Seit _____ Jahren.

13. Hier finden Sie eine Liste von verschiedenen Krankheiten: Leiden Sie an...?

(Bitte kreuzen Sie bei jeder Frage eine Antwortmöglichkeit an.)

	Ja	Nein	Früher ja, aber nicht jetzt
Allergie			
Asthma			
Diabetes			
Augenkrankheiten (z.B. Katarakte)			
Kardiovaskuläre Krankheiten (z.B. hoher Blutdruck)			
Leberkrankheiten (z.B. Hepatitis)			
Magen-Darm-Krankheiten (z.B. Magengeschwür)			
Lungenkrankheiten (z.B. chronisch-obstruktives Lungenleiden)			
Krebs, bitte angeben _____			
Andere Krankheiten, bitte angeben _____			

14. Hier finden Sie eine Liste mit Medikamenten. Falls Sie welche verwenden, wie oft?

(Bitte kreuzen Sie bei jeder Frage eine Antwortmöglichkeit an.)

	Nein	Nach Notwendigkeit	Regelmäßig	Täglich
Schmerzmittel				
Schlafmittel				
Medikamente gegen psychische Erkrankungen				
Medikamente gegen Allergien				
Medikamente gegen Asthma				
Medikamente für das Herz				
Medikamente gegen Diabetes				
Natur- und Pflanzenmedizin (z.B. Homöopathie)				
Andere Medikamente, bitte angeben _____				

15. Wie viele Medikamente nehmen Sie insgesamt täglich ein? _____ *(Anzahl)*

16. Wie würden Sie Ihren allgemeinen Gesundheitszustand einschätzen? *(Bitte kreuzen Sie die für Sie zutreffende Antwort an.)*

Sehr gut	
Gut	
Mittel	
Eher Schlecht	
Schlecht	

C 1 Studien 2 und 3: Fußball-Wissenstest**Grundschüler**

1. *Interessierst du Dich für Fußball?*
 - Ja, sehr
 - Ja, etwas
 - Ich weiß es nicht recht
 - Nein, weniger
 - Nein, gar nicht

2. *Schaust Du Dir die Fußballspiele an, die im Fernsehen übertragen werden?*
 - Immer
 - Meistens
 - Manchmal
 - Selten
 - Nie

3. *Spielst du selbst Fußball?*
 - Sooft es geht
 - Häufig
 - Manchmal
 - Kaum
 - Nie

4. *Was ist im Fußball ein Auswärtsspiel?*
 - Die Mannschaft kommt in einem Fußballturnier eine Runde weiter
 - Ein Spieler erreicht den Ball erst hinter der Auslinie
 - Ein Spieler spielt probeweise in einem anderen Verein mit
 - Die Zuschauer unterstützen die Mannschaft des Gegners
 - Die Mannschaft spielt auf dem Platz des Gegners
 - Keine Ahnung

5. *Welche Mannschaft darf den Anstoß am Anfang des Spiels ausführen?*
 - Die Mannschaft, die das Münzwerfen verliert
 - Die Mannschaft, die der Schiedsrichter bestimmt
 - Die Gastmannschaft
 - Die schlechtere Mannschaft
 - Die Heimmannschaft (auf deren Platz das Spiel stattfindet)
 - Keine Ahnung

6. *Was bedeutet im Fußball, „eine Mannschaft spielt offensiv“?*
- Die Mannschaft ist ihrem Gegner überlegen
 - Die Spieler lassen der gegnerischen Mannschaft viel Spielraum
 - Die Mannschaft spielt in die Breite
 - Die Mannschaft ist besonders auswärtsstark
 - Die Mannschaft spielt auf Angriff
 - Keine Ahnung
7. *Wo auf dem Spielfeld geht ein Spiel weiter, nachdem ein Tor gefallen ist?*
- An der Eckfahne
 - Im Mittelkreis
 - Im Strafraum
 - Im Torraum
 - An der Seitenauslinie
 - Keine Ahnung
8. *Wie wird gemeldet, dass ein Spieler im Abseits steht?*
- Der Trainer macht dem Schiedsrichter vom Spielfeldrand einen Zuruf
 - Der Linienrichter geht mit gesenkter Fahne in Richtung Mittellinie
 - Bei der nächsten Spielunterbrechung sagt der Linienrichter dem Schiedsrichter Bescheid
 - Der Linienrichter hebt mit ausgestrecktem Arm die Fahne
 - Der Schiedsrichter zeigt dem Spieler die gelbe Karte
 - Keine Ahnung
9. *Was ist eine Vorlage?*
- Wenn eine Mannschaft mit einem Tor führt
 - Wenn ein Torwart den Ball zu seinem Verteidiger spielt
 - Wenn ein Stürmer einen Ball so gut zugespielt bekommt, dass er ein Tor schießen kann
 - Das Dach über der Bank, auf der die Auswechselspieler und der Trainer sitzen
 - Der Teil der Stadionüberdachung, der am weitesten in das Spielfeld hineinreicht
 - Keine Ahnung
10. *Wievielmahl wird in einem Fußballspiel mindestens ein Anstoß ausgeführt?*
- 3 Mal
 - 5 Mal
 - 4 Mal
 - 1 Mal
 - 2 Mal
 - Keine Ahnung

11. *Aus welcher Stadt kommt der VfB?*

- Bremen
- Stuttgart
- Dortmund
- Berlin
- Hamburg
- Keine Ahnung

12. *Was ist der Strafraum?*

- Der Raum, in den die Spieler geschickt werden, die ein schlimmes Foul begangen haben
- Der Kreis in der Mitte des Spielfeldes
- Der Bereich, in dem man den Gegner nicht foulen darf
- Der Bereich, in dem der Torwart den Ball in die Hand nehmen darf
- Der Raum, in dem die Spieler ihre Strafzeit absitzen müssen
- Keine Ahnung

Jugendliche, jüngere und ältere Erwachsene

Wir möchten wissen, wie gut Sie sich mit Fußball auskennen und wie sehr Sie sich für Fußball interessieren!

Deshalb finden Sie hier ein paar Fragen zum Thema Fußball, die Sie nun beantworten sollen. Viel Spaß dabei!

1. Was ist eine Notbremse beim Fußball?

2. Was ist eine Flanke? (Kreuzen Sie die richtige Antwort an.)

- Ein hoher Pass in den gegnerischen Strafraum/ vor das gegnerische Tor
- Ein Pass von der linken Spielfeldseite auf die Rechte
- Ein langer Ball von hinten nach vorne

3. Wie viele Mannschaften spielen in der 1. Bundesliga?

4. Wann steht ein Spieler Abseits? (Kreuzen Sie die richtige Antwort an.)

- Wenn er beim Anspiel näher zum Tor steht als zwei Verteidiger.
- Wenn er beim Anspiel näher zum Tor steht als ein Verteidiger.
- Wenn er beim Anspiel näher zum Tor steht als der Torhüter.
- Wenn er das Spielfeld verlässt.

5. In welchem Land fand die letzte WM statt?

6. Was bedeutet „schwalben“? (Kreuzen Sie die richtige Antwort an.)

- Sich fallen lassen, um einen Freistoß zu bekommen
- Hoch über den Gegenspieler hinweg springen
- Den Ball besonders hoch schießen
- Sich lautstark beschweren

7. Wo darf der Torwart den Ball mit der Hand berühren? (Kreuzen Sie die richtige Antwort an.)

- Im Fünfmeterraum
- Im Sechzehnmeterraum
- Im Siebenmeterraum

8. Wo fand die erste Fußball-Weltmeisterschaft statt? (Kreuzen Sie die richtige Antwort an.)

- Spanien
- Deutschland
- Uruguay
- USA
- Russland

9. Welches Land ist Rekordweltmeister im Fußball?

10. Was ist eine Viererkette beim Fußball?

Wie sehr interessieren Sie sich für Fußball? (Kreuzen Sie an, was für Sie zutrifft.)

- Gar nicht Wenig Mittelmäßig Stark Sehr stark

Sehen Sie sich die Fußballspiele an, die im Fernsehen übertragen werden? (Kreuzen Sie an, was für Sie zutrifft.)

- Nie Selten Manchmal Meistens Sehr oft

Spielen Sie selbst Fußball? (Kreuzen Sie an, was für Sie zutrifft.)

- Nie Selten Manchmal Häufig Sehr oft

C 2 Studie 4: Strategietraining

Grundschüler

Es gibt verschiedene Möglichkeiten, sich Informationen besser zu merken. Zum Beispiel kann man sich Dinge besser merken, wenn mal sie wiederholt.

Eine andere und sehr gute Möglichkeit, sich Informationen besser zu merken, möchte ich dir jetzt erklären.

Wenn du dir etwas Schwieriges merken sollst, und dir dazu ein Bild vorstellst, fällt es dir leichter dich später daran zu erinnern.

Ich möchte dir das an einem Beispiel deutlich machen. Wenn ich dir den Satz sage: *Ein Zweibein sitzt auf einem Dreibein und isst dabei ein Einbein. Da kommt ein Vierbein und nimmt dem Zweibein das Einbein weg.*

Kannst du mir diesen Satz nachsprechen? (...) Es fällt dir wahrscheinlich etwas schwer.

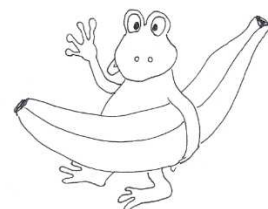
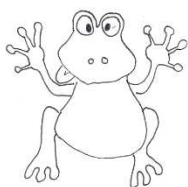
Wenn ich dir diesen Satz sage: *Ein Mensch sitzt auf einem Hocker und isst eine Hühnerkeule. Da kommt ein Hund und nimmt dem Menschen die Hühnerkeule weg.*

Kannst du mir diesen Satz nachsprechen? (...) Es fällt dir wahrscheinlich leichter, diesen Satz zu wiederholen, weil du dir ein Bild zu den Wörtern gemacht hast. Das kann sich unser Gehirn leichter merken.

Wenn du versuchst, dir Wortpaare zu merken, kannst du das genauso mit Bildern machen.

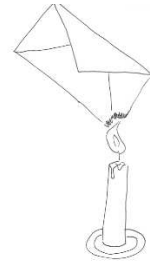
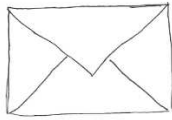
Ich sag dir noch mal ein Beispiel. Stell dir vor, du möchtest dir die Wörter „Frosch“ und „Banane“ zusammen merken. Wenn ich dir also später das Wort „Frosch“ sage, sollst du noch wissen, dass „Banane“ dazu gehört.

Jetzt sollst du dir also ein Bild dazu vorstellen. Zum Beispiel einen Frosch, der eine Banane trägt. Dann kannst du dir „Frosch“ und „Banane“ leichter zusammen merken. Dabei ist es wichtig, dass bei dem Bild die beiden Wörter vorkommen und miteinander zu tun haben. Es funktioniert besser, sich das Wortpaar zu merken, wenn du dir nicht nur einen Frosch und eine Banane nebeneinander vorstellst, sondern wenn beide Wörter im Bild verbunden sind.



Hier siehst du noch ein Beispiel:

Wenn du „Brief“ und „Kerze“ zusammen lernen sollst, so dass du, wenn ich dir später „Brief“ sage, noch weißt, dass da das Wort „Kerze“ dazugehört, ist es hilfreich, sich beide Wörter als ein Bild zusammen vorzustellen. Also zum Beispiel eine Kerze, die einen Brief verbrennt.



Am besten ist es, wenn in dem Bild, das du dir vorstellst, eine Handlung vorkommt, also etwas gemacht wird, etwas passiert.

Wir werden dies nun an ein paar Beispielen üben. Bitte überleg dir sich für die folgenden Wortpaare jeweils ein Bild, in dem beide Wörter miteinander zu tun haben, und sag mir, welches Bild du dir vorstellst.

1. Wiese – Bleistift
2. Kiste – Zitrone
3. Rose – Wolle
4. Käse – Heft
5. Brille – Gabel
6. Stier – Ring
7. Lampe – Pferd
8. Fliege – Eimer
9. Besen – Apfel
10. Maulwurf – Wolke

Bsp. Ein Bleistift zeichnet eine Wiese...

Wichtig für den Testleiter: Die Bilder müssen wirklich interaktiv sein, d.h. beide Wörter müssen zusammenhängend vorkommen. Nach jeder Antwort der Versuchsperson muss rückgemeldet werden, ob die Elaboration gut gelungen ist oder nicht. Wenn nicht, muss gemeinsam nach einem interaktiven Bild gesucht werden.

Jugendliche, jüngere und ältere Erwachsene

Es gibt verschiedene Möglichkeiten, sich Informationen besser zu merken. Zum Beispiel kann man sich Dinge besser merken, wenn mal sie wiederholt oder wenn man sie sinnvoll sortiert.

Eine andere und sehr wirksame Möglichkeit, sich Informationen besser zu merken, stellt die sog. „Elaboration“ dar.

„Informationen, die wir uns einprägen wollen, werden dabei erweitert, indem noch sprachliches oder bildhaftes Material hinzugefügt wird. Mit anderen Worten, es soll eine Art „Eselsbrücke“ gebildet werden. Denn wenn man sich zusätzlich ein Bild merkt, kommt es zu einer doppelten Verarbeitung im Gehirn, nämlich einmal als Wort und einmal als Bild. Dadurch wird die Merkleistung erhöht.

Ich möchte Ihnen das an einem Beispiel deutlich machen. Wenn ich Ihnen den Satz sage: *Ein Zweibein sitzt auf einem Dreibein und isst dabei ein Einbein. Da kommt ein Vierbein und nimmt dem Zweibein das Einbein weg.*

Können Sie mir diesen Satz nachsprechen? (...) Es fällt Ihnen wahrscheinlich etwas schwer.

Wenn ich Ihnen diesen Satz sage: *Ein Mensch sitzt auf einem Hocker und isst eine Hühnerkeule. Da kommt ein Hund und nimmt dem Menschen die Hühnerkeule weg.*

Können Sie diesen Satz nachsprechen? (...) Es fällt Ihnen wahrscheinlich leichter, diesen Satz zu wiederholen, weil Sie sich ein Bild zu den Wörtern gemacht haben. Das kann unser Gehirn leichter verarbeiten und speichern.

Das gleiche Prinzip kann man auch anwenden, wenn man versucht, sich Wortpaare zu merken.

Nehmen wir an, Sie möchten versuchen, die die beiden Wörter „Banane“ und „Frosch“ zusammen zu merken, so dass Sie, wenn ich Ihnen später „Banane“ sage, noch wissen, dass da das Wort „Frosch“ dazu gehört.

Wenn man nun dem Prinzip der „Elaboration“ folgt, ist es leichter, sich dieses Wortpaar zu merken, wenn Sie sich ein Bild dazu vorstellen, z.B. einen Frosch, der eine Banane trägt. Dabei ist es wichtig, dass bei dem Bild die beiden Wörter vorkommen und miteinander zu tun haben. Es funktioniert besser, sich das Wortpaar zu merken, wenn Sie sich nicht nur eine Banane und einen Frosch nebeneinander vorstellen, sondern wenn beide Wörter im Bild verbunden sind [*Weiter wie bei den Grundschülern*].

C 3 Studie 4: Fragebogen zum Strategietraining

Strategietraining JA

- Hast du die Strategie, dir zwei zusammenhängende Bilder vorzustellen, verwendet?
- Bei wie viel Prozent der Wortpaare hast du die Strategie verwendet? Bzw. bei Kindern: Wie oft hast du die Strategie verwendet (Smiley-Skala)?



- Ist es dir leicht gefallen, dir Bilder vorzustellen?
- Denkst du, dass diese Strategie dir geholfen hat, dir die Wortpaare zu merken?
- Hast du noch andere Strategien verwendet, um dir die Wortpaare zu merken?
- Anmerkungen des Testleiters:

Strategietraining NEIN

- Hast du eine Strategie verwendet, um dir die Wortpaare besser merken zu können?
- Wenn ja: Welche? (Können auch mehrere sein)
- Wenn ja: Bei wie viel Prozent der Wortpaare hast du eine Strategie verwendet? Bzw. bei Kindern: Wie oft hast du eine Strategie verwendet (Smiley-Skala)?



- Anmerkungen des Testleiters:

C 4 Studie 6: Fragebogen zum Lernanreiz

- Wie wichtig war es dir / Ihnen, im zweiten Lerndurchgang möglichst viele Punkte zu sammeln?

Sehr unwichtig sehr wichtig

- Wie wichtig war es dir / Ihnen, möglichst viel von den Extra-Geschenken / von dem Geld zu bekommen?

Sehr unwichtig sehr wichtig

- Denkst du / denken Sie, dass die Extra-Belohnung dich / Sie motiviert hat, die Wortpaare mit den fünf Punkten besser und länger zu lernen?

Nein, gar nicht ja, sehr stark

- **Nur bei Kindern und Jugendlichen:** Wie gut haben dir die Extra-Geschenke gefallen?

Überhaupt nicht sehr gut

- Anmerkungen des Testleiters: