universität wien

Diplomarbeit

Titel der Diplomarbeit

Szenenerkennung durch wiederholt betrachtete, natürliche Reizmerkmale

Verfasser

Gerhard Böck

Angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat)

Wien, im Jänner 2013

Studienkennzahl:	298
Studienrichtung:	Psychologie
Betreuer:	Univ.-Prof. Dr. Ulrich Ansorge

Danksagung

Ich danke in erster Linie Prof. Dr. Ulrich Ansorge für die erstklassige Betreuung meiner Diplomarbeit, sowie für sein herzliches Bemühen und Entgegenkommen in verschiedensten Situationen auf dem Weg zur Entstehung dieser Arbeit. Ich fühlte mich immer respektiert und herzlich behandelt, was meine Arbeit wesentlich erleichterte.

Außerdem möchte ich Mag. Christian Valuch für all die herzlichen und zugleich sehr kompetenten Hilfestellungen, welche im Zuge dieser Arbeit immer wieder nötig waren, danken. Er war mir in jeder Phase dieser Arbeit, mit Rat und Tat zur Seite und stets geduldig bemüht alle Unklarheiten zu beseitigen. Diese großartige Unterstützung trug wesentlich zum Entstehen dieser Arbeit bei.

Ein herzliches Dankeschön möchte ich auch an all die Versuchspersonen richten, welche freiwillig und unentgeltlich die Mühe auf sich nahmen, um an diesem Experiment teilzunehmen. Ohne diese Menschen wäre diese Arbeit nicht möglich gewesen.

Besonders möchte ich mich auch bei meiner Mutter Eva Böck und meinem Vater Walter Böck für all die tatkräftige, finanzielle und praktische Unterstützung bedanken, welche mir die Arbeit wesentlich erleichterte.

Abschließend möchte ich mich bei all meinen Freunden und Freundinnen bedanken, welche mich immer wieder moralisch stützten und mir auch praktische Tipps gaben. Besonders möchte ich an dieser Stelle all jenen lieben Freunden und Freundinnen danken, welche sich die Zeit nahmen um diese Arbeit Korrektur zu lesen.

Überblick

Können foveal fixierte Bereiche natürlicher Szenen wiedererkannt werden, wenn der Szenenkontext nicht zur Verfügung steht? Bisherige Forschungen zum visuellen Gedächtnissystem zeigen teilweise widersprüchliche Auffassungen, was die Art und Weise der visuellen Verarbeitung betrifft. Die vorliegende Arbeit soll die unterschiedlichen Paradigmen auf diesem Gebiet aufzeigen und einen Beitrag dazu leisten, die einleitende Fragestellung zu klären. Zu diesem Zweck wurde die Wiedererkennungsleistung von fixierten Bildbereichen natürlicher Szenen getestet, wobei in der Testphase nur so kleine Bildausschnitte zur Wiedererkennung präsentiert wurden, dass der Szenenkontext großteils ausgeblendet wurde. Dabei wurde ein Blickverfolgungsgerät verwendet, welches die Fixationen der Versuchspersonen während der Betrachtung von Bildern natürlicher Szenen aufzeichnete. Nachdem alle Bilder betrachtet worden waren, wurde die Gedächtnisleistung gemessen, in dem fixierte und nicht fixierte Bereiche der zuvor betrachteten Bilder, sowie Bildbereiche neuer nicht gezeigter Bilder, vorgegeben und getestet wurden. Bei jeder Versuchsperson wurden die Fixationspunkte, die Reaktionsleistungen sowie die Wiedererkennungsrate gemessen und anschließend statistisch ausgewertet. Fixierte Bildbereiche konnten signifikant besser und schneller wiedererkannt werden als nicht fixierte Bildbereiche. Die Ergebnisse werden im Anschluss mit den bisherigen Forschungsergebnissen verglichen und diskutiert.

Abstract

Is it possible to remember fixated locations of presented images of natural scenes, when the scene context is not available? Previous investigations in the visual memory system show different views on visual processing. The present study illustrates the different paradigms in this field and contributes to answer the introductory question. For that purpose the recognition performance for fixated regions of natural scenes was tested. The pictures in the recognition test where so small that the scene context was mostly cut off. During the observation of the natural images an eye tracker was used to record the participants fixations. After the participants saw all the images, the recognition performance was measured by showing small fixated and not fixated cut outs of the previously shown images as well as cut outs of new, not presented images. For each subject the fixations, reaction times and recognition performance was measured and subsequently statistically evaluated. Recognition performance and reaction times were significantly better for fixated regions of the natural images than for not fixated regions. Subsequently the results will be discussed and compared with past research.

Inhaltsverzeichnis

1 Einleitung

Um auf dieser Erde selbstständig überleben zu können, ist es notwendig sich mit allen Sinnesreizen bewusst oder unbewusst auseinandersetzen zu können und sie richtig zu interpretieren. Gerade in einer multimedialen Umwelt, so wie wir sie heute antreffen, spielt die Verarbeitung visueller Reize eine besondere Rolle. Ein Großteil der eintreffenden und zu verarbeitenden Informationen sind visueller Natur und müssen, um sich im Raum orientieren zu können oder eine Aufgabe zu meistern, verarbeitet und interpretiert werden. Um große Mengen an visuellen Eindrücken verarbeiten zu können, verlassen etwa eine Million Fasern jedes Auge, welche die Informationen kürzester Zeit befördern können (Koch & Tsuchiya, 2007). Nachdem unsere Aufnahmekapazität nicht unbegrenzt ist (vgl. Broadbent, 1958; Kinchla, 1992), ist es notwendig die Flut an Reizen nach Wichtigkeit zu selektieren (Carrasco, 2011). Für diesen Prozess sind komplexe neuronale Verrechnungen notwendig (für einen Überblick siehe Corbetta & Shulman, 2002). Unser visuelles System ist aber nicht für die gleichzeitige Aufnahme der gesamten, vor uns liegenden visuellen Umwelt ausgerichtet. Lediglich ein kleiner Teil der elektromagnetischen Wellen werden detailreich wahrgenommen (Henderson, 2003). Um den vor uns liegenden Bereich zur Gänze wahrnehmen zu können, bedarf es also mehrerer Augenbewegungen. Aus der Kombination dieser Fixationen, wird dann ein Gesamtbild konstruiert.

Doch was leitet unsere Augenbewegungen bzw. was zieht den Blick eines Menschen an? Im Allgemeinen unterscheiden wir zwischen reizgesteuerter (engl. „bottom up“) und zielgesteuerter (engl. „top down“) Aufmerksamkeit (z.B. Corbetta & Shulman, 2002; Theeuwes, 2010). Sofern wir nicht auf der Suche nach bestimmten Objekten oder visuellen Reizen sind, orientieren wir uns an hervorstechenden Reizeigenschaften der visuellen Umwelt (z.B. Car-

rasco, 2011; Treisman & Gelade, 1980). Suchen wir hingegen zielgerichtet nach bestimmten visuellen Reizen oder Reizkonstellationen, z.B. einer Person mit blauem Pullover und blondem Haar, so kommen top-down-Prozesse zur Anwendung (Theeuwes, 2010).

Das gesamte visuelle System ist aber noch weitaus komplexer. Bei der Betrachtung des visuellen Umfeldes werden auch Inhalte des Kurzzeitgedächtnisses, des Langzeitgedächtnisses, räumliche und semantische Informationen von anderen ähnlichen Szenen, sowie Ziele und Pläne miteinbezogen (Henderson, 2003). Viele dieser Inhalte sind bewusst nicht abrufbar, dennoch beeinflussen sie das visuelle System. Diese unbewussten Inhalte werden unter dem Begriff implizites Gedächtnis zusammengefasst (Schacter, Chiu & Ochsner, 1993). Um den Einfluss dieser impliziten Inhalte experimentell zu testen, verwendet man Primingaufgaben. In diesen Primingaufgaben stellte man fest, dass Menschen einen Großteil von betrachteten Bilder wiedererkennen (vgl. Haber, 1970; Shepard, 1967; Standing, 1973), auch wenn sie diese Tage zuvor betrachteten. (Standing, Conezio & Haber, 1970). Andererseits konnten Studien belegen, dass sogar größere Veränderungen von Bildern die alternierend gezeigt wurden, kaum erkannt werden (Simons & Rensink, 2005; Simons & Levin, 1997). Nachdem wir, wie schon oben beschrieben, unser visuelles Umfeld durch sprunghafte Augenbewegungen erkunden, ist es nachvollziehbar, dass Veränderungen von Bildteilen, die nicht betrachtet wurden, nicht erkannt werden. Wenn dem so ist sollten betrachtete Bildausschnitte eher und besser wiedererkannt werden als nicht betrachtete Teile eines Bildes.

Die vorliegende Arbeit geht dieser Fragestellung nach, indem das Blickverhalten beim Betrachten von natürlichen Szenen mitverfolgt wird und im Anschluss betrachtete und nicht betrachtete Bildausschnitte gezeigt werden. Ein Vergleich der Reaktionszeiten und der Wiedererkennungsleistungen von betrachteten und nicht betrachteten Bildteilen, soll Unterschiede in der Behaltensleistung und der visuellen Verarbeitung verdeutlichen. Zur theoretischen Begründung der Arbeit, sowie zum besseren Verständnis, werden im Folgenden die mit der Stu-

die in Verbindung stehenden und schon zu Beginn kurz erläuterten Paradigmen genauer beleuchtet und relevante Untersuchungen näher beschrieben.

Zu Beginn folgt nun eine kurze Einführung in die physiologische Beschaffenheit und Funktionsweise des Auges. Anschließend werde ich reizgesteuerte und zielgesteuerte Faktoren des Blickverhaltens näher beleuchten und schließlich werde ich auf den Einfluss von impliziten Gedächtnisinhalten auf die Verarbeitung von visuellen Reizen eingehen. Zusätzlich wird die in Verbindung mit dieser Arbeit stehende bisherige Forschung näher beleuchtet. Eine nähere Beschreibung der relevanten Experimente soll die Zusammenhänge illustrieren und dem besseren Verständnis dienen. Abschließend wird, aufbauend auf dem theoretischen Hintergrund, die Forschungsfrage der gegenwärtigen Studie präzisiert und in weiterer Folge werden die gefundenen Ergebnisse präsentiert und interpretiert. In den letzten Kapiteln erfolgt schließlich eine Diskussion der Ergebnisse, sowie ein Ausblick über weiterführende zukünftige Forschungsmöglichkeiten zur Abklärung der aus dieser Arbeit entstandenen Fragestellungen.

1.1 Selektive Aufmerksamkeit

Das Auge ist einer Unzahl an visuellen Reizen ausgesetzt, die nicht gleichzeitig verarbeitet werden können. Nur jener Teil der visuellen Umwelt, welcher auf einem kleinen Teil der Netzhaut, der Foeva Centralis, eintrifft, wird klar und deutlich wahrgenommen. Nur dieser winzige Bereich des Auges besitzt die höchste Auflösung, was Farbe und Raum betrifft (Henderson, 2003). Dort liegen die Zapfen, welche für Farb- und Raumwahrnehmung verantwortlich sind, in größter Dichte vor (Ansorge & Leder, 2011). Das Licht fällt zunächst durch die Iris und Linse auf die lichtempfindliche Netzhaut, um dann in einen Nervenimpuls transformiert und weitergeleitet zu werden. Etwa eine Million Fasern transportieren dann die umgewandelten Impulse (Koch & Tsuchiya, 2007) zum Nucleus geniculatum laterale (LGN) welche schließlich als Hauptverbindung im primären visuellen Cortex (V1) münden. Weitere

Verbindungen führen zum suprachiasmatischen Nucleus (SCN), sowie über die superior Colliculli (SC) zum posterioren Parietalcortex (PPC) (Abbildung 1).

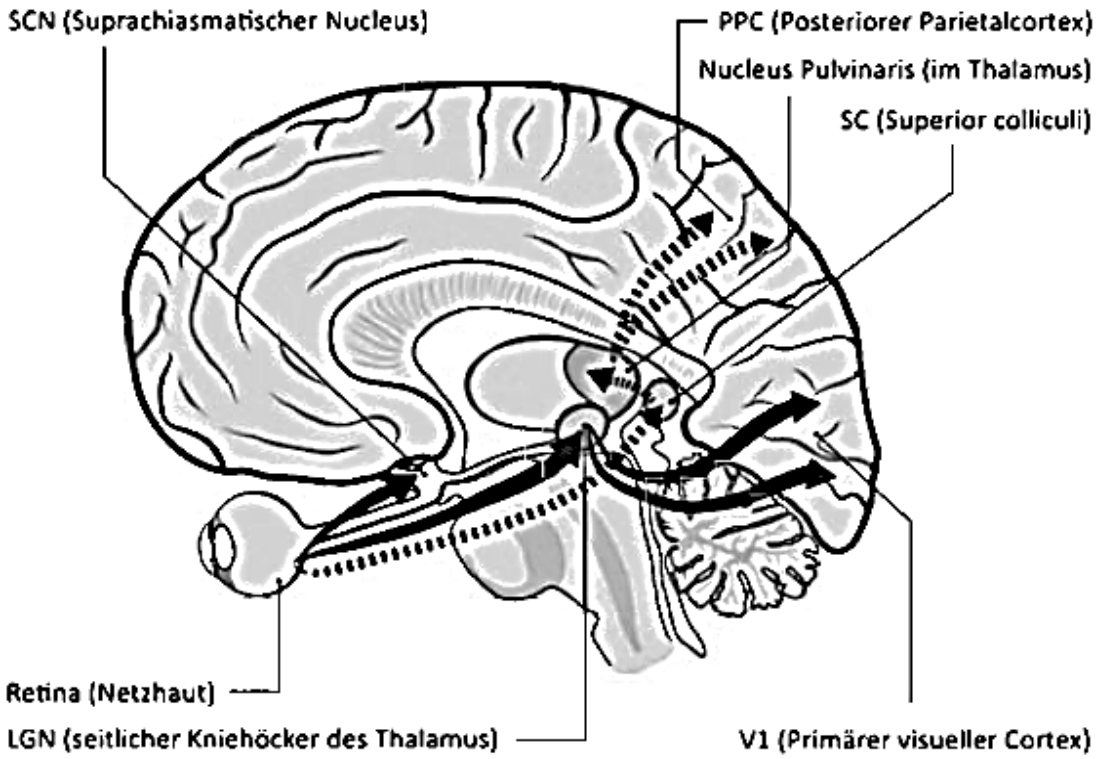


Abbildung 1: Schematische Darstellung der drei retino-zentralen Projektionen. (Ansorge, Leder, 2011)

Bei der Betrachtung der visuellen Umwelt verhält sich das menschliche Auge wie ein Scheinwerfer, der so eingestellt ist, dass er nur einen gewissen kreisrunden Spot hell erleuchtet. Dieser Spot entspricht etwa einem Sehwinkel von ca. 2° (Henderson, 2003). Um eine Szene vollständig visuell zu erfassen, bedarf es mehrerer solcher Fixationspunkte. Der Sprung von einem zu einem anderen Punkt nennt man Sakkade. Etwa dreimal pro Sekunde fixiert das Auge eine andere Stelle, durchschnittlich alle 300ms (Henderson, 2007). Studien zur Veränderungsblindheit (engl. „change blindness“), wo im Zuge einer Betrachtung sogar größere Veränderungen des Bildmaterials unbemerkt bleiben (Simons, 2000; Simons & Levin, 1997), zeigen eindrucksvoll, dass längst nicht alle Teile einer Szene wahrgenommen werden. Manche dieser Studien änderten das Bild während einer Sakkade (Currie, McConkie, Carlson-

Radvansky & Irwin, 2000; Henderson & Hollingworth, 1999, 2003; Hollingworth & Henderson, 2002; McConkie & Currie, 1996), während des Liedschlages (O'Regan, Deubel, Clark & Rensink, 2000), zwischen Präsentationsunterbrechungen (Aginsky & Tarr, 2000; Hollingworth & Henderson, 2000; Rensink, ORegan & Clark, 1997) oder auch bei der Betrachtung von bewegten Objekten (Levin & Simons, 1997). So verschieden die Versuchsdesigns auch sein mögen, die Ergebnisse sind einheitlich (Simons & Levin, 1997) und scheinen zu belegen, dass jene Bereiche, die nicht fixiert werden auch nicht wahrgenommen werden.

Welche Teile der visuellen Umwelt werden also bevorzugt betrachtet? Einerseits beeinflusst die besondere Beschaffenheit der visuellen Reize unsere Wahrnehmung, andererseits sind es unsere eigenen Absichten, Pläne oder Aufgaben. Visuelle Reize, welche den Blick auf sich ziehen, können starke Farben, Kontraste, besonders helle Bereiche oder aber auch bewegte Objekte sein, wie es Treisman und Gelade (1980) in ihrer Merkmalsintegrationstheorie (engl. „feature integration theory") aufzeigen. Angeregt von den Ergebnissen ihrer Studien unterscheiden sie dabei zwischen paralleler und serieller Suche. Die parallele Suche tritt auf, wenn es die Aufgabe verlangt einen Zielreiz zu suchen und dieser sich durch ein einzelnes spezifisches Merkmal von allen anderen Reizen unterscheidet. Die Suchzeiten sind dabei kurz und unabhängig von der Größe des Suchsets. Ist der Zielreiz durch eine Kombination von zwei Merkmalen, z.B. Farbe und Form definiert, wobei jedes dieser Merkmale auch von einem Distraktor getragen werden kann, sind die Suchzeiten länger und steigen mit zunehmender Anzahl an Distraktoren.

Ein Modell der reizgesteuerten Aufmerksamkeit, ist das Salienzmodell, welches auf den Arbeiten von Koch und Ullman (1985) aufbaut und mehrfach modifiziert wurde (Itti, Koch & Niebur, 1998; Parkhurst, Law & Niebur, 2002; Walther & Koch, 2006). In diesem Modell geht man grundsätzlich davon aus, dass der Blick durch lokale Merkmalskontraste gesteuert wird. Die beiden gingen in ihrem hypothetischen Modell von Neuronen im visuellen Cortex

aus, welche auf spezifische visuelle Reizeigenschaften reagieren. In frühen Stadien der Verarbeitung soll demnach die Verarbeitung parallel und retinotop erfolgen. Dabei sollen benachbarte Bereiche der Netzhaut die entsprechenden angrenzenden Neuronen in den visuell verarbeitenden Arealen aktivieren. Neuronen in diesen Arealen reagieren auf bestimmte Merkmalskontraste, die sich aus lokalen Merkmalsausprägungen der Dimensionen Farbe, Intensität und Orientierung und deren räumlicher Verteilung ergeben. Natürliche Bilder z.B. enthalten eine Fülle solcher Kontraste. Welche werden nun bevorzugt betrachtet bzw. wodurch wird die Reihenfolge der Sakkaden bestimmt? Koch und Ullman (1985) gehen von einem Alles-oder-Nichts-Prinzip (engl. „winner-take-all“) aus, wobei jener Bereich mit dem höchsten Salienzwert zuerst betrachtet wird, gefolgt von der Stelle mit dem zweithöchsten Wert usw. Der Mechanismus der Hemmung der Rückkehr (engl. „inhibition of return“) reduziert dabei die Wahrscheinlichkeit gleiche Merkmale kurzfristig noch einmal zu betrachten, wodurch eine effiziente Abtastung des Reizmaterials möglich wird (vgl. Klein, 2000).

Studien zur Ermittlung des Blickverhaltens bedienen sich der Methode des *Eye-trackings*. Dabei werden verschiedene visuelle Szenen oder Reize auf einem Bildschirm präsentiert und das Blickverhalten wird aufgezeichnet (Itti et al., 1998). Nachdem sich mehrere Versuchspersonen (Vpn) die Bilder angesehen haben und die Fixationspunkte aufgezeichnet wurden, werden die meist fixierten Punkte ermittelt. Um dieses Blickverhalten prognostizieren zu können, erarbeiteten Koch und Ullman (1985), auf ihren Modellannahmen basierend, erstmals sogenannte Salienzkarten, welche dann mit den gemessenen Fixationspunkten der Vpn verglichen werden konnten. Vereinfacht dargestellt werden dabei zunächst, wie beispielsweise im Modell von Itti et al. (1998), die Ausprägungen der drei Merkmalskanäle Farbe, Intensität und Orientierung ermittelt und in Merkmalskarten gespeichert. Für die Ermittlung der lokalen Merkmalskontraste der jeweiligen Karten werden spezielle Filter verwendet, welche die Kontraste verstärken und der Reaktionen des zentralen Nervensystems nachempfunden sein sol-

len. Ganz nach dem Alles-oder-Nichts-Prinzip wird, verglichen mit den umgebenden Reizintensitäten des gleichen Merkmals (z.B. Farbe), die stärkste Ausprägung hervorgehoben und die restlichen werden gehemmt. Diese merkmalsdimensionsspezifischen „Auffälligkeitskarten“ (engl. „conspicuity maps“) werden schließlich zu einer Salienzkarte kombiniert, die unabhängig von der Merkmalsdimension die Bereiche der höchsten Salienz im Bild widerspiegelt (Abbildung 2). Je heller die Bereiche sind, desto höher ist die Salienz und die Wahrscheinlichkeit, dass diese Bereiche betrachtet werden.

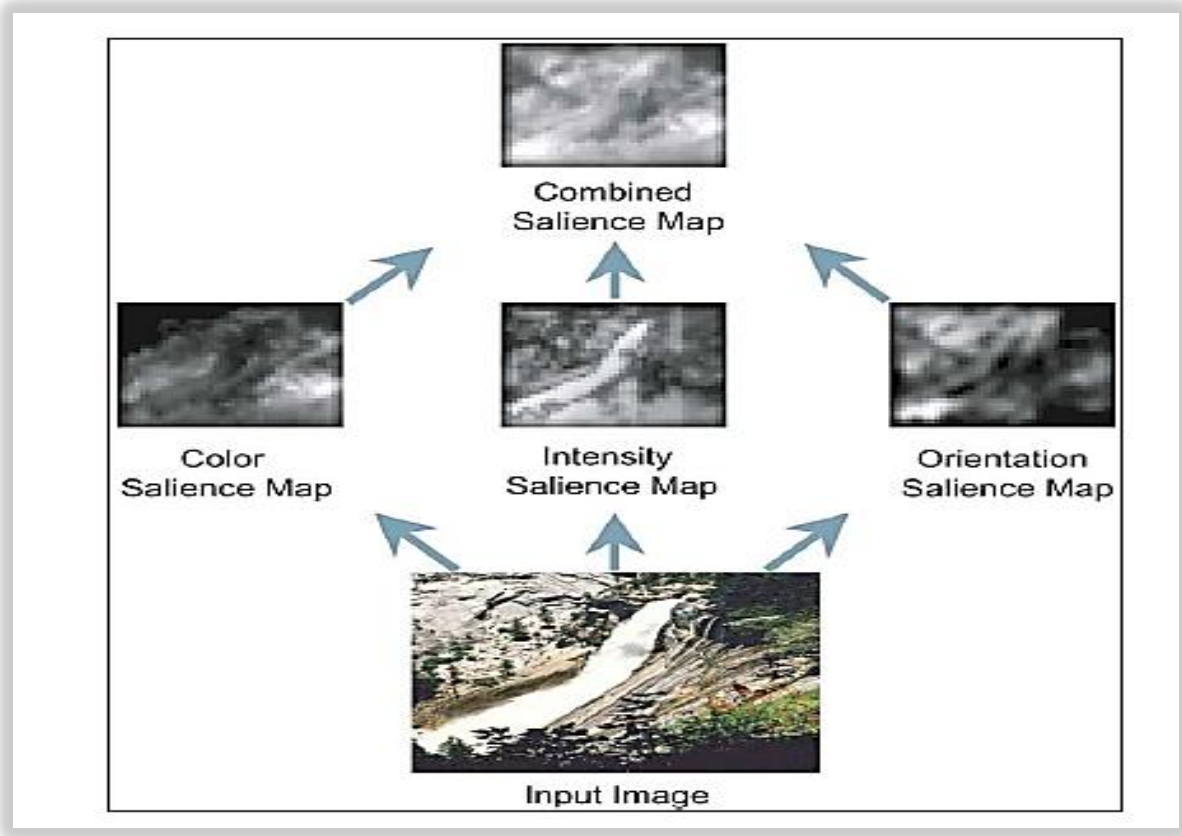


Abbildung 2: Salienzkarte (Parkhurst, Klinton & Niebur, 2002). Die drei Bereiche Farbe, Intensität und Orientierung werden zu einer kombinierten Salienzkarte zusammengefasst.

Sofern keine konkrete Aufgabenstellung, wie die Suche nach bestimmten Objekten oder Reizen gegeben ist, sind es Stellen mit starken Kontrasten oder Farben, die fixiert werden. Suchen wir hingegen nach einem bestimmten Reiz oder einer Konstellation von Reizen, so werden saliente Merkmale zweitrangig behandelt und das Suchverhalten verläuft zielgesteuert

nach spezifischen Reizmerkmalen (z.B. Einhäuser, Rutishauser & Koch, 2008; Henderson, 2003; Schütz, Braun & Gegenfurtner, 2011; Theeuwes, 2010).

Zielgesteuerte Aufmerksamkeit wird in der Literatur oft als ein völlig willentlicher Vorgang beschrieben (z.B. Theeuwes, 2010), bei dem nicht nur Absichten eine wesentliche Rolle spielen, sondern auch Pläne, Erwartungen und Vorwissen oder Erfahrungen (z.B. Ansorge, 2006; Ansorge, Leder, 2011; Schütz et al., 2011; Theeuwes, 2010). Bei der Suche nach einer verlorengegangenen Geldbörse werden beispielsweise hervorstechende Reizeigenschaften unwichtiger und jene der gesuchten Geldbörse treten in den Vordergrund, sodass die Suche erleichtert wird und nicht alle Inputs weiterverarbeitet werden müssen. Um diese Leistung vollbringen zu können, sind explizite Gedächtnisinhalte notwendig, welche als Muster eingesetzt und mit dem visuellen Input abgeglichen werden können.

Das Auffinden eines bestimmten, als Muster bereits abgespeicherten Reizes, kann durch einen Hinweis erleichtert werden. Posner (1980) untersuchte die willentliche Steuerung des Blickverhaltens, indem er vor der Präsentation eines Reizes einen Hinweisreiz (engl. „cue“) darbot, der die Lokalisation des Zielreizes verriet. Sofern der Hinweisreiz den Zielreiz richtig vorhersagte, resultierte dies in kürzeren Reaktionszeiten und einer verminderten Fehlerquote. Dieses Beispiel soll illustrieren wie das Vorwissen, in diesem Fall ein Hinweisreiz, nachfolgende Verarbeitung erleichtern kann. Der Hinweis scheint alle beteiligten Systeme dahingehend zu sensibilisieren, dass die Verarbeitung des betreffenden Reizes vereinfacht wird.

So bedeutend bewusste Pläne und Absichten für unsere Aufmerksamkeit sind, einen nicht unbeträchtlichen Einfluss haben unbewusste Gedächtnisinhalte. Diese impliziten Inhalte werden nicht oder nur teilweise wahrgenommen und bestehen aus vergangenen Erfahrungen, Tätigkeiten oder Wahrnehmungen und beeinflussen nachfolgende Verarbeitung unbewusst (Schacter et al., 1993). Um implizite Inhalte handelt es sich auch bei automatisierten Abläufen wie Autofahren oder Fußballspielen. Die Koordination des Körpers mit der Flugbahn eines

herannahenden Balles basiert auf erlernten unbewussten Abläufen. Würde man z.B. bei jedem Kopfball die Flugbahn berechnen, würde wohl selten ein Tor geschossen werden. Bei klassischen Studien in denen implizite Einflüsse erforscht werden, wird unter anderem ein Reiz dargeboten und untersucht ob dieser einen Einfluss auf folgende Darbietung selbigen Reizes bezüglich der Verarbeitung aufweist (z.B. Becker, 2008; Henderson, Pollatsek & Rayner, 1987; Kristjansson, 2006; Maljkovic & Nakayama, 1994). Viele dieser Studien verwenden Suchaufgaben und bedienen sich oft nur einzelner einfacher Reize die sich in Eigenschaften wie Farbe, Form, Ausrichtung, oder Größe unterscheiden. Ein Primingeffekt macht sich dabei durch verbesserte Reaktionszeiten und Leistungen von zuvor geprimten Reizen bemerkbar. Vorerfahrungen spielen demnach eine große Rolle bei der späteren Verarbeitung von visuellen Reizen. Wie stark der Einfluss von unbewussten Erfahrungen auf nachfolgende Verarbeitung ist beschreibt Theeuwes (2010). Auch wenn, entgegen der Intentionen der Betrachter, die Aufmerksamkeit auf andere Faktoren, wie emotionale Inhalte der Reize (z.B. verärgerte Gesichter) oder Vorerfahrungen gelenkt wird, wirkt sich dieser Einfluss auf die darauf folgende visuelle Selektion merklich aus. In Experimenten in denen der Einfluss von früherer Erfahrung auf die nachfolgende visuelle Verarbeitung und Verhaltensleistung geprüft wird (intertrial priming), wird ungewollt, bei der Suche nach einem neuen Objekt, weiterhin das Zielobjekt des vorangegangenen Durchgangs selektiert (Theeuwes, 2010). Insofern können uns Vorerfahrungen auch gegen unseren Willen beeinflussen.

1.2 Wahrnehmung und Gedächtnis

Eine der ersten, welche sich mit der Behaltenskapazität von visuellen Inhalten beschäftigten, waren Haber (1970), Shepard (1967) und Standing (1973). Shepard (1967) zeigte seinen Versuchspersonen nacheinander 600 einzelne Bilder für wenige Sekunden und prüfte anschließend ihre Wiedererkennungsleistungen, indem er Bildpaare präsentierte. Jedes Paar enthielt

ein zuvor gesehenes und ein neues Bild und die Vpn sollten entscheiden welches das gesehene war. Mit einer Wiedererkennungsrate von 98 %, konnte eine nahezu perfekte Wiedererkennungsleistung festgestellt werden. Etwas später führte Standing (1973) eine ähnliche Untersuchung durch, nur gab er seinen Vpn bis zu 10,000 Bilder kurz zum Einprägen vor. Sogar nach Tagen konnten die Testpersonen noch bis zu 90 % der betrachteten Bilder wiedererkennen. Der Autor schloss aus diesen Ergebnissen auf eine nahezu unbegrenzte Gedächtniskapazität für visuelle Inhalte.

Im Gegensatz dazu zeigen Studien zur Veränderungsblindheit, dass unserem visuellen System in bestimmten Situationen sehr wohl Kapazitäten fehlen (z.B. Currie, McConkie, Carlson-Radvansky & Irwin, 2000; McConkie & Currie, 1996; Simons, 2000; Simons & Levin, 1997). Details unserer visuellen Repräsentationen können scheinbar von einer Fixation zur nächsten nicht vollständig gespeichert werden (Simons & Levin, 1997). In den ersten Versuchen zur Veränderungsblindheit wurden Objekte (Egeth, 1966) oder nur ähnliche Buchstabenreihen (Taylor, 1976) miteinander verglichen, wobei Unterschiede festgestellt werden sollten. In neueren Experimenten werden Teilbereiche von Szenen während der Betrachtung verändert (z.B. Currie, McConkie, Carlson-Radvansky & Irwin, 2000; McConkie & Currie, 1996; Simons, 2000; Simons & Levin, 1997). Im Zuge der Betrachtung des Bildes wird dann zwischen einer Fixation und der anderen ein Detail oder auch ein Objekt verändert, ausgetauscht oder entfernt. Ein Großteil der Vpn konnten dabei nicht einmal größere Veränderungen registrieren, wodurch viele Forscher auf eine stark begrenzte Verarbeitungskapazität schlossen. Wie schon weiter oben erwähnt, wird ein Bild oder eine Szene nur punktuell erfasst, insofern wundert es nicht, wenn nicht alle Details einer Präsentation erfasst werden.

In diesem Zusammenhang entstanden zwei sehr ähnliche Theorien, welche mit den zuvor besprochenen Ergebnissen im Einklang stehen: die *coherence theory* (Rensink, 2000a, 2000b, 2000c, 2002) und die *object file theory of transsaccadic memory* (Irwin & Zelinsky, 2002;

Irwin, 1992). Beide beschreiben wie Objektmerkmale bei der Betrachtung von Szenen in dynamischer Weise und in Abhängigkeit der Verteilung der Aufmerksamkeit repräsentiert werden. Sie postulieren, dass beim Betrachten einer Szene ein sehr geringer Teil der visuellen Oberflächeninformation einer weiteren Verarbeitung unterzogen wird und lediglich semantische Informationen erhalten bleiben. Aufgrund der erheblichen Überschneidungen wurden beide schließlich als *visual transience-Theorien* zusammengefasst (Hollingworth, 2003; Hollingworth & Henderson, 2002). Diesem zusammenfassenden Ansatz nach, werden einzelne Merkmale von Objekten nur sehr kurz behalten. Repräsentationen natürlicher Szenen sind demnach sehr oberflächlich und ungenau, da sie während der Betrachtung einer Szene kaum detaillierte visuelle Objektmerkmale enthalten; langfristig werden sogar nur abstrakte, semantische Merkmale der Objekte, sowie Informationen zum thematischen Kerngehalt und der räumlichen Struktur der gesamten Szene gespeichert. Diesem Paradigma folgend sollten fixierte visuelle Inhalte einer Szene nicht oder sehr begrenzt im Gedächtnis gespeichert werden.

Studien, welche Aspekte des semantischen Gehalts einer Szene berücksichtigen, zeigen, dass Veränderungen welche dem Szenenkontext sehr nahe sind, leichter erkannt werden, als Merkmale die für den Kontext bedeutungslos sind (Kelley, Chun & Chua, 2003; Shore & Klein, 2000). Forschungen und Überlegungen diesbezüglich werden unter dem „contextual cuing"-Paradigma subsumiert. Darunter versteht man grundsätzlich die begünstigende visuelle Verarbeitung von Szenen, aufgrund von Erfahrungen visuell-räumlicher Invarianten (Chun, 2003). Im Laufe unserer reichhaltigen visuellen Erfahrungen lernen wir, dass unsere Umwelt gewisse Regularitäten aufweist und dass sogar Abweichungen einigermaßen vorhersehbar sind (Chun & Nakayama, 2000). So würden wir in einer Küche kein Motorrad vermuten und auf dem Flugfeld keinen Topf. Folgend dessen sollten Bereiche, welche mit dem Szenenkontext konsistent sind, kürzer betrachtet werden als inkonsistente Bereiche (Chun, 2003; Henderson, 2003; J. M. Henderson, Weeks & Hollingworth, 1999). Gewohnte Szenen-

Objekt-Konstellationen könnten demnach die Verarbeitung begünstigen. So zeigten z.B. Palmer (1975) und Biederman, Mezzanotte und Rabinowitz (1982), dass der Szenenkontext die Verarbeitung positiv beeinflusst. Gewohnte, valide Szenen wie z.B. eine Küche in der eine Brotdose stand, wurden schneller und besser wiedererkannt als unübliche Szenen. In einer weiteren Studie wurden Szenen mit jeweils zwei Veränderungen präsentiert. Eine hatte hohe Bedeutung für den Szenenkontext, die andere eine geringere Bedeutung. Die bedeutungsvollere Veränderung wurde schneller und besser erkannt als jene, die weniger wichtig für den Kontext war. Wurden die Bilder umgedreht, so verringerte sich der Effekt signifikant (Kelley et al., 2003). Henderson & Hollingworth (1999) beschäftigten sich in ihrer Studie zur Veränderungsblindheit mit der Verweildauer und Dichte der Augenfixationen. In ihren Studien konnten sie einen Zusammenhang zwischen Fixationsdichte und -länge und visuell sowie semantisch informative Bildregionen feststellen. Die Forscher folgerten aus ihren Ergebnissen, dass die ersten Fixationen mit den visuellen Merkmalen und den globalen, semantischen Charakteristika zu tun haben könnten und die folgenden Fixationen mit der visuellen und semantischen Bedeutung der Region. Die visuelle Informationsbeschaffung bei der Betrachtung inkonsistenter Bildbereiche beeinflusst demnach die Fixationsdauer der jeweiligen Bildregion. In den Studien von Henderson et al. (1999) variierten die Fixationszeiten je nach Aufgabenstellung. Sollten die Bilder memoriert werden, so wurde eine längere Verweildauer gemessen als bei freien Betrachtungsaufgaben.

Die Szenenbedeutung allein kann aber die hohe Wiedererkennungsrate von Bildern nicht vollständig erklären. Konkle, Brady, Alvarez und Oliva (2010) konnten zeigen, dass die hohen Behaltensleistungen auch bei vielen vom Kerngehalt her sehr ähnlichen Bildszenen verhältnismäßig stabil bleiben. Chun (2003), schlägt diesbezüglich ein Zweiwegemodell der Szenenverarbeitung vor. Demnach soll zunächst eine globale, räumliche Orientierung stattfinden, welche nachfolgende Augenabtastung beeinflusst. Beim Scannen der Umgebung durch

die einzelnen Fixationen, dienen die betrachteten Bereiche einerseits der Szenenidentifikation und andererseits ermöglichen sie eine erleichterte Verarbeitung im Sinne von Priming, also eine erleichterte Verarbeitung durch entsprechende, implizite Inhalte. Sanocki und Epstein (1997) führten diesbezüglich Versuche durch, indem sie den Informationsgehalt einer natürlichen Szene variierten und anschließend den räumlichen Abstand zweier neu hinzugefügter Objekte (z.B. Sessel) abfragten. Vpn, welche vollständige Bilder der Szene zuvor gesehen hatten, antworteten wesentlich schneller als jene welche nur einfache Strichzeichnungen der gleichen Szene betrachtet haben und diese wiesen wiederum kürzere Reaktionszeiten auf als jene, welche einen leeren Rahmen zu sehen bekamen. Die Autoren folgerten daraus, dass die zuvor dargebotenen Informationen unbewusst in den Verarbeitungsprozess miteinbezogen wurden und somit die räumliche Verrechnung vereinfacht wurde.

Bartram (1974) stellte in seiner Primingstudie fest, dass gleiche oder sehr ähnliche Objekte bei der wiederholten Darbietung den größten Effekt im Vergleich zu gleichnamigen aber verschiedenen Objekten erzielen. Bezogen auf den Informationsgehalt, konnten Snodgrass und Feenan (1990) zeigen, dass fragmentiert dargestellte Objekte eine gute Wiedererkennungsrate aufweisen, sofern sie genügend Informationen liefern. Dass sich der Effekt nicht nur auf direkt nachfolgende Testdurchgänge bezieht sondern noch mehrere Testphasen weiterbesteht, konnten Maljkovic und Nakayama (1994, 1996) in ihren Forschungsarbeiten beobachten.

Bei vielen Primingstudien ist jedoch eine Konfundierung mit expliziten Gedächtnisinhalten nicht auszuschließen. Einen Versuch explizite Inhalte konstant zu halten, machten Pertzov, Zohary und Avidan (2009). Um eine Vermischung explizierter Inhalte zu vermeiden, teilten sie ihre Studie zur Veränderungserkennung (engl. „change detection“) in einen expliziten und impliziten Teil, indem sie zu Beginn des Versuches einen Prime oder einen Distraktor darboten und am Ende die Wiedererkennungsleistung testeten. Nachdem die Vpn den fragmentierten Prime/Distraktor zu Beginn wiedererkannten, folgten Durchgänge mit Szenen, welche

alternierend mit und ohne Veränderung gezeigt wurden, wobei der zuvor präsentierte Prime/Distraktor integriert wurde. Sofern der Prime zu Beginn gezeigt wurde, war dies auch der veränderte Zielreiz, ansonsten wurde ein anderes Objekt verändert und der Distraktor wurde integriert. Die Vpn sollten durch Tastendruck die bemerkte Veränderung der Szene bekannt geben und anschließend in einem Raster den genauen Ort der Veränderung eintragen. Abschließend wurden wieder zwei Objekte präsentiert und die Vpn sollten entscheiden welcher der Beiden anfänglich als Prime/Distraktor präsentiert wurde. Wurde der Prime/Distraktor wiedererkannt, so deklarierte man diesen Durchgang als explizit, konnte er nicht wiedererkannt werden, so wurde er als implizit bewertet. Zusätzlich wurde das Blickverhalten der Probanden aufgezeichnet. Auch bei den impliziten Durchgängen, in denen der anfängliche Prime nicht wiedererkannt wurde, konnten kürzere Reaktionszeiten bei der Veränderungsaufgabe gemessen werden. Wurde der Zielreiz (Prime) anfänglich präsentiert, so wurde dieser auch öfter in der nachfolgenden Veränderungsaufgabe fokussiert als wenn ein Distraktor anfänglich gezeigt wurde. Die Autoren folgerten, dass durch den Prime unbewusste Repräsentationen abgelegt werden, welche zukünftig bei der Betrachtung von Szenen als Abgleich dienen könnten.

Eine Theorie, welche den Abgleich von fixierten Muster einer Szene als Grundlage für spätere Wiedererkennung der gleichen Szene versteht, ist die *scanpath Theorie* (vgl. Noton & Stark, 1971; Stark & Ellis, 1981). Sie versteht das Wiedererkennen von Szenen als sensumotorischen Prozess. Beim Abtasten einer Szene durch foveale Fixierungen des Auges, werden nicht nur die einzelnen Fixierungen als Muster im Gedächtnis abgespeichert, sondern auch die dazugehörigen motorischen Augenbewegungen. Wiedererkennung ist dieser Theorie zufolge, abhängig von der Wiederholung dieses sensumotorischen Musters, wobei die Abfolge der Fixationen und Sakkaden gespeichert und mit einer späteren Betrachtung verglichen wird. So

erfolgt eine exakte Wiedererkennung einer Szene durch das gleiche Muster und die gleiche Abfolge der Fixationen bei ein und demselben Bild (vgl. Brandt & Stark, 1997).

In Blickverfolgungsstudien versuchen Forscher oft herauszufinden welche Bereiche einer Szene betrachtet werden, welche physikalischen Eigenschaften diese Bereiche aufweisen (z.B. Van der Linde, Rajashekar, Bovik & Cormack, 2009) und in welchem Zusammenhang diese Bereiche mit der visuellen Verarbeitung stehen. Doch nicht nur fixierte Bereiche einer Szene können Gegenstand unserer Aufmerksamkeit sein sondern auch Bereiche einer Szene, die wir nicht fixieren. Darüber hinaus können auch kognitive Prozesse unsere gesamte Wahrnehmung in den Bann ziehen und das, obwohl die Augen auf ein bestimmtes Objekt gerichtet bleiben. Es ist z.B. möglich einen Text zu lesen und danach nicht mehr zu wissen was man eigentlich gelesen hat, weil man in Gedanken gerade woanders war. Posner (1980) untersuchte dieses Phänomen der verdeckten Aufmerksamkeit (engl. „covert attention“). Verdeckte Aufmerksamkeit bezeichnet laut Posner die Zuwendung der Aufmerksamkeit auf einen bestimmten Ort, ohne dabei die Augen oder den Kopf zu bewegen. In den klassischen Experimenten dazu wurden Hinweisreize zur Verlagerung der Aufmerksamkeit gegeben, wobei die Augen nicht bewegt werden durften. Es stellte sich heraus, dass die Aufmerksamkeit unabhängig von den Augenbewegungen verschoben werden konnte. Aus den Ergebnissen dieser Studien schloss der Forscher auf eine Trennung von Aufmerksamkeit und Blickverhalten. Posner, Snyder und Davidson (1980) gingen folglich von einem zusätzlichen Aufmerksamkeitssystem aus, dass mit dem visuellen System interagiert aber nicht gekoppelt ist und folglich getrennt agieren kann.

Fraglich ist nun, ob visuelle Inhalte, welche bewusst nicht registriert wurden, implizit verarbeitet werden und somit die darauf folgende Verarbeitung beeinflussen. Ob nicht fixierte Bereiche einer Szene im Gedächtnis kurz oder auch längerfristig abgelegt werden zeigen zwei ähnliche Studien von Irwin und Zelinsky (2002) und Zelinsky und Loschky (2005). In diesen

Experimenten wurde die Merkleistung für fixierte und nicht fixierte Objekte untersucht. In beiden Designs wurden mehrere Objekte zur freien Betrachtung präsentiert und kurz danach abgefragt. Bei der Studie von Irwin und Zelinsky (2002) war die Anzahl der Fixationen begrenzt, es wurde aber von Durchgang zu Durchgang zwischen einer und 15 Fixationsmöglichkeiten variiert. Sobald das Limit erreicht wurde, verschwand das Bild und das schwarze Testdisplay mit den gleichen aber anders sortierten Gegenständen wurde gezeigt. Die Vpn sollten jenen Gegenstand zeigen, welcher an einer bestimmten Position des zuvor präsentierten Sets lag. Durch die aufgezeichneten Blickbewegungen konnte die Fixationsreihenfolge berechnet und in Beziehung zur Merkleistung gesetzt werden. Zusammengefasst konnten die Forscher eine erhöhte Wiedererkennungsleistung von 80 % bis 90 % für fixierte Objekte feststellen, aber nur wenn das Zielobjekt nicht mehr als drei Fixationen zurücklag. Wurden nach dem getesteten Zielobjekt mehr als drei weitere Stellen des Bildes fixiert, so fiel die Wiedererkennungsleistung auf rund 65 % was der Wiedererkennungsleistung von nicht fixierten Objekten (59 %) sehr nahe kam. Nicht fixierte Objekte konnten bei bis zu neun Fixationsmöglichkeiten zu durchschnittlich etwas mehr als 50 % wiedererkannt werden. Konnten die Vpn 15mal fixieren, wurde kein signifikanter Unterschied mehr zwischen fixierten und nicht fixierten Zielobjekten festgestellt. Ab einer gewissen Anzahl von Fixationen in einer Szene, landen die Blickpunkte sehr nahe bei den Objekten, auch wenn diese nicht direkt fixiert werden (Irwin & Zelinsky, 2002). Folglich können die fovealen wahrgenommenen Bereiche so dicht werden, dass es kaum Stellen gibt, die nicht foveal (oder parafoveal) wahrgenommen werden. Wie groß der wahrgenommene Bereich rund um den durchschnittlichen Mittelpunkt der Fixation ist, zeigen die Ergebnisse von Hollingworth, Schrock und Henderson (2001) und Nelson und Loftus (1980), welche dem fixierten Bereich eine bedeutende Rolle bei Veränderungsaufgaben zuweisen. Nelson und Loftus (1980) untersuchten beispielsweise, in welchem Abstand ein Objekt fixiert werden muss, um eine Veränderung bei einem bestimmten Objekt

zu erkennen. Bei der Betrachtung von Szenen zeichneten sie die Blickbewegungen ihrer Probanden auf und unterzogen sie danach einem Test mit zwei Antwortalternativen. Ein Vergleich zwischen Erkennungsleistung und Fixationsnähe zeigte eine Abhängigkeit der Wiedererkennungsleistung von der Nähe der tatsächlichen Fixation. Die besten Gedächtnisleistungen wurden bei Fixationen bis zu einem Blickwinkel von 1.8° festgestellt, ab dieser Grenze fielen die Leistungen rapide ab. Nelson und Loftus (1980) ermittelten in ihren Untersuchungen schließlich einen Radius von 3-4° in dem Objekte in natürlichen Szenen wahrgenommen werden können. Die hier besprochenen Untersuchungen beziehen sich auf die Erkennung von Veränderungen während oder kurz nach der Betrachtung von Bildern. Sie sprechen also das Kurzzeitgedächtnis an. Im Folgenden soll nun auf die Wiedererkennungsleistung von fixierten Bildbereichen natürlicher Szenen über längere Zeit eingegangen werden. Werden fixierte Bereiche auch im Langzeitgedächtnis gespeichert?

Die ersten Studien zur Behaltensleistung von natürlichen Bildern überprüften, ob gesamte Bilder über längere Zeit behalten werden können, unabhängig davon welche Bereiche fixiert wurden (z.B. Standing, 1973). Wie schon zuvor besprochen, kann oft nicht eindeutig gesagt werden, ob einzelne visuelle Szeneninhalte oder die Kernbedeutung der Szene memoriert wird. Vereinzelte Studien untersuchten deshalb das Langzeitgedächtnis für einzelne Objektmerkmale innerhalb einer Szene. Dabei wurde die Behaltensleistung mit Hilfe von Distraktoren, welche sich in einem bestimmten Merkmal unterschieden (z.B. Friedman, 1979; Goodman, 1980; Mandler & Johnson, 1976; Mandler & Parker, 1976; Pezdek, Whetstone, Reynolds, Askari & Dougherty, 1989; Salmaso, Baroni, Job & Peron, 1983) oder mit Distraktoren welche sich durch ihre Orientierung unterschieden (Mandler & Parker, 1976; Mandler & Ritchey, 1977), gemessen. Sogar nach einem Tag konnten bis zu 65 % der Veränderungen erkannt werden (Mandler & Ritchey, 1977). Viele dieser anfänglichen Studien verwendeten jedoch nur einfache schematische schwarz-weiß-Zeichnungen und die Augenbewegungen

wurden nur in der Studie von Friedman (1979) gemessen. Eine neuere Studie gibt es diesbezüglich von Hollingworth und Henderson (2002). Sie untersuchten den Einfluss der Fixationslokalität auf die Entdeckung von Veränderungen an einzelnen Objekten während der Betrachtung von gewohnten Szenen. Die Vpn bekamen zunächst einige Szenen auf einem Monitor zu sehen, wobei sich während der Betrachtung des Bildes ein Objekt der Szene änderte, entweder bevor oder nachdem es fokussiert wurde. Dabei sollten die Vpn eine Taste drücken, sobald eine Veränderung wahrgenommen wurde. Die Veränderung wurde erst dann vollzogen, wenn der Blick weit genug vom veränderten Objekt entfernt lag. Nachdem alle Szenen betrachtet wurden, sollten die Vpn zwischen dem originalen und dem veränderten Bild entscheiden, je nachdem welches dem zuvor betrachteten Bild entsprach. In Abhängigkeit davon, wann das Bild betrachtet und getestet wurde, variierte die Zeit in der das Objekt behalten werden musste zwischen 5 und 30 Minuten. Die Forscher unterteilen in zwei verschiedenen Arten von Veränderungen. Zum einem wurde die Type (z.B. ein Block durch eine Diskette) verändert und zum anderen die Ausführung (z.B. ein Notizblock durch einen Ringblock). Wurde das Zielobjekt vor der Veränderung betrachtet, konnte in allen drei Experimenten, welche sich durch die Art der Veränderung (Rotation, Type, Ausführung) und den Zeitpunkt der Testphase im Wesentlichen unterschieden, eine durchschnittliche Erkennungsleistung der Veränderung von 80 % festgestellt werden. Diese Leistungen waren für einen Zeitraum von 5 bis 30 Minuten nachweislich stabil, auch wenn in dieser Zeit noch mehrere Objekte und sogar Szenen nach der Fixation des Zielobjektes betrachtet wurden. Diese Ergebnisse decken sich mit jenen von Standing et al. (1970) für ganze Szenen und jenen von Friedman (1979) und Parker (1978) für einzelne Objekte (Hollingworth & Henderson, 2002). Der wesentliche Unterschied zu vorangegangenen Untersuchungen zur Veränderungsblindheit besteht den Autoren zufolge darin, dass das veränderte Objekt nachweislich vor der Veränderung fixiert wurde. Einen wesentlichen Einfluss hatte auch die Fixationsdauer und –dichte bezüglich des Zielobjektes auf

die Erkennungsleistung. Einen Einfluss der Anzahl der nach dem Zielobjekt folgenden Fixationen konnte nicht einheitlich nachgewiesen werden. Aufgrund dieser und in anderen Experimenten gewonnen Einsichten entstand die *visual memory theory of scene representation* (Henderson & Hollingworth, 2003; Hollingworth & Henderson, 2002; Hollingworth, Williams & Henderson, 2001; Hollingworth, 2003, 2004, 2005). Dieser Theorie zufolge werden Objektmerkmale während der Betrachtung einer Szene nicht nur kurzfristig im Arbeitsgedächtnis gespeichert, sondern auch im Langzeitgedächtnis. Dabei werden nicht nur visuelle und semantische Informationen von Objekten in beiden Gedächtnissystemen abgelegt, sondern auch von der gesamten Szene, sowie räumliche Informationen des Szenenlayouts. Sofern Objektrepräsentationen im Arbeits- oder Langzeitgedächtnis bereits aufgebaut wurden, reicht es dann die Aufmerksamkeit auf einen beliebigen Bereich der Szene zu lenken, um den Zugriff auf die korrespondierenden Repräsentationen im Arbeits- oder Langzeitgedächtnis auszulösen. Der umgebende Szenenkontext dient dabei als Hinweisreiz für die Aktivierung der Objektrepräsentationen. Hollingworth und Hendersons *visual memory theory* kann als Gegenvorschlag zu den *visual transience*- Theorien gesehen werden: Visuelle Objektmerkmale werden nicht nur kurzfristig und begrenzt gespeichert, sondern bleiben dem Langzeitgedächtnis auch längerfristig erhalten. Allen Theorien gemeinsam, ist jedoch die Voraussetzung der visuellen Aufmerksamkeit auf das betreffende Objekt beim Aufbau von Gedächtnisrepräsentationen von Szenen.

In einer Untersuchung, deren Design der vorliegenden Studie ähnlich ist, untersuchten Van der Linde, Rajashekar, Bovik und Cormack (2009) die Wiedererkennungsleistung von Bildbereichen natürlicher Szenen. Sie präsentierten ihren Vpn eine Reihe von natürlichen Szenen in schwarz-weiß, die frei betrachtet werden konnten, zeichneten die Blickbewegungen auf und testeten nach jedem Bild die Wiedererkennungsleistung, indem ein fixierter Bereich der Szene und ein kleiner Ausschnitt aus einer anderen nicht präsentierten Szene zur Auswahl standen.

Die Vpn sollten entscheiden, welcher Ausschnitt zuvor betrachtet wurde. Im Gegensatz zu den hier zuletzt besprochenen Studien, konnte eine Wiedererkennungsrate von 68 % erzielt werden. Da der Wiedererkennungstest sofort nach der Präsentation des Bildes stattfand, wurde diese Leistung hauptsächlich mithilfe des Kurzzeitgedächtnisses vollbracht.

Zusammenfassend kann man von einer sehr guten Wiedererkennungsleistung von gesamten Szenen ausgehen. Bezüglich visuell fixierter Teilbereiche natürlicher Szenen ergeben sich teilweise widersprüchliche Forschungsergebnisse. Einerseits konnte gezeigt werden, dass sogar größere Veränderungen während der Betrachtung einer Szene nicht oder kaum bemerkt wurden (Simons, 2000; Simons & Levin, 1997), andererseits gibt es Hinweise, dass diese Blindheit für Veränderungen durch mangelnde Fixationen in den betreffenden Bereichen zustande komm (Hollingworth et al., 2001; Nelson & Loftus, 1980). Die Zielobjekte wurden also nicht oder in zu großer Entfernung fixiert. Weiters könnten der Szenenkontext sowie der Kerngehalt einer Szene die Wiedererkennungsleistung von fixierten Bereichen begünstigen, wie die oben angeführten Studien zeigen. Inwiefern fixierte Bereiche von natürlichen, komplexen Szenen, wie wir sie in unserer Umwelt vorfinden, losgelöst vom Szenenkontext längerfristig im Gedächtnis behalten werden, wurde bislang noch kaum erforscht. Zum einem soll in dieser hier vorliegenden Arbeit dieser Frage nachgegangen werden, zum anderen sollen zusätzliche Erkenntnisse bezüglich der langfristigen Wiedererkennungsleistung von fixierten und nicht fixierten Bereichen komplexer, natürlicher Szenen zu bestehenden Forschungsarbeiten gewonnen werden.

2 Fragestellung

Das Hauptanliegen der im Rahmen der vorliegenden Arbeit durchgeführten Untersuchung war es zu erforschen, ob fixierte Bereiche komplexer natürlicher Szenen, wie wir sie im Alltag vorfinden, langfristig und losgelöst vom Szenenkontext erinnert werden können und inwiefern vergleichsweise nicht fixierte Bereiche der gleichen Szene im Gedächtnis längerfristig repräsentiert werden. Dem aktuellen Forschungsstand zufolge bestehen diesbezüglich zwei teilweise widerstrebende Theorien. Die *transience-Theorien* (Rensink, 2000a, 2000b, 2000c, 2002; Irwin, 1992; Irwin & Zelinsky, 2002) gehen von einem kurzfristigen, begrenzten Speicher aus, welcher sehr fragmentarisch Objektmerkmale nur für kurze Zeit speichert. Dieser Theorie zufolge sollten nur wenige fixierte Bereiche (2-4) für kurze Zeit behalten werden. Der *visual memory theory of scene representation* (Henderson & Hollingworth, 2003; Hollingworth, 2003, 2004, 2005; Hollingworth & Henderson, 2002; Hollingworth, Williams & Henderson, 2001) zufolge, werden Repräsentationen der Merkmale fixierter Bereiche nicht nur im Arbeitsgedächtnis, sondern auch im Langzeitgedächtnis gespeichert. Wird eine Szene wiederholt betrachtet, so lösen die fixierten Bereiche die Aktivierung korrespondierender bereits gespeicherter Inhalte aus, wodurch die Szene schneller und besser erkannt wird. Diese Theorie wird unter anderem durch die Versuche zu *contextual cueing* gestützt. Diese konnten eine bessere und schnellere räumliche und detaillierte Verarbeitung von geläufigen, im Vergleich mit ungewohnten oder veränderten Szenen, nachweisen. Folgend dessen sollten Bereiche, welche mit dem Szenenkontext konsistent sind, kürzer betrachtet werden als inkonsistente Bereiche (Chun, 2003; Henderson, 2003; Henderson, Weeks & Hollingworth, 1999). Wird die gesamte Szene wiederholt betrachtet, erleichtert dies die Verarbeitung der gleichen Szene. Demzufolge sollte ein Wegfall des Szenenkontextes die Verarbeitung beeinträchtigen. Inwie-

fern fixierte Bereiche natürlicher Szenen losgelöst vom Szenenkontext wiedererkannt werden, wurde bisher nicht hinreichend erforscht.

In der vorliegenden Studie wurden zunächst einige natürliche Szenen des Alltags auf einem Bildschirm präsentiert. Nachdem alle Bilder betrachtet wurden, wurde die Wiedererkennungsleistung für die am längsten fixierten Teilbereiche im Vergleich zu nicht fixierten, salienten Teilbereichen getestet. Zur Kontrolle wurden zusätzlich Ausschnitte aus neuen, nicht gelernten Bildern, in der Testphase präsentiert. Die neuen, nicht gezeigten Bildteile wurden in hoch saliente und zufällige Bereiche unterteilt. Ausgehend von den oben angeführten Überlegungen und dem aktuellen Kenntnisstand wurden folgende Hypothesen untersucht:

- Wenn die visuelle Verarbeitung fixierter Bereiche natürlicher Szenen gemäß der *transience-Theorien* verläuft, sollten die fixierten Bildbereiche nicht ins Langzeitgedächtnis eingehen. Es sollten sich keine Unterschiede zwischen fixierten, nicht fixierten und neuen Teilbereichen hinsichtlich der Wiedererkennungsleistung und Reaktionsgeschwindigkeit ergeben. Die Wiedererkennungsrate sollte die Zufallswahrscheinlichkeit von .5 nicht signifikant übersteigen.
- Wenn die visuelle Verarbeitung natürlicher Szenen gemäß der *visual memory theory of scene representation* verläuft, sollten sich Unterschiede zwischen fixierten und nicht fixierten Bereichen der Bilder und zwischen präsentierten und nicht präsentierten Bildbereichen ergeben.
    - Präsentierte Bildbereiche sollten von nicht präsentierten Bildbereichen unterschieden werden können.
    - Fixierte Bildbereiche sollten schneller und besser wiedererkannt werden als nicht fixierte.
    - Nicht präsentierte Bildbereiche sollten mit ähnlich hoher Leistung zurückgewiesen werden wie präsentierte Bildbereiche erkannt werden.

- Bei den nicht präsentierten Bildbereichen sollten sich keine Unterschiede hinsichtlich der Reaktionszeiten und der Anzahl an Zurückweisungen zwischen hoch salienten und zufälligen Bildteilen ergeben.

3 Methode

3.1 Versuchspersonen, Apparatur und Reize

Versuchspersonen. An der Studie nahmen 24 Vpn teil, 15 davon waren weiblich, das Durchschnittsalter betrug 25 Jahre *(SD*=4). Die Vpn waren einerseits Studierende der Fakultät für Psychologie der Universität Wien, welche freiwillig und unentgeltlich teilnahmen oder durch die Teilnahme eine Teilleistung im Zuge einer Lehrveranstaltung erbrachten. Weitere Teilnehmer stammten aus dem eigenen Bekanntenkreis. Alle Vpn verfügten über normale oder korrigierte Sehfähigkeit, was durch einen standardisierten Visus-Test vor dem jeweiligen Versuch erhoben wurde. Bevor das eigentliche Experiment begann, wurden alle Vpn über den Ablauf und rechtliche Aspekte der Teilnahme informiert und unterschrieben folgend dessen eine Einverständniserklärung.

Apparatur. Die Bilder wurden auf einem 19“ CRT Farbmonitor der Marke Sony (Multiscan G400) mit einer Auflösung von 800 × 600 Pixel und einer Wiederholungsfrequenz von 100 Hz gezeigt. Während der ersten Betrachtungsphase der Bilder wurden die Blickbewegungen von einem Blickverfolger des Typs EyeLink 1000 Desktop Mount (SR Research, Mississauga, Ontario , Canada) mit einer Abtastungsfrequenz von 1000 Hz, aufgezeichnet. Das Gerät befand sich unterhalb des Bildschirmes und zeichnete die Bewegungen des dominanten Auges auf. Die Experimentalprozedur zur Darbietung der Bilder und die Reaktionsaufzeichnung wurde in MATLAB mit der Psychophysics Toolbox implementiert (Brainard, 1997; Pelli, 1997) und auf einem Standard-PC unter Windows XP durchgeführt. Um eine konstante Distanz zu gewährleisten und Bewegungen zu vermeiden, wurde eine Kopf- und Kinnstütze im Abstand von 72 cm angebracht, wodurch eine Fläche von 28° × 21° sichtbar wurde. In der

Testphase sollten die Vpn mittels Tastendruck entscheiden, ob der jeweilige Ausschnitt aus einem bekannten (d.h. bereits gesehenen) oder aus einem neuen (d.h. in der Lernphase nicht präsentierten) Bild stammt. Realisiert wurde die Antwort mit einer herkömmlichen USB-Tastatur, wobei die Tasten „F“ und „J“ für ja und nein standen und mit dem linken und rechten Zeigefinger bedient wurden.

Reize. Das Reizset bestand aus 60 Photographien natürlicher Szenen (Abbildung 3). Die Bilder bestanden aus Außenaufnahmen des Alltags, wobei sorgfältig darauf geachtet wurde, dass keine ungewöhnlichen oder seltsamen Objekte, bekannte Orte oder spezifische Personen im Vordergrund zu sehen waren.

Abbildung 3: Verwendete natürliche Bildszenen der Lernphase, in der die Blickbewegungen aufgezeichnet wurden.

Um eine gründliche Betrachtung der Bilder während der Lernphase zu sichern, wurden die Szenen so ausgewählt, dass sie ausreichend heterogene Objekte enthielten. Die Bilder der Lernphase wurden mit einer Auflösung von 800 × 600 Pixel × 32 Bit × 100 Hz dargeboten.

Nachdem alle Bilder in der Lernphase betrachtet wurden und der Computer die Fixationspunkte verrechnet hatte, wurden in der Testphase kleine Bildteile mit einer Auflösung von 100 × 100 Pixel und einer sichtbaren Fläche von 3.5° × 3.5° jeweils in der Mitte des Bildschirms (d.h. am Ort der zentralen Fixation) präsentiert (Abbildung 4).

Abbildung 4: Bildbereiche der Testphase nachdem alle Bilder der Lernphase betrachtet wurden. (A) Bildbereiche der am längsten fixierten Bereiche und nicht fixierte, saliente Bereiche wurden aus präsentierten Bildern der Lernphase ausgewählt. (B) Aus nicht präsentierten Bildern, wurden saliente und zufällige Bildbereiche ausgewählt. (C) Präsentiert/fixierte, präsentiert/nicht fixierte, nicht präsentiert/saliente und nicht präsentiert/zufällige Bildbereiche der Testphase

Dabei wurde einerseits zwischen fixierten und nicht fixierten Bereichen der präsentierten Bilder der Lernphase unterschieden und andererseits zwischen präsentierter und nicht präsentierter Bildteile. Teilbereiche von Bildern, die präsentiert, aber nicht betrachtet wurden, wurden bewusst so ausgewählt, dass sie saliente Merkmale enthielten. Die nicht präsentierten neuen Bildbereiche wurden jeweils zur Hälfte in hoch saliente und zufällige unterteilt

3.2 Ablauf und Design

Alle Teilnehmer unterschrieben zunächst eine Einverständniserklärung und wurden anschließend auf Fehlsichtigkeit, Farbsehen und ihr dominantes Auge getestet. Anschließend wurde Alter, Händigkeit, Geschlecht und Sehbehelfe erfasst und der Blickverfolger wurde auf das dominante Auge justiert und kalibriert. Jede Vp nahm dann vor dem Bildschirm Platz, der Raum wurde verdunkelt und ein Hintergrundlicht eingeschaltet um Spiegelungen, welche die Messungen des Blickverfolgers stören könnten, zu vermeiden. Die Höhe des Stuhles wurde so eingestellt, dass die Vpn einerseits ca. eine halbe Stunde bequem sitzen konnten und der Kopf andererseits angenehm in die Kopfstütze passte. Die Kopfstütze wurde verwendet, um nachträgliche Bewegungen zu vermeiden und somit konstant, genaue Messungen des Blickverfolgungsgerätes zu gewährleisten, sowie den Betrachtungsabstand für alle Vpn konstant zu halten. Nachdem alles eingestellt war, konnte das eigentliche Experiment beginnen, welches mit folgender Anleitung begann:

Die folgende Studie untersucht die Rolle von Aufmerksamkeit für die Wiedererkennung von Bildern.

Im ersten Teil des Experiments wirst Du eine Reihe von Fotos sehen. Schau Dir diese Fotos gut an, damit Du sie später gut wiedererkennen kannst. Während Du die Fotos ansiehst, werden wir Deine Augenbewegungen aufzeichnen.

Jedes Foto wird genau einmal für fünf Sekunden gezeigt. Bevor ein Foto gezeigt wird, erscheint in der Mitte des Bildschirms immer ein kleiner Punkt. Schau auf diesen Punkt, damit der Durchgang gestartet und das nächste Bild gezeigt werden kann.

Versuche während eines Durchgangs, d.h. während Du Dir ein Bild ansiehst, möglichst wenig zu blinzeln. Du kannst immer blinzeln, sobald ein Durchgang vorbei ist – also das Bild verschwindet.

Nachdem Du alle Bilder gesehen hast, wirst Du Gelegenheit zu einer kurzen Pause haben. Danach werden wir Dir Ausschnitte aus Fotos zeigen, und Du solltest bei jedem Ausschnitt entscheiden, ob dieser aus einem der Bilder stammt, die Du gesehen hast, oder aus einem anderen Bild, welches nicht gezeigt wurde.

Noch Fragen? Dann wende Dich bitte an die Versuchsleitung.

Ansonsten viel Spaß beim Betrachten der Bilder.

Drücke die Taste ´s´ auf der Tastatur, um das Experiment zu starten!

Anschließend wurden 30 Bilder gezeigt, welche natürliche Szenen des Alltags enthielten. Die restlichen 30 Fotografien wurden für die Testphase verwendet. Die Zuweisung der Bilder in Lern- oder Testphase erfolgte abwechselnd von einer Vp zu nächsten, so dass die nachfolgende Person jeweils das andere Set zu sehen bekam. Vor jedem Bild sollten die Vpn kurz ein Fixationskreuz betrachten und anschließend erfolgte die Präsentation des Bildes für 5000ms (Abbildung 5).

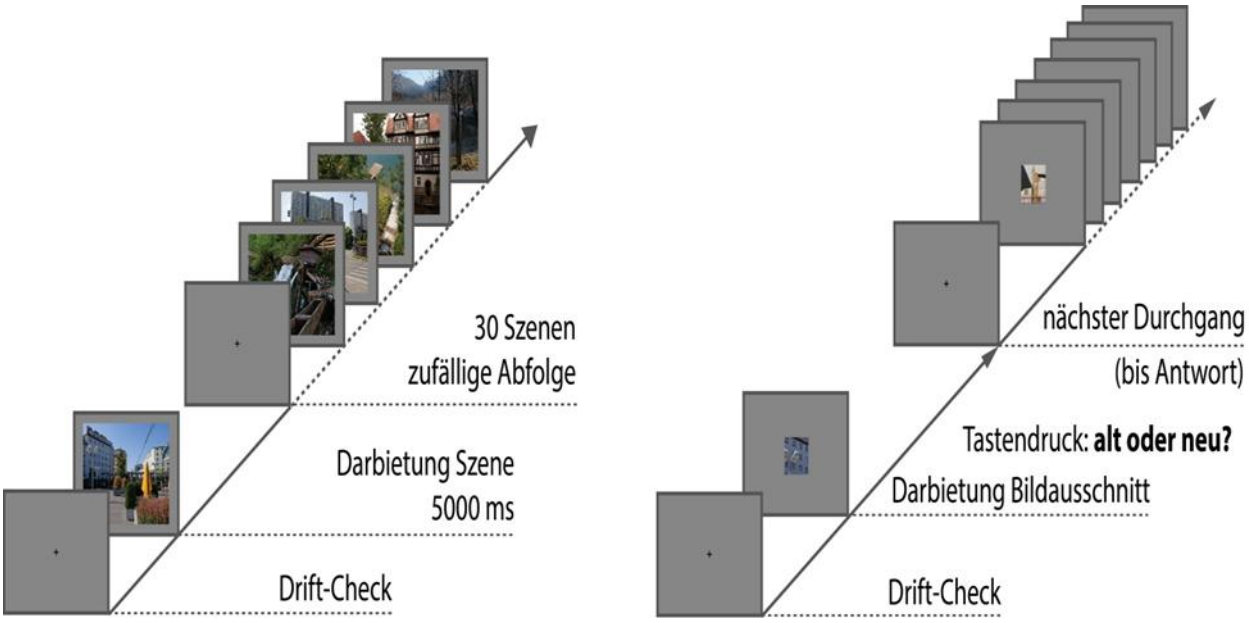


Abbildung 5: Versuchsablauf in der Lern- und Testphase. Jeweils nachdem das Fixationskreuz fixiert wurde, wurde zum nächsten Bild (in voller Größe) gewechselt (links). In der Testphase wurden die kleinen Bildbereiche in der Mitte des Bildes präsentiert (rechts).

Sobald die gemessene Blickposition um mehr als 1° vom Fixationskreuz abwich, wurde eine 9-Punkte Rekalibrierung des Blickverfolgers durchgeführt. Der Blickverfolger zeichnete während der Betrachtung die Fixationspunkte des Auges auf und speicherte dann den am längsten betrachteten Bereich, welcher im folgenden Testdurchgang präsentiert wurde. Nachdem alle 30 Bilder gezeigt wurden, erfolgte eine Pause von 10 Minuten und folgende Information erschien am Bildschirm:

Du hast nun alle Bilder gesehen und kannst jetzt eine kurze Pause machen um Dich zu lockern und zu entspannen. Im zweiten Teil des Experiments wirst Du nun Ausschnitte aus Bildern sehen. Diese Ausschnitte werden jetzt automatisch vorbereitet, was einige Minuten dauert. Sobald Du fortfahren kannst, wird dies rechts unten am Bildschirm angezeigt.

Die Ausschnitte werden in der Mitte des Bildschirms angezeigt und stammen zum Teil aus den Bildern die Du gesehen hast und zum Teil aus anderen Bildern, die Du noch nicht gesehen hast. Deine Aufgabe ist, bei jedem dieser Ausschnitte zu entscheiden, ob dieser aus einem Bild

stammt, welches Du zuvor gesehen hast, oder aber aus einem Bild, welches Du noch nicht gesehen hast.

Triff Deine Entscheidung möglichst schnell, denn wir werden die Zeit auswerten, die Du dafür brauchst. Arbeite jedoch trotzdem konzentriert und genau! Benutze die folgenden Tasten:

Drücke immer die Taste ´F‘, wenn es sich um einen Ausschnitt handelt, der aus einem bekannten Bild stammt, welches Du zuvor gesehen hast!

Drücke immer die Taste ´J´, wenn es sich um einen Ausschnitt aus einem unbekannten Bild handelt, also Du dieses Bild vorher noch nicht gesehen hast.

Schaue dabei immer auf den Bildausschnitt und lasse den linken und den rechten Zeigefinger während des Experiments immer auf den Tasten ´F´ und ´J´ ruhen, so dass Du möglichst schnell sein kannst. Vor jedem Durchgang wird wieder ein kleiner Punkt in der Mitte angezeigt, den Du anschauen musst, damit der Durchgang gestartet werden kann.

Während dieser Pause wurden die aufgezeichneten Blickdaten (Fixationspunkte) automatisiert ausgewertet und für die aktuelle Vp ein individuelles Set an Bildausschnitten erstellt. Dieses Set bestand aus insgesamt 120 Bildausschnitten in einer Größe von 3.5° x 3.5°. Dieser Bereich ist bei der Betrachtung von Szenen besonders sensitiv für Veränderungen (Henderson, Williams, Castelhano & Falk, 2003). Die eine Hälfte stammte aus Bildern, die in der Lernphase präsentiert wurden, die andere Hälfte bestand aus Teilen neuer Bilder. Von den gezeigten Bildern wurden die 30 jeweils am längsten fixierten Bereiche präsentiert, die restlichen 30 waren nicht betrachtete, jedoch hoch saliente Bereiche unterschiedlicher Bilder, welche mit Hilfe der Saliency Toolbox für MATLAB (Walther & Koch, 2006) generiert wurden. Die 60 Teilbereiche aus nicht präsentierten Bildern, teilten sich in 30 hoch saliente und 30 zufällige Bildbereiche. Die Salienz wurde durch Salienzkarten ermittelt. Es folgte ein zweiter Durchgang, in dem die Bildausschnitte in zufälliger Reihenfolge in der Mitte des Bildschirms präsentiert wurden und die Vpn zu entscheiden hatten, ob es sich um bereits in der Lernphase

gesehene Bilder handelte (Abbildung 5). Realisiert wurde diese Einschätzung durch Tippen der jeweiligen Taste auf der Tastatur wobei die Tastenbelegung („J“ für ja und „F“ für nein) von Vp zu Vp alternierend gewechselt wurde. Die Testung dauerte für jede Vp etwa 35 bis 50 Minuten, je nachdem wie zügig die Einstellung des Blickverfolgungsgerätes voran ging.

3.3 Datenanalyse

Als Rohdaten wurden die Reaktionszeiten, die Wiedererkennungsleistungen und die Fixationspunkte der Vpn verwendet. Um die Wiedererkennungsleistungen zu evaluieren, wurde zunächst die Anzahl der korrekt wiedererkannten fixierten und nicht fixierten Bildteile sowie der korrekten Zurückweisungen der nicht präsentierten Bildteile erhoben und mit den nicht präsentierten, salienten und zufälligen Bildteilen verglichen. Zusätzlich zur Behaltensleistung wurden die Reaktionszeiten für alle vier Gruppen (präsentiert/fixiert, präsentiert/nicht fixiert, nicht präsentiert/salient und nicht präsentiert/zufällig) ermittelt, wobei diese Analyse auch ausschließlich für Durchgänge, in denen korrekt geantwortet wurde, wiederholt wurde. Anschließend erfolgte eine weitere Analyse der Daten mittels Varianzanalysen (ANOVA vom engl. analysis of variance), geplanten Kontrasten und post-hoc Vergleiche.

Wurden die Tasten mehr als einmal betätigt, so wurde nur der erste Tastendruck in die Berechnungen miteinbezogen. Es wurden nur jene Fixationen gewertet, bei welchen sich die aufgenommenen Blickbewegungen um weniger als 0.1° änderten, die Blickgeschwindigkeit weniger als 30°/sek. war und die Blickbeschleunigung unter 8000°/sek. lag.

Auswahl der präsentierten Bildbereiche. Die Vpn konnten in der Lernphase grundsätzlich alle Bereiche der Bilder frei betrachten, es wurden jedoch nur jene Bereiche, welche außerhalb eines Sehwinkels von 3.5° (100 Pixel Durchmesser) von der Mitte und 70 Pixel (2.5°) vom Bildrand entfernt lagen, in die Berechnungen miteinbezogen. Ausgehend von Studien

(z.B. Hollingworth & Henderson, 2002), in denen die Fixationsdauer signifikant mit der Entdeckungsleistung für Veränderungen korrelierte, wurden für die fixierten Bildbereiche in der Testphase, die am längsten fixierten Bereiche in der Lernphase ausgewählt. Nichtfixierte Teilbereiche der Testphase, wurden aus präsentierten Bildern der Lernphase entnommen. Damit auch saliente Bereiche der präsentierten aber nicht fixierten Bilder der Lernphase in der Testphase gezeigt werden konnten, wurden vorberechnete Salienzkarten verwendet (Walther & Koch, 2006). Um Überschneidungen und Wiederholungen zu verhindern, wurden Bereiche in einem Umkreis von 3.5° um die getätigten Fixationen der Lernphase, sowie die oben erwähnten Randbereiche und der Mittelpunkt in der Salienzkarte auf null gesetzt und somit ausgeschlossen. Basierend auf dieser modifizierten Salienzkarte, wurden nicht fixierte aber saliente Bildbereiche aus der Lernphase ausgewählt. Diese durften sich nicht mehr als 5 % mit den fixierten Bereichen überlappen. Um auszuschließen, dass die Wiedererkennung von fixierten oder nicht fixierten Bereichen auf einen Vorteil für fixierte Bereiche, aufgrund höherer Salienz dieser Bereiche erfolgte, wurde ein post-hoc-Vergleich basierend auf der ursprünglich, unveränderten Salienzkarte, zwischen fixierten und nicht fixierten Teilbereiche der Lernphase durchgeführt (siehe Abschnitt 4.2). Fixationen unter 100 ms und über 2000 ms wurden ausgeschlossen. Die erste Fixation in jedem Durchgang wurde ebenfalls von den Analysen ausgeschlossen, da diese durch das zentrale Fixationskreuzes (welches vor jedem Durchgang präsentiert wurde) zustande kam. Die verbleibenden Daten wurden anschließend mit MATLAB und SPSS weiterführend analysiert. Bei allen folgenden statistischen Tests wurde das α-Niveau auf .05 gesetzt. Bei allen post-hoc Vergleichen wurde das α-Niveau Bonferroni korrigiert. Um die Daten auf Normalverteilung zu testen, wurde vor jeder ANOVA ein Kolmogorov-Smirnov-Test durchgeführt.

4 Ergebnisse

Alle jene Resultate, nämlich 9.1 %, welche 1.5 Standardabweichungen der durchschnittlichen individuellen Reaktionszeiten (RZn) übertrafen, wurden aus den Berechnungen entfernt. D.h. alle Resultate mit einer RZ von über 7354 Millisekunden (ms) wurden ausgeschlossen. Die folgende Tabelle gibt einen allgemeinen Überblick zu den Ergebnissen.

Tabelle 1

Verhaltensmaße der Wiedererkennungsleistung in Abhängigkeit der Versuchsbedingung

	präsentiert				nicht präsentiert			
	fixiert		nicht fixiert		salient		zufällig	
Leistung	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
RZ allgemein (ms)	1626	476	1822	639	1872	648	1936	739
RZ korrekt (ms)	1497	417	1714	632	1900	653	1957	714
% korrekt	61.1	12.5	47.5	14.3	65.1	14.5	67.9	12.9
Sensitivität (*d'*)	0.75	0.45	0.38	0.39				
Antwortbias (*c*)	0.17	0.28	0.26	0.32				

Wiedererkennungsleistungen. Zunächst wurde eine faktorielle ANOVA mit Messwertwiederholung für die durchschnittlichen Reaktionszeiten der vier Bedingungen (präsentiert/fixiert, präsentiert/nicht fixiert, nicht präsentiert/salient, nicht präsentiert/zufällig) berechnet. Es konnte ein Haupteffekt für die Art der Bildteile $F(3.69) = 7.3$, $p < .001$, bei einer Effektgröße von $\eta_p^2 = .24$ nachgewiesen werden. Um zwischen den Bedingungen zu unterscheiden, wurden geplante Kontraste (engl. „planned contrasts“) berechnet. Dabei konnte ein

signifikanter Effekt für fixiert/präsentierte Bilder festgestellt werden. In dieser Gruppe antworteten die Vpn schneller als in allen anderen Gruppen, wie präsentiert/nicht fixiert [$F(1.23) = 12.0$, $p < .01$, $\eta_p^2 = .34$], nicht präsentiert/salient [$F(1.23) = 12.1$, $p < .01$, $\eta_p^2 = .35$], oder nicht präsentiert/zufällige Bildteile [$F(1.23) = 13.4$, $p < .01$, $\eta_p^2 = .37$]. RZ zwischen den letzten drei Gruppen (präsentiert/nicht fixiert, nicht präsentiert/salient, nicht präsentiert/zufällig) waren nicht signifikant.

Eine weitere ANOVA wurde für die Mittelwerte der Reaktionszeiten für richtige Antworten durchgeführt. Die Ergebnisse zeigen ein identisches Muster wie die zuvor präsentierten. Wieder konnte ein signifikanter Effekt für die Art der Bildteile festgestellt werden $F(3.69) = 14.1$, $p < .001$, $\eta_p^2 = .38$ und geplante Vergleiche zeigten schnellere RZ für korrekte Antworten welche präsentiert und fixiert wurden als bei präsentiert/nicht fixierten [$F(1.23) = 12.0$, $p < .01$, $\eta_p^2 = .34$], nicht präsentiert/salienten [$F(1.23) = 22.5$, $p < .001$, $\eta_p^2 = .49$], oder nicht präsentiert/zufälligen [$F(1.23) = 23.5$, $p < .01$, $\eta_p^2 = .51$] Bildteilen. Ein anschließender post-hoc Vergleich (Bonferoni korrigiert) zeigte auch hier im Vergleich zu allen anderen Gruppen, signifikant verkürzte RZ für präsentiert/fixierte Bereiche (alle p`s $< .05$). Alle anderen Vergleiche fielen nicht signifikant aus.

Um anschließend die Wiedererkennungsgenauigkeit zu überprüfen, wurde erneut eine ANOVA mit Messwertwiederholung für die durchschnittliche Unterscheidungsleistung durchgeführt. Die Wiedererkennungsrate wurde, nach Beseitigung von Ausreißern bei den RZ (siehe oben), als Anteil korrekter Durchgänge berechnet. Die ANOVA, welche auf durchschnittliche Diskriminationsfähigkeit der einzelnen, variablen Bildteile (z.B. korrekte Wiedererkennung bei präsentiert/fixierten und präsentiert/nicht fixierten Bildteilen in % oder korrekte Zurückweisungen für nicht präsentiert/saliente und nicht präsentiert/zufällige Bildteile in %), welche während der Wiedererkennungsphase gezeigt wurden, testete, ergab einen signifikanten Effekt der korrekten Antworten, $F(3.69) = 10.9$, $p < .001$, $\eta_p^2 = .32$. Ein post-hoc

Paarvergleich zeigte, dass die Rate der korrekten Antworten in der präsentiert/nicht fixiert Gruppe in signifikanter Weise niedriger waren als in allen anderen verglichenen Gruppen (alle $p < .05$). Alle anderen Paarvergleiche waren nicht signifikant. Weiters wurde die Rate korrekter Antworten mit der Ratewahrscheinlichkeit von .5 für eine richtige Antwort verglichen. Die Wiedererkennungsleistung oder die korrekte Zurückweisung war bei präsentiert/fixierten [$t(23) = 4.3$, $p < .001$], nicht präsentiert/salienten [$t(23) = 5.1$, $p < .001$], und nicht präsentiert/zufälligen [$t(23) = 6.8$, $p < .001$] Bildteilen signifikant höher als die Zufallswahrscheinlichkeit. Präsentiert/nicht fixierte Bereiche unterschieden sich bei diesem Vergleich nicht signifikant von der Zufallswahrscheinlichkeit $t(23) = -0.9$.

4.1 Signalentdeckungsanalyse der Wiedererkennungsleistung

Um zu überprüfen, ob die hohe Wiedererkennungsrate der präsentiert/fixierten Bildbereiche auch tatsächlich auf Wiedererkennung beruht und nicht einer Antworttendenz unterliegt, wurden Maße der Signalentdeckungstheorie von Green und Swets (1966) verwendet. Hierbei werden die relativen Häufigkeiten der richtig erkannten Bildteile (Treffer) sowie der falschen Alarme (falsch wiedererkannte Bildteile) verwendet, um anschließend ein Sensitivitätsmaß bestimmen zu können. Dieses Sensitivitätsmaß (d') wird aus der Differenz zwischen Treffer und falscher Alarme berechnet und gibt Auskunft darüber wie genau die Wiedererkennungsleistung der Treffer tatsächlich ist. Weiters konnten die oben erwähnten relativen Häufigkeiten zur Bestimmung der Antworttendenz c verwendet werden. Diese spiegelt die individuelle Tendenz wieder, sich eher für gesehen oder nicht gesehen zu entscheiden. c repräsentiert also die Antworttendenz. Ein c-Wert von 0 bedeutet, sich für keine der beiden Antwortalternativen bevorzugt zu entscheiden. Ist das c signifikant abweichend unter 0, so stellt dies eine Antworttendenz in Richtung präsentiert dar. Ein signifikant positives c spricht für eine Antworttendenz in Richtung nicht präsentiert (Stanislaw & Todorov, 1999).

Um die Maßzahlen der Signalentdeckungstheorie zu erhalten, wurden zunächst die relativen Häufigkeiten der Treffer (richtig wiedererkannte Bildteile; präsentiert/fixiert, präsentiert/nicht fixiert) und der falschen Alarme (fälschlich wiedererkannte Bildteile bei nicht präsentiert/salienten und nicht präsentiert/zufälligen Bildteilen) ausgewertet. Um *d'* zu erhalten, wurde die falsche-Alarm-Rate, als Wahrscheinlichkeit nicht präsentierte Bildteile als präsentiert zu werten, berechnet. Um die Sensitivität zwischen fixiert und nicht fixiert unterscheiden zu können, wurde die Trefferquote für präsentiert/fixierte und präsentiert/nicht fixierte Bildteile getrennt berechnet. Ein *t*-Test für verbundene Stichproben ergab eine höhere Sensitivität für präsentiert/fixierte Bildbereiche als für präsentiert/nicht fixierte Bildteile, $t(23) = 6.1$, $p < .001$ (Abbildung 6).

Im Allgemeinen konnte eine durchschnittlich ablehnende Antworttendenz in beiden Bedingungen festgestellt werden. Betrachtet man jedoch präsentiert/fixierte und präsentiert/nicht fixierte Bildbereiche getrennt, so ist diese Tendenz nur bei präsentiert/nicht fixierten Bildteilen zu bemerken, was sich durch die signifikant von 0 differierenden durchschnittlichen c-Werte zeigt, $t(23) = 4.0$, $p < .01$ (Abbildung 6). Folglich entspricht die Wiedererkennungsleistung der präsentiert/fixierten Bereiche einer korrekten Diskriminationsfähigkeit zwischen präsentierten und nicht präsentierten Bildteilen welche keiner Antworttendenz unterliegt.

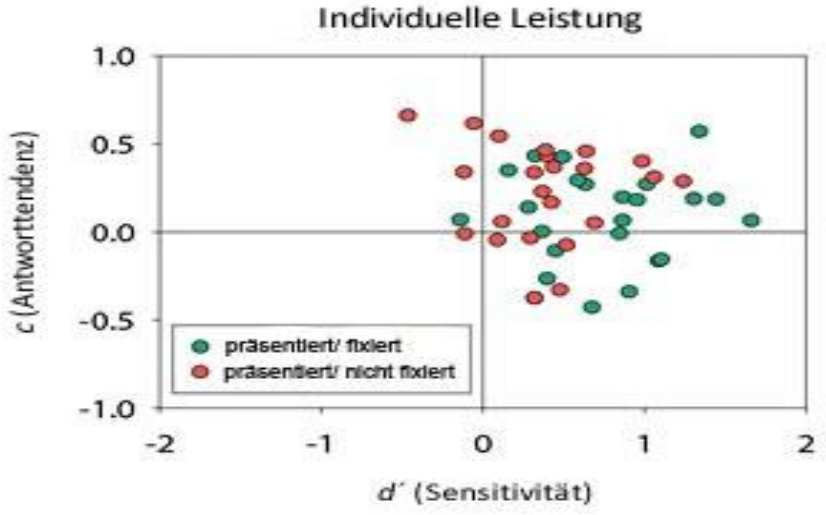


Abbildung 6: Streuung der individuellen Werte für *d'* und *c*.

4.2 Wie hängen nun Salienz, Wiedererkennungsleistung und Fixationsorte zusammen?

Wie einleitend beschrieben, spielt die Salienz eine bedeutende Rolle beim Blickverhalten (Elazary & Itti, 2008). Demnach werden saliente Bereiche eher betrachtet als nicht saliente. Um möglichst ähnliche Bedingungen in der Testphase zu gewährleisten, wurden nicht fixierte Bereiche nach Salienz ausgewählt. Es wurden also Bildteile ausgewählt, welche hoch salient waren und nicht fixiert wurden, um eine verbesserte Erinnerungsleistung aufgrund von Salienz auszuschließen. Um zu testen, ob sich präsentiert/fixierte und präsentiert/nicht fixierte Bildbereiche in der Testphase bezüglich Salienz unterscheiden, wurde ein post-hoc Vergleich, basierend auf der ursprünglich erstellten Salienzkarte, durchgeführt. Dafür wurde die relative Häufigkeit jener Durchgänge berechnet, in der die durchschnittliche Salienz der präsentiert/fixierten Bereiche im gleichen Bild jeder Vp höher war als in präsentiert/nicht fixierten Bildteilen. Die Ergebnisse sprechen hier sogar für eine niedrigere relative Wahrscheinlichkeit salientere fixierte Bildbereiche zu erhalten, als salient nicht fixierte Bildteile (M = .17, SD = .079). Diese Wahrscheinlichkeit lag sogar unter dem Zufallsniveau von .5, $t(23) = -20.2$, $p < .001$, bei dem die Salienz in beiden Gruppen (präsentiert/fixiert und präsentiert/nicht fixiert) gleich ist. Insofern ist es auch nicht möglich, dass die signifikant verbesserte Wiedererkennungsleistung der präsentiert/fixierten Bildbereiche durch höhere Salienz dieser Bildbereiche zustande kam.

Zusammenhang zwischen Salienz und Blickverhalten. Um der Frage nachzugehen, inwiefern Salienz in dieser Studie das Blickverhalten der Probanden beeinflusst, wurde eine Korrelation zwischen den Fixationspunkten der Probanden und der Salienzausprägung der fixierten Bereiche der Bilder der Lernphase errechnet. Dafür wurden Salienzkarten als binomiale Klassifikatoren für Fixationsorte verwendet (vgl. Wilming, Betz, Kietzmann & König, 2011). Zu-

nächst wurde für jedes der 60 Bilder, mit Hilfe der Saliency Toolbox für MATLAB (Walther & Koch, 2006) eine Salienzkarte erstellt. Dabei wurden die Standardeinstellungen verwendet. Um zu evaluieren wie gut die generierten Salienzkarten dem Blickverhalten (Fixationspunkte) der Vpn entsprechen, wurde eine Grenzwertoptimierungskurve (engl. „receiver operating characteristic“; ROC) verwendet. ROC ist eine Methode der visuellen Bewertung von Analysestrategien. Sie soll die Passung von Fixationen und Salienzkarte aufgrund der Fehlerrate darstellen. Die ROC-Kurve wurde durch eine stufenweise Schwellenwertbildung jeder einzelnen Salienzkarte, ausgehend vom minimalsten und bis zum höchsten Wert reichend, abgeleitet. Jene Salienzwerte, die sich bis zu einer maximalen Entfernung von 2° um die Fixationspunkte befinden, wurden dann mit den einzelnen Schwellenwerten jeder Stufe verglichen. Die ROC-Kurve wird anhand von Treffern und falschen Alarmen berechnet. In dieser Berechnung wurden fixierte Bereiche, welche einen bestimmten Schwellenwert der Salienz überstiegen, als Treffer gewertet und Salienzwerte darunter als Verpasser. Bildbereiche, welche nicht fixiert wurden, aber den Schwellenwert überstiegen, wurden als falsche Alarme gewertet. Zusätzlich zur ROC-Kurve wird die Fläche unter der Kurve berechnet (engl. „area under the curve“, AUC), welche die Klassifikationsgenauigkeit der Salienzkarte beschreibt. Diese kann einen Wert zwischen 0 und 1 annehmen, wobei ein durchschnittlicher über die Bilder generierter AUC-Wert welcher nicht signifikant von .5 abweicht bedeuten würde, dass das Salienzmodell, das Blickverhalten nicht überzufällig prognostiziert. Ist der durchschnittliche AUC-Wert hingegen signifikant abweichend von .5, so bedeutet dies, dass die Salienzkarten das Fixationsverhalten überzufällig vorhersagt. Nachdem einige Studien (z.B. Parkhurst et al., 2002) den höchsten Einfluss der Salienz in den ersten Blickbewegungen vorhersagen, wurde die Analyse einerseits für alle Fixationen durchgeführt und andererseits zusätzlich, getrennt, für die ersten fünf Fixationen.

Wurden alle Fixationen der Lernphase miteinbezogen, konnte ein durchschnittlicher AUC-Wert von .53 *(SD* = .028) ermittelt werden. Dieser Wert unterscheidet sich signifikant von der Zufallswahrscheinlichkeit von .5, *t(*59) = 8.97, *p* < .001 und spiegelt eine überzufällige Passung zwischen Fixation und Salienz wieder. Die getrennte Betrachtung der ersten fünf Fixationen führte zu einem ähnlichen Ergebnis von .53 *(SD* = .039) für die Fläche unter der Kurve, was wieder signifikant über dem Zufall liegt, *t(*59) = 6.18, *p* < .001. Ein Vergleich der eben genannten Ergebnisse führt zu keinem signifikanten Unterschied, *t(*59) = -0.43, also keinem Vorteil der ersten fünf Fixationen bezüglich Salienz. Die Ergebnisse sprechen für einen Zusammenhang zwischen Salienz und Blickverhalten und rechtfertigen somit die Auswahl salienter, präsentiert/nicht fixierter Bildteile in der Testphase.

5 Diskussion

Die vorliegende Arbeit untersuchte die Wiedererkennungsleistung von fixierten Bildbereichen natürlicher Szenen, wobei in der Testphase nur kleine Bildausschnitte zur Wiedererkennung präsentiert wurden und somit der Szenenkontext großteils ausgeschaltet wurde. Die Wiedererkennungsleistungen der in der Lernphase fixierten Bildbereiche wurden mit nicht fixierten aber präsentierten Bildteilen und mit nicht präsentierten Bildteilen, die in der Testphase gezeigt wurden, verglichen. Um einen Vorteil der Wiedererkennung aufgrund von Salienz auszuschließen, achtete man darauf, dass auch nicht fixierte aber präsentierte Bildbereiche der Lernphase, genauso salient waren wie fixierte Bereiche. Die Kontrollbedingung bestand zu gleichen Teilen aus nicht präsentiert/salienten und nicht präsentiert/zufälligen Bildteilen. Die 120 getesteten Bildteile wurden in zufälliger Reihenfolge vorgegeben und enthielten gleich viele Anteile in jeder der vier Gruppen (präsentiert/fixiert, präsentiert/nicht fixiert, nicht präsentiert/salient und nicht präsentiert/zufällig). Wie erwartet, konnten präsentiert/fixierte Bildteile signifikant besser und schneller wiedererkannt werden als präsentiert/nicht fixierte Bildteile. Letztere unterschieden sich, statistisch betrachtet, nicht von der Zufallswahrscheinlichkeit von .5 und das obwohl diese Bereiche bewusst so ausgewählt wurden, dass sie zumindest gleich salient waren wie fixierte Bereiche. Daher kann eine erleichterte Verarbeitung von fixierten Bereichen, nur aufgrund von Salienz, ausgeschlossen werden. Nicht präsentierte Bildteile konnten ebenfalls signifikant und korrekt zurückgewiesen werden. Dabei machte es kaum einen Unterschied ob diese Bildteile salient oder zufällig ausgewählt wurden (65.1 zu 67.9 %). Präsentiert fixierte Bereiche konnten zwar wesentlich besser wiedererkannt werden als präsentiert/nicht fixierte, die Leistungen lagen jedoch klar unter jenen Studien (z.B. Shepard, 1967; Standing, 1973; Standing et al., 1970), in welchen in der Testphase ganze Bilder

präsentiert wurden. Ein Vergleich der Fixationsdaten mit Salienzkarten der präsentierten Bilder zeigte einen schwachen aber signifikanten Zusammenhang. Die Berechnung der Diskriminationsfähigkeit von präsentierten und nicht präsentierten Bildteilen ergab einen Vorteil für fixierte Bereiche. Eine Antworttendenz konnte im Vergleich nur bei nicht fixierten Bildbereichen festgestellt werden.

Bezüglich der eingangs definierten Hypothesen stützen die gegenwärtigen Ergebnisse, die *visual memory theory of scene representation*. Fixierte Bildbereiche der Lernphase können signifikant besser und schneller von neuen, nicht präsentierten Bildbereichen in der Testphase unterschieden werden, als nicht fixierte Bildbereiche. Jene kleinen Bereiche einer Szene, die von der Versuchsperson fixiert, und somit auch in höchster Auflösung, detailliert wahrgenommen werden, werden somit im Langzeitgedächtnis gespeichert und können größtenteils wiedererkannt werden, auch ohne den Szenenkontext zur Verfügung zu haben. Nachdem nicht fixierte Bildbereiche nur auf Zufallsniveau wiedererkannt werden können, kann die Voraussetzung der Aufmerksamkeit für den zu erinnernden Bereich bestätigt werden. Nur jene Bildbereiche, welche fixiert werden, scheinen genügend Aufmerksamkeit zu erhalten, um auch später wiedererkannt zu werden. Bei den nicht präsentierten Bildern konnte kein signifikanter Unterschied zwischen salienten und zufällig ausgewählten Bildteilen festgestellt werden.

Wie erwartet konnte im Einklang mit der *visual memory theory of scene representation* (Henderson & Hollingworth, 2003; Hollingworth & Henderson, 2002; Hollingworth et al., 2001; Hollingworth, 2004, 2005) eine signifikant höhere Erinnerungsleistung von nachweislich betrachteten Bildbereichen bestätigt werden. Vergleicht man aber die Ergebnisse mit den ersten klassischen Studien zur Wiedererkennungsleistung von gesamten Bildern (Haber, 1970; Shepard, 1967; Standing et al., 1970; Standing, 1973), in denen bis zu 98 % von 600 Bildern wiedererkannt werden konnten (Shepard, 1967), so fällt in dieser Studie die Wieder-

erkennungsleistung von 61 % relativ niedrig aus. Die grundsätzliche Annahme war, dass die fixierten Bereiche die Wiedererkennung der Szene zu einem späteren Zeitpunkt erleichtert. Diese Annahme konnte zwar durch die signifikante Wiedererkennungsrate der fixierten Bildteile bestätigt werden, eine hinreichende Erklärung bietet sie jedoch nicht. Abgesehen von methodischen Einflüssen (z.B. die Instruktion), legen die hier erarbeiteten Ergebnisse den Schluss nahe, dass die alleinige Darbietung der Fixationen eine vollständige Erinnerung der Bilder eher behindert bzw. unzureichend ist. In der vorliegenden Arbeit wurde ja nur jener kleine Teilbereich in der Testphase dargeboten, welcher in der Lernphase mit dem Auge am längsten fixiert wurde. Es fehlte, im Gegensatz zu oben angesprochenen Studien, das restliche Bild. Einerseits können zwar fixierte Bereiche einer Szene relativ gut wiedererkannt werden, andererseits scheint für die vollständige Wiedererkennung, zu wenig Information in den fixierten Teilbereichen enthalten zu sein, um die Gesamtszene mit der gleichen Zuverlässigkeit zu erkennen, wie bei der Präsentation der vollständigen Szene. Ob und wie genau betrachtete Bereiche einer Szene wiedererkannt werden, wurde in der Vergangenheit immer wieder erforscht (z.B. Friedman, 1979; Goodman, 1980; Van der Linde et al., 2009; Mandler & Johnson, 1976; Pezdek et al., 1989; Salmaso et al., 1983) und deren Ergebnisse scheinen letztere Überlegungen zu stützen . Ein Großteil der älteren Studien verwendete zwar nur einfache schematische schwarz-weiß Zeichnungen mit merkmalsveränderten Distraktoren und zeichnete die Fixationen nicht auf, die Ergebnisse lassen jedoch den Schluss zu, dass alleiniges präsentieren der in der Lernphase betrachteten Bereiche einen Informationsverlust bedeutet. In der bereits angesprochenen Studie von Irwin und Zelinsky (2002) wurden die Fixationen aufgezeichnet und fixierte und nicht fixierte Objekte wurden anschließend ohne den ursprünglichen in der Lernphase gezeigten Hintergrund in der Testphase präsentiert. Obwohl diese Studie das Kurzzeitgedächtnis testete, indem die Testdurchgänge sofort nach der Lernphase gezeigt wurden, fiel die Wiedererkennungsrate der betrachteten Objekte auf ungefähr 65 %.

Nicht fokussierte Bereiche konnten jedoch zu 59 % wiedererkannt werden. Auch Van der Linde et al. (2009) berichtete eine Wiedererkennungsrate von 68 % für kurz nach der Präsentation einer Gesamtszene gezeigte fixierte Bildausschnitte. Wie wichtig der Szenenkontext beim Betrachten und Erinnern von natürlichen Szenen ist, beschreibt Torralba (2003). In Abhängigkeit von der Betrachtungsintention werden beim Betrachten einer Szene verschiedene Bereiche fixiert. Soll die gesamte Szene memoriert werden, so sind es jene Bereiche die fixiert werden, welche für das Behalten der Szene am wichtigsten sind. Die einzelnen Fixationen werden somit nicht nur einzeln gespeichert, sie stehen außerdem in Verbindung zueinander. Beim neuerlichen Betrachten der gleichen Szene wird das Wiedererkennen durch diese memorierten kontextuellen Verbindungen erleichtert.

Beim erstmaligen Betrachten einer Szene wird schon auf den ersten Blick die Szenenbedeutung erkannt (z.B. Oliva & Schyns, 1997; Thorpe, Fize & Marlot, 1996), wobei dies eine gewisse Vertrautheit mit der Szene voraussetzt. Sofern die Szene und die Relationen zwischen den Objekten der Szene bekannt sind, reichen dann auch schon einige wenige undeutliche Details um die Szene wiederzuerkennen. Dass System assoziiert die Lokalisation der Objekte mit den Eigenschaften ihres Hintergrundes. Je mehr Erfahrung man mit dieser Szene hat, umso schneller kann über die Anwesenheit eines bestimmten Objektes entschieden werden (Torralba, 2003). Der Kontext einer Szene scheint demnach unerlässlich für eine vollständige Erinnerung der Szene zu sein. Insofern ist es naheliegend, dass das Fehlen dieser kontextuellen Informationen zu Leistungseinbußen bei der Wiedererkennung von Szenen führt. Nachdem der Informationsgehalt von kleinen fixierten Bildteilen für die Wiedererkennung scheinbar nicht ausreichend ist um alle fixierten Bildbereiche im Gedächtnis zu behalten und tausende von ganzen Bildern größtenteils wiedererkannt werden, sofern sie auch vollständig präsentiert werden, liegt es nahe, dass mehr als nur ein einzelner, zuvor fixierter Bildausschnitt notwendig ist, um eine höhere Wiedererkennungsleistung zu erhalten. Dass möglichst viele visuelle

Informationen in einer ersten Präsentation notwendig sind um die nachfolgende Verarbeitung des dargebotenen Materials zu erleichtern, zeigen Sanocki und Epstein (1997). Sie führten diesbezüglich eine Studie durch, in der sie den Informationsgehalt von präsentierten natürlichen Szenen variierten und anschließend räumliche Aufgaben zu den zuvor präsentierten Szenen vorgaben. Sobald genügend Information in den ersten Bildern vorhanden waren – die Bandbreite reichte von einfachen fragmentarischen Strichzeichnungen bis zur kompletten Szene – konnten die Aufgaben gut gelöst werden. So gesehen ist auch anzunehmen, dass umgekehrt bei der Wiedererkennung eine möglichst informative Darstellung des zu wiedererkennenden Materials vorgegeben werden muss, um hohe Wiedererkennungsleistungen zu erzielen.

Nicht beachtete Bereiche einer Szene scheinen zumindest ohne Kontext nicht im Langzeitgedächtnis gespeichert und somit auch kaum erinnert zu werden. Nachdem nur zwei Antwortalternativen (ja oder nein) bei der vorliegenden Studie zur Verfügung standen und die Wiedererkennungsleistungen in dieser Gruppe auf Zufallsniveau lagen, ist es durchaus denkbar, dass auch ein Großteil der Treffer durch Raten zustande kam. Diese Annahme wird auch durch bereits erwähnte Studien zu Veränderungsblindheit unterstützt, die fanden, dass es schwer ist Veränderung an visuellen Vorlagen zu entdecken, selbst wenn danach aktiv gesucht wird (Rensink et al., 1997). Diese Blindheit tritt vor allem dann auf, wenn die visuelle Vorlage während der Veränderung nicht sichtbar ist. Weiterführende Studien zu diesem Thema konnten zeigen, dass auch Veränderungen während eines Lidschlages (O'Regan et al., 2000), einer Sakkade (Currie et al., 2000; Henderson & Hollingworth, 1999, 2003; Hollingworth & Henderson, 2002; McConkie & Currie, 1996) während das Zielobjekt durch ein anderes verstellt wird (Levin & Simons, 1997; Simons & Levin, 1998) oder während eines Filmschnittes (Levin & Simons, 1997) nicht oder kaum erkannt wurden. Bei der Veränderung wurde also darauf geachtet, dass das Auge den veränderten Bereich des Bildes während der

Veränderung nicht fixieren kann. Ähnlich verhält es sich mit *change detection* Untersuchungen, in denen natürliche Szenen zum Einsatz kamen und der thematische Kerngehalt nicht verändert wurde bzw. keine Anomalien verwendet wurden (z.B. ein umgedrehter Baum). Es wurde also lediglich die Form (Henderson & Hollingworth, 2003), die Farbe (Aginsky & Tarr, 2000), die Größe (McConkie & Currie, 1996), die räumliche Orientierung (Henderson & Hollingworth, 1999), die konzeptuelle Klassenzugehörigkeit (Henderson & Hollingworth, 2003) oder die Position (O'Regan et al., 2000) der Objekte verändert. Auch bei diesen Untersuchungen fiel die Entdeckungsrate der Veränderungen relativ niedrig aus, was nicht verwunderlich ist bei der steigenden Komplexität, die natürliche Szenen mit sich bringen. Wurde jedoch der veränderte Bereich eines Bildes kurz vor der Veränderung fixiert, wurden die Veränderungen sogar nach mehreren darauf folgenden Durchgängen erkannt (Hollingworth & Henderson, 2002). Die hier vorliegende Arbeit stützt also diese Ergebnisse, insofern als dass Bereiche beachtet werden müssen, um sie für längere Zeit im Gedächtnis zu behalten.

Des Weiteren könnte die Dauer der Betrachtung einen gewissen Einfluss auf die Erinnerungsleistung sowohl fixierter als auch nicht fixierter Bereiche haben. Je länger ein Bild betrachtet werden kann, umso mehr Fixationen werden vollzogen und desto näher werden die Fixationspunkte aneinander liegen, so dass die räumliche Nähe zu nicht fixierten Bereichen kleiner wird. So gesehen beeinflusst nicht nur die Verweildauer der Fixationen die Erinnerungsleistung sondern auch die gesamte Betrachtungsdauer der Szene. Nelson und Loftus (1980) führten diesbezüglich eine Studie durch, in der die Fixationspunkte beim Betrachten von Bildern der Vpn aufgezeichnet wurden und dann einen Rekognitionstest mit zwei Antwortalternativen (Testobjekt und Distraktor) durchführten. Je näher die gemessenen Fixationen an den getesteten Zielobjekten waren, umso eher wurden sie wiedererkannt. Eine ähnliche Studie führten Irwin und Zelinsky (2002) durch. Auch sie zeichneten die Blickbewegungen ihrer Probanden auf und variierten die Anzahl der Fixationsmöglichkeiten, in dem nach einer

gewissen vorgegebenen Anzahl an Fixationen das Bild verschwand. Das Bild in der Lernphase bestand aus sieben Objekten, die halbkreisförmig angeordnet waren. Je öfters die Vpn fixieren konnten und je mehr Zeit sie hatten das Bild zu betrachten, desto mehr Objekte konnten sie in der darauf folgenden Testphase im Gedächtnis behalten - bis zu 78 % konnten korrekt wiedererkannt werden. Insofern wäre es auch möglich, dass sich die Wiedererkennungsleistung für nicht fixierte Bereiche durch längere Präsentationszeiten verbessern ließe, da dann mehr Fixationen durchgeführt werden könnten. Somit würde die Dichte der Fixationspunkte zunehmen, wodurch sich die räumliche Nähe zu nicht fixierten Bereichen verringern würde und dieser Bereich dadurch eher registriert und im Gedächtnis behalten werden würde.

Wie eingangs erwähnt, postuliert die Scanpath-Theorie, dass die Wiedererkennung von Szenen durch die exakte Wiederholung der ursprünglichen Fixationsreihenfolge der Lernphase in der darauf folgenden Betrachtung erfolgt (vgl. Noton & Stark, 1971; Stark & Ellis, 1981). Das Wiedererkennen einer Szene ist demnach das Ergebnis des Vergleichs interner Repräsentationen mit dem aktuell wahrgenommenen sensumotorischen Muster. In der hier vorliegenden Studie, wurde die Möglichkeit, eine Szene anhand des gleichen Fixationsmusters erneut vollständig zu betrachten, unterbunden, in dem nur ein kleiner Ausschnitt der zuvor betrachteten Szene in der Testphase präsentiert wurde. Wenn die Wiedererkennung einer Szene durch den Vergleich eines gespeicherten sensumotorischen Musters mit dem aktuellen sensumotorischen Muster erfolgt, so sollte die Wiedererkennungsrate bzw. die Rate der korrekten Zurückweisungen, die Zufallswahrscheinlichkeit von .5 bei den präsentiert/fixierten Bildausschnitten, nicht übersteigen. Die Wiedererkennungsrate von 61 % war zwar in dieser Studie nicht besonders hoch, aber trotzdem signifikant, was heißt, dass das sensumotorische Fixationsmuster nicht ausschließlich für die Wiedererkennung verantwortlich sein kann, denn auch ohne diesen Abgleich konnte ein Großteil der fixierten Bereiche wiedererkannt werden.

Bezüglich des Salienzmodells konnte zwar ein signifikanter Zusammenhang zwischen Salienz und Blickverhalten festgestellt werden, doch zeigen die ermittelten Ergebnisse auch, dass die bewusst ausgewählten präsentiert/nicht fixierten (aber salienten) Bildteile der Testphase durchschnittlich salienter waren, als die präsentiert/fixierten Bildteile. Auch wenn ein signifikant von .5 abweichender AUC-Wert ermittelt werden konnte, so ist der Wert von .53 eher gering. Folglich werden zwar großteils saliente Bereiche betrachtet, ein beträchtlicher Teil der fixierten Bereiche ist aber nicht oder nur wenig salient. Es wurden also nicht ausschließlich hoch saliente Bereiche betrachtet. Ein Vergleich zeigte auch, dass die Wahrscheinlichkeit präsentiert/nicht fixierte saliente Bildteile zu erhalten höher war, als präsentiert/fixierte. Das heißt, dass mehr und höher saliente Bereiche in den Bildern vorhanden waren als tatsächlich fixiert wurden. Unter diesem Aspekt ist Salienz nur zu einem gewissen Teil für das Blickverhalten verantwortlich. Die schon erwähnte Studie von Van der Linde et al. (2009) bietet eine mögliche Erklärung. In dieser vom Design her ähnlichen Studie, wurden fixierte Bereiche mehrerer schwarz-weiß Bilder, auf ihre Eigenschaften hin untersucht und mit der Wiedererkennungsrate verglichen. Die Forscher fanden heraus, dass Luminanz zwar der beste Prädiktor für die Wiedererkennung fixierter Bildbereiche war, luminante Bereiche aber weniger oft fixiert wurden als andere Bereiche des gleichen Bildes. Saliente Bereiche konnten also gut behalten werden, wurden aber nicht vorrangig betrachtet. Auch wenn in der vorliegenden Arbeit die Eigenschaften der fixierten Bildbereiche nicht direkt untersucht wurden, weisen die oben erwähnten Ergebnisse auch hier darauf hin, dass nicht ausschließlich hoch saliente Bereiche von Szenen fixiert werden, auch wenn diese zu einer verbesserten Wiedererkennungsleistung der fixierten Bildbereiche beitragen könnten.

Erwartungsgemäß konnte bei den nicht präsentierten neuen Bildteilen der Testphase ebenfalls kein Unterschied zwischen hoch salienten und zufälligen Bildteilen für die Gedächtnisleistung (richtige Zurückweisungen) festgestellt werden. Salienz trägt demnach, zumindest in

diesem Experiment, nicht wesentlich dazu bei, fixierte Bildbereiche von neuen nicht, präsentierten Bildbereichen zu unterscheiden. Man könnte vermuten, dass saliente Bildbereiche eher betrachtet werden, weil sie einen größeren Informationsgehalt beinhalten und somit die Gedächtnisleistung unterstützen. Für die Unterscheidung zwischen fixierten und nicht präsentierten Bildbereichen scheint dies aber keine Rolle zu spielen.

In der Vergangenheit wurde immer wieder darauf hingewiesen, dass Aufmerksamkeit und Blickverhalten nicht immer im Einklang sind (Posner, 1980). Demzufolge ist es möglich, dass der Blick auf einen bestimmten Punkt fixiert ist, jedoch die Aufmerksamkeit auf eine andere Stelle einer Szene gerichtet ist. Insofern sollte es möglich sein, auch andere Bereiche einer Szene zu beachten, obwohl diese mit dem Auge nicht fixiert werden. Bei der Betrachtung von natürlichen Szenen sollten demnach auch nicht fixierte Bereiche ins Bewusstsein treten und folglich auch erinnert werden können. In der vorliegenden Studie konnte keine verdeckte Aufmerksamkeit festgestellt werden. Nicht fixierte Bereiche der präsentierten Bilder wurden auf Zufallsniveau erkannt. Dieses Ergebnis legt den Schluss nahe, dass die Aufmerksamkeit nur von einem Fixationspunkt zum nächsten wandert. Eine mögliche Erklärung dafür könnte die Aufgabenstellung sein. In Abhängigkeit von der Intention eine Szene zu betrachten, werden verschiedene Bildbereiche bevorzugt betrachtet. Es könnte auch sein, dass wenn die Aufgabenstellung es verlangt, eine Szene zu memorieren, bestimmte für den Szenenkontext wichtige Details, welche aber nicht unbedingt hoch salient sind, betrachtet werden. Torralba (2003) fand diesbezüglich heraus, dass kontextuelle Einflüsse, saliente Merkmale in den Schatten stellen können, sofern diese für die Aufgabe nicht wichtig sind. Aus vergangenen Studien ist bekannt, dass die Aufgabeninstruktion die Fixationsmuster während der Betrachtung einer Szene beeinflusst (vgl. Castelhano, Mack & Henderson, 2009) Die Instruktion verlangte im gegenwärtigen Versuch, die Bilder so zu betrachten, dass sie in einer folgenden Testphase möglichst gut wiedererkannt werden können. Es ist denkbar, dass die Vpn ein ihr

individuelles spezifisches Vorwissen verwendeten, um möglichst effizient ihren Blick zu lenken, um die 30 Bilder zu memorieren (vgl. Underwood, Foulsham & Humphrey, 2009). Bei freier Betrachtung, ohne Instruktion könnte der Blick breitgefächerter gestreut werden, so dass auch nicht fixierte Bereiche Zuwendung erfahren.

Abgesehen von der Aufgabenstellung, stellt sich auch die Frage der Repräsentativität des Versuchsdesigns in Alltagsbedingungen. Bei näherer Betrachtung ergeben sich eine Unzahl von Möglichkeiten natürliche Szenen zu betrachten. Man kann auf einer Parkbank sitzen und völlig im Gedanken versunken sein oder gezielt nach einem bestimmten Objekt suchen oder man streift wie ein Tourist mit dem Blick durch die Gegend und verweilt an Orten, die interessant erscheinen. Geht man in ein Museum, werden wieder die einzelnen Objekte genau betrachtet und die restliche Umgebung bleibt vermutlich weitgehend unbeachtet. Dies ist nur ein kleiner Auszug an möglichem Blickverhalten, insofern ist es möglich, dass unter anderen Bedingungen auch andere Ergebnisse zustande kommen. In Abhängigkeit vom jeweiligen, individuellen Betrachtungsmodus, sind also verschiedene Kombinationen von Fixationen möglich, welche die Wiedererkennungsleistungen beeinflussen könnten.

6 Ausblick

Die vorliegende Studie untersuchte die Wiedererkennungsleistung fixierter Bereiche natürlicher Szenen. Im Vergleich zu Studien in denen auch in der Testphase das vollständige Bild der Lernphase präsentiert wurde, konnte eine verringerte Wiedererkennungsleistung festgestellt werden. Folgend dessen bedeutet der Wegfall des Szenenkontextes einen Verlust an Information, welcher die verminderten Leistungen erklären könnte. Ungewiss ist jedoch, welche und wie viele Informationen genau notwendig sind, um eine nahezu perfekte Wiedererkennungsleistung von natürlichen Szenen zu gewährleisten. Wie viel der ursprünglich betrachteten Szene muss in der Testphase präsentiert werden um eine vollständige Wiedererkennungsleistung zu erhalten? Um dieser Frage nachzugehen, könnten die präsentierten Bildbereiche der Testphase sukzessive vergrößert und Unterschiede in den Wiedererkennungsleistungen ermittelt werden. Es ist wohl anzunehmen, dass mit der Größe der Bildteile die Wiedererkennungsleistung der zuvor präsentierten vollständigen Bilder ansteigt.

Um den Informationsgehalt der fixierten Bereiche in Kombination zu messen, könnten auch mehrere fixierte Bereiche einer Szene an den entsprechenden originalen Stellen ohne den Rest des Bildes präsentiert werden. Verschiedenste Kombinationen sind dabei möglich wobei auch die Anzahl der präsentierten Fixationen variiert werden könnte. Die erhaltenen Daten würden Auskunft über die Zweckmäßigkeit, sowie der Anzahl an benötigten Fixationen beim Memorieren geben. Sich nur die fixierten Bereiche beim Betrachten einer Szene zu merken, scheint für den Alltag nicht sinnvoll, da wir ja immer gesamte Szenen zu Verfügung haben. Insofern wundert es auch nicht, wenn ein einzelner fixierter Bildbereich nicht ausreicht, um bekannte Szenen mit hoher Sicherheit wiederzuerkennen. Eine entsprechende Aufgabenstellung könnte die Ergebnisse beträchtlich beeinflussen, und zwar dann, wenn darauf

hingewiesen wird, dass die fixierten Bereiche memoriert werden sollen um die Wiedererkennung anschließend abzufragen.

Literatur

Aginsky, V. & Tarr, M. J. (2000). How are different properties of a scene encoded in visual memory? *Visual Cognition*, *7*(1-3), 147–162.

Ansorge, U. (2006). Die Rolle von Absichten bei der automatischen Verarbeitung visuell-raeumlicher Reizinformation. *Psychologische Rundschau*, *57*(1), 2–12.

Ansorge, Leder. (2011). *Wahrnehmung und Aufmerksamkeit* (1. Aufl.). Wiesbaden: VS-Verl. für Sozialwiss.

Bartram, D. (1974). Role of Visual and Semantic Codes in Object Naming. *Cognitive Psychology*, *6*(3), 325–356. doi:10.1016/0010-0285(74)90016-4

Becker, S. I. (2008). The mechanism of priming: Episodic retrieval or priming of pop-out? *Acta Psychologica*, *127*(2), 324–339. doi:10.1016/j.actpsy.2007.07.005

Biederman, I., Mezzanotte, R. & Rabinowitz, J. (1982). Scene Perception - Detecting and Judging Objects Undergoing Relational Violations. *Cognitive Psychology*, *14*(2), 143–177. doi:10.1016/0010-0285(82)90007-X

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. doi:10.1163/156856897X00357

Brandt, S. A. & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, *9*(1), 27–38. doi:10.1162/jocn.1997.9.1.27

Broadbent, D. E. (1958). *Perception and communication*.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. doi:10.1016/j.visres.2011.04.012

Castelhano, M. S., Mack, M. L. & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of vision*, *9*(3).

Chun, M. M. (2003). Scene perception and memory. *Psychology of Learning and Motivation: Advances in Research and Theory: Cognitve Vision, Vol 42*, *42*, 79–108.

Chun, M. M. & Nakayama, K. (2000). On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Visual Cognition*, *7*(1-3), 65–81.

Corbetta, M. & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. doi:10.1038/nrn755

Currie, C. B., McConkie, G. W., Carlson-Radvansky, L. A. & Irwin, D. E. (2000). The role of the saccade target object in the perception of a visually stable world. *Perception & Psychophysics*, *62*(4), 673–683. doi:10.3758/BF03206914

Egeth, H. (1966). Parallel Versus Serial Processes in Multidimensional Stimulus Discrimination. *Perception & Psychophysics*, *1*(8), 245–252. doi:10.3758/BF03207389

Einhäuser, W., Rutishauser, U. & Koch, C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision*, *8*(2). doi:10.1167/8.2.2

Elazary, L. & Itti, L. (2008). Interesting objects are visually salient. *Journal of Vision*, *8*(3). doi:10.1167/8.3.3

Friedman, A. (1979). Framing Pictures - Role of Knowledge in Automatized Encoding and Memory for Gist. *Journal of Experimental Psychology-General*, *108*(3), 316–355. doi:10.1037//0096-3445.108.3.316

Goodman, G. S. (1980). Picture memory: How the action schema affects retention. *Cognitive Psychology*, *12*(4), 473–495. doi:10.1016/0010-0285(80)90017-1

Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Haber, R. (1970). How We Remember What We See. *Scientific American*, *222*(5), 104.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504. doi:10.1016/j.tics.2003.09.006

Henderson, J. M. & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science, 10*(5), 438–443. doi:10.1111/1467-9280.00183

Henderson, J. M. & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics*, *65*(1), 58–71. doi:10.3758/BF03194783

Henderson, J. M., Weeks, P. A. & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology-Human Perception and Performance*, *25*(1), 210–228. doi:10.1037//0096-1523.25.1.210

Henderson, J. M., Williams, C. C., Castelhano, M. S. & Falk, R. J. (2003). Eye movements and picture processing during recognition. *Perception & Psychophysics*, *65*(5), 725–734. doi:10.3758/BF03194809

Henderson, J., Pollatsek, A. & Rayner, K. (1987). Effects of Foveal Priming and Extrafoveal Preview on Object Identification. *Journal of Experimental Psychology-Human Perception and Performance*, *13*(3), 449–463. doi:10.1037/0096-1523.13.3.449

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, *16*(4), 219–222. doi:10.1111/j.1467-8721.2007.00507.x

Henderson, J. M. & Hollingworth, A. (1999). High-level scene perception. In J. T. Spence (Hrsg.), *Annual Review of Psychology* (Bd. 50, S. 243–271).

Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology-Human Perception and Performance*, *29*(2), 388–403. doi:10.1037/0096-1523.29.2.388

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology-Human Perception and Performance*, *30*(3), 519–537. doi:10.1037/0096-1523.30.3.519

Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, *31*(3), 396–411. doi:10.1037/0278-7393.31.3.396

Hollingworth, A. & Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, *7*(1-3), 213–235.

Hollingworth, A. & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology-Human Perception and Performance*, *28*(1), 113–136. doi:10.1037//0096-1523.28.1.113

Hollingworth, A., Schrock, G. & Henderson, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*, *29*(2), 296–304. doi:10.3758/BF03194923

Hollingworth, A., Williams, C. C. & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, *8*(4), 761–768. doi:10.3758/BF03196215

Irwin, D. (1992). Memory for Position and Identity Across Eye-Movements. *Journal of Experimental Psychology-Learning Memory and Cognition*, *18*(2), 307–317. doi:10.1037/0278-7393.18.2.307

Irwin, D. E. & Zelinsky, G. J. (2002). Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics*, *64*(6), 882–895. doi:10.3758/BF03196793

Itti, L., Koch, C. & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *Ieee Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259. doi:10.1109/34.730558

Kelley, T. A., Chun, M. M. & Chua, K. P. (2003). Effects of scene inversion on change detection of targets matched for visual salience. *Journal of Vision*, *3*(1), 1–5. doi:10.1167/3.1.1

Kinchla, R. (1992). Attention. *Annual Review of Psychology*, *43*, 711–742. doi:10.1146/annurev.ps.43.020192.003431

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147. doi:10.1016/S1364-6613(00)01452-2

Koch, C. & Ullman, S. (1985). Shifts in Selective Visual-Attention - Towards the Underlying Neural Circuitry. *Human Neurobiology*, *4*(4), 219–227.

Koch, C. & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. doi:10.1016/j.tics.2006.10.012

Konkle, T., Brady, T. F., Alvarez, G. A. & Oliva, A. (2010). Scene Memory Is More Detailed Than You Think: The Role of Categories in Visual Long-Term Memory. *Psychological Science*, *21*(11), 1551–1556. doi:10.1177/0956797610385359

Kristjansson, A. (2006). Simultaneous priming along multiple feature dimensions in a visual search task. *Vision Research*, *46*(16), 2554–2570. doi:10.1016/j.visres.2006.01.015

Levin, D. T. & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, *4*(4), 501–506. doi:10.3758/BF03214339

Linde, I. van der, Rajashekar, U., Bovik, A. C. & Cormack, L. K. (2009). Visual memory for fixated regions of natural images dissociates attraction and recognition. *Perception*, *38*(8), 1152 – 1171. doi:10.1068/p6142

Maljkovic, V. & Nakayama, K. (1996). Priming of pop-out .2. The role of position. *Perception & Psychophysics*, *58*(7), 977–991. doi:10.3758/BF03206826

Maljkovic, V. & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, *22*(6), 657–672. doi:10.3758/BF03209251

Mandler, J. & Johnson, N. (1976). Some of Thousand Words a Picture Is Worth. *Journal of Experimental Psychology-Human Learning and Memory*, *2*(5), 529–540. doi:10.1037//0278-7393.2.5.529

Mandler, J. & Parker, R. (1976). Memory for Descriptive and Spatial Information in Complex Pictures. *Journal of Experimental Psychology-Human Learning and Memory*, *2*(1), 38–48. doi:10.1037//0278-7393.2.1.38

Mandler, J. & Ritchey, G. (1977). Long-Term-Memory for Pictures. *Journal of Experimental Psychology-Human Learning and Memory*, *3*(4), 386–396. doi:10.1037//0278-7393.3.4.386

McConkie, G. W. & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology-Human Perception and Performance*, *22*(3), 563–581. doi:10.1037//0096-1523.22.3.563

Nelson, W. & Loftus, G. (1980). The Functional Visual-Field During Picture Viewing. *Journal of Experimental Psychology-Human Learning and Memory*, *6*(4), 391–399. doi:10.1037/0278-7393.6.4.391

Noton, D. & Stark, L. (1971). Scanpaths in Saccadic Eye Movements While Viewing and Recognizing Patterns. *Vision Research*, *11*(9), 929–&. doi:10.1016/0042-6989(71)90213-6

O'Regan, J. K., Deubel, H., Clark, J. J. & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, *7*(1-3), 191–211.

Oliva, A. & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*(1), 72–107. doi:10.1006/cogp.1997.0667

Palmer,. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. doi:10.3758/BF03197524

Parkhurst, D., Law, K. & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107–123. doi:10.1016/S0042-6989(01)00250-4

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. doi:10.1163/156856897X00366

Pertzov, Y., Zohary, E. & Avidan, G. (2009). Implicitly perceived objects attract gaze during later free viewing. *Journal of vision*, *9*(6).

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N. & Dougherty, T. (1989). Memory for Real-World Scenes - the Role of Consistency with Schema Expectation. *Journal of Experimental Psychology-Learning Memory and Cognition*, *15*(4), 587–595. doi:10.1037//0278-7393.15.4.587

Posner, M. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, *32*(FEB), 3–25. doi:10.1080/00335558008248231

Posner, M., Snyder, C. & Davidson, B. (1980). Attention and the Detection of Signals. *Journal of Experimental Psychology-General*, *109*(2), 160–174. doi:10.1037//0096-3445.109.2.160

Rensink, R. A. (2000a). Seeing, sensing, and scrutinizing. *Vision Research*, *40*(10-12), 1469–1487. doi:10.1016/S0042-6989(00)00003-1

Rensink, R. A. (2000b). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition*, *7*(1-3), 345–376. doi:10.1080/135062800394847

Rensink, R. A. (2000c). The dynamic representation of scenes. *Visual Cognition, 7*(1-3), 17–42. doi:10.1080/135062800394667

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology, 53*, 245–277. doi:10.1146/annurev.psych.53.100901.135125

Rensink, R. A., ORegan, J. K. & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368–373. doi:10.1111/j.1467-9280.1997.tb00427.x

Salmaso, P., Baroni, M., Job, R. & Peron, E. (1983). Schematic Information, Attention, and Memory for Places. *Journal of Experimental Psychology-Learning Memory and Cognition, 9*(2), 263–268. doi:10.1037//0278-7393.9.2.263

Sanocki, T. & Epstein, W. (1997). Priming spatial layout of scenes. *Psychological Science, 8*(5), 374–378. doi:10.1111/j.1467-9280.1997.tb00428.x

Schacter, D., Chiu, C. & Ochsner, K. (1993). Implicit Memory - a Selective Review. *Annual Review of Neuroscience, 16*, 159–182. doi:10.1146/annurev.ne.16.030193.001111

Schütz, A. C., Braun, D. I. & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11*(5). doi:10.1167/11.5.9

Shepard, R. (1967). Recognition Memory for Words Sentences and Pictures. *Journal of Verbal Learning and Verbal Behavior, 6*(1), 156–&. doi:10.1016/S0022-5371(67)80067-7

Shore, D. I. & Klein, R. M. (2000). The effects of scene inversion on change blindness. *Journal of General Psychology, 127*(1), 27–43.

Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition, 7*(1-3), 1–15. doi:10.1080/135062800394658

Simons, D. J. & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review, 5*(4), 644–649. doi:10.3758/BF03208840

Simons, D. J. & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. doi:10.1016/j.tics.2004.11.006

Simons, D. J. & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267. doi:10.1016/S1364-6613(97)01080-2

Snodgrass, J. & Feenan, K. (1990). Priming Effects in Picture Fragment Completion - Support for the Perceptual Closure Hypothesis. *Journal of Experimental Psychology-General*, *119*(3), 276–296. doi:10.1037/0096-3445.119.3.276

Standing, L. (1973). Learning 10,000 Pictures. *Quarterly Journal of Experimental Psychology*, *25*(MAY), 207–222. doi:10.1080/14640747308400340

Standing, L., Conezio, J. & Haber, R. (1970). Perception and Memory for Pictures - Single-Trial Learning of 2500 Visual Stimuli. *Psychonomic Science*, *19*(2), 73–74.

Stanislaw, H. & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods Instruments & Computers*, *31*(1), 137–149. doi:10.3758/BF03207704

Stark, L. & Ellis, S. (1981). Scanpaths revisited: cognitive models direct active looking. In D. Fisher (Hrsg.), *Eye movements: cognition and visual perception* (S. 193–226). Lawrence Erlbaum Associates.

Taylor, D. (1976). Effect of Identity in Multi-Letter Matching Task. *Journal of Experimental Psychology-Human Perception and Performance*, *2*(3), 417–428. doi:10.1037//0096-1523.2.3.417

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. doi:10.1016/j.actpsy.2010.02.006

Thorpe, S., Fize, D. & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. doi:10.1038/381520a0

Torralba, A. (2003). Modeling global scene factors in attention. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *20*(7), 1407–1418. doi:10.1364/JOSAA.20.001407

Treisman, A. M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. doi:10.1016/0010-0285(80)90005-5

Underwood, G., Foulsham, T. & Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: Eye movements during encoding and recognition. *Visual Cognition*, *17*(6-7), 812–834. doi:10.1080/13506280902771278

Walther, D. & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*(9), 1395–1407. doi:10.1016/j.neunet.2006.10.001

Wilming, N., Betz, T., Kietzmann, T. C. & Konig, P. (2011). Measures and limits of models of fixation selection. *PloS one*, *6*(9).

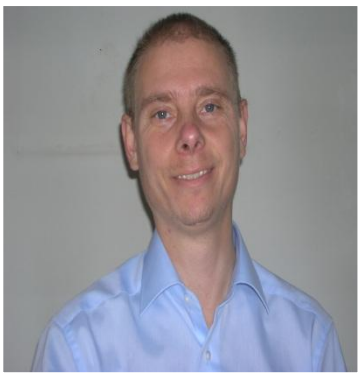
Zelinsky, G. J. & Loschky, L. C. (2005). Eye movements serialize memory for objects in scenes. *Perception & Psychophysics*, *67*(4), 676–690. doi:10.3758/BF03193524

Abbildungsverzeichnis

Tabellenverzeichnis

Gerhard Böck

Geboren am 1. November 1974,
Wohnhaft in 1140 Wien
Mail:gerhard-boeck@gmx.at
Österreichischer Staatsbürger



Ausbildung

2005-2013	Diplomstudium Psychologie: Zweiter Studienabschnitt, Universität Wien, Abschluss im März 2013 (geplant)
2000-2005	Diplomstudium Psychologie: Erster Studienabschnitt, Universität Wien, 1. Diplomprüfungszeugnis vom 25. Februar 2005
1994-2000	Bundeshandelsakademie für Berufstätige, 1100 Wien Reife- und Diplomprüfung am 26. Juni 2000
1990-1991	Bundesrealgymnasium Kandlgasse, 1070 Wien
1985-1990	Bundesrealgymnasium Marchettigasse, 1060 Wien
1981-1985	Volksschule Sonnenuhrgasse, 1060 Wien

Berufliche Tätigkeiten

Seit September 2010	Behindertenbetreuung, Auftakt, Freizeitassistenz
November 2011	Praktikum, OWS, Neuropsychologische Diagnostik
2005-2010	Vereinsarbeit, Tüwi (Verein zur Förderung von Kommunikation, Interaktion und Integration), Bar- und Kochdienste sowie Organisations- und Veranstaltungsmanagement
1999-2009	Lebensmittelkontrolle, Kontrollstelle für artgemäße Nutztierhaltung, Eierkontrollen im Rahmen des Konsumentenschutzes
1992-1998	Berufskraftfahrer, Fa. J. Kandler, Belieferungen, Übersiedlungen, Eröffnungen
1990-1992	Diverse Tätigkeiten: Paketdienst, Post; Schweißtechnik, Security
nebenbei seit 2000	Promotion, Laborexperimente der Meduni Wien

Sprachkenntnisse

Muttersprachen	Deutsch
	Slowakisch
Fremdsprachen	Englisch
	Französisch

Zusätzliche

Weiterbildungen	Systemische Familienaufstellung bei Lorenz Wiest, progressive Muskelrelaxation nach Jakobson, Kinesiologie, Massage
Führerschein	A, B, zwanzigjährige Fahrpraxis
Computerkenntnisse	Linux, Microsoft Windows
DTP	Open Office.org, Microsoft Office
Scientific	SPSS

Interessen

Sport	Laufen, Klettern, Schitouren, Wandern, Radfahren, Inlineskaten, Snowboarden, Schifahren, Yoga
Kreatives	Malen, Holzarbeiten, Knüpfen, Nähen, Töpfern, Kochen
Diverses	Meditation, kontemplative Psychologie, Nachhaltigkeit, Ökologie und Abfallwirtschaft