



**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2006

Adaptives Lernen mit Wissensstrukturkarten : Mess- und Skalierungsmodelle zur Modellierung von klassifikatorischem Expertenwissen auf dem Prüfstand

Schlatter, Katharina

Abstract: Die vorliegende Arbeit beleuchtet diejenigen kognitiven Prozesse, die das objektbasierte Merkmalswissen organisieren. Die drei Wissensgebiete, Ornithologie, Psychopathologie und Geographie lieferten hierzu die Datenbasis. Das Verfahren der Nonmetrischen Multidimensionale Skalierung (NMDS) bietet die Möglichkeit, Wissen einzelner Personen relational als geometrische Karte darzustellen (=Wissensstrukturkarte). Im Zentrum der Forschung stand die Frage, mit welchen didaktischen Möglichkeiten ein Lerner in der Ausbildung seines Wissensmodells effizient unterstützt werden kann. Die dazu durchgeführten Lernexperimente zeigen, dass sich Kognitive Karten besonders gut dazu eignen, um die Integration von Faktenwissen beim Lernen zu beschleunigen. Zwei weitere Lernexperimente weisen einen positiven Zusammenhang zwischen dem Merkmalswissen eines Lerners und der Güte seiner Wissensstruktur nach. Ein auf Wissensstrukturkarten zurückgreifendes Diagnostikmodell für die Messung individuellen Wissens wurde entwickelt. Darauf basierend wurde ein computergestütztes Intelligentes Tutorielles System (IST) konzipiert. Neben diesem praktischen Ertrag dokumentiert die Arbeit auch allgemeinere kognition-psychologische Einsichten über das Wechselspiel assimilativer und akkomodativer Prozesse beim Lernen von deklarativem Faktenwissen sowie über die Annahme von Hierarchien in der kognitiven Organisation objektbasierten deklarativen Wissens. The present thesis illuminates cognitive processes organizing object-based feature knowledge. The database was gained from three fields of knowledge, ornithology, psychopathology, and geography. Non-metric multidimensional scaling (NMDS) provides the opportunity to relationally depict the knowledge of an individual person in a geometrical map (knowledge structure map). The aim of this research was to determine the didactical possibilities which can efficiently support a learner while shaping a knowledge structure. The conducted learning studies demonstrate that cognitive maps accelerate the integration of factual knowledge in the learning process. Two further learning experiments exhibit a positive correlation between feature knowledge of a learner and the quality (goodness?) of his/her structure of knowledge map. Based on these maps a diagnostic model was created to measure individual knowledge. Adapted from this model a computer-assisted intelligent tutorial system (ITS) was developed. In addition to this practical profit general insights of cognitive psychology are documented concerning the interplay of assimilated and accommodated processes while learning factual knowledge on the one hand and the assumption of hierarchies in the cognitive organisation of object-based declarative knowledge on the other hand.

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-163362>

Dissertation

Published Version

Originally published at:

Schlatter, Katharina. Adaptives Lernen mit Wissensstrukturkarten : Mess- und Skalierungsmodelle zur Modellierung von klassifikatorischem Expertenwissen auf dem Prüfstand. 2006, University of Zurich, Faculty of Arts.

Adaptives Lernen mit Wissensstrukturkarten

-

**Mess- und Skalierungsmodelle zur Modellierung von
klassifikatorischem Expertenwissen auf dem Prüfstand**

Abhandlung
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der Universität Zürich

vorgelegt von
Katharina Schlatter
Buchs (ZH)

Angenommen auf Antrag von
Prof. Dr. Damian Läge und
Prof. Dr. Wolfgang Marx

Zürich, 2006

Zusammenfassung

Die vorliegende Arbeit beleuchtet diejenigen kognitiven Prozesse, die das objektbasierte Merkmalswissen organisieren. Die Wissensgebiete Ornithologie, Psychopathologie und Geographie bilden die exemplarische Datenbasis für diese psychologische Untersuchung. Um Klassifikationssysteme für die in diesen Gebieten relevanten Objekte (Vogelarten, Störungsbilder und Länder) modellieren zu können, wurde vor allem auf Nonmetrische Multidimensionale Skalierung (NMDS) zurückgegriffen. Dieses Verfahren bietet die Möglichkeit, das Wissen einzelner Personen relational als geometrische Karten darzustellen (in Form von Wissensstrukturkarten, auch „Kognitive Karten“ genannt).

Im Zentrum der vorgelegten Forschung steht die Frage, mit welchen didaktischen Möglichkeiten ein Lerner in der Ausbildung seines Wissensmodells effizient unterstützt werden kann. Die dazu durchgeführten Lernexperimente zeigen, dass sich Kognitive Karten durch ihre relationale Struktur besonders gut eignen, um die Integration von Faktenwissen beim Lernen zu beschleunigen. Zwei weitere Lernexperimente weisen einen positiven Zusammenhang zwischen dem Merkmalswissen eines Lerners und der Güte seiner Kognitiven Karte nach: Je mehr ein Laie weiss, desto mehr ähnelt seine Karte derjenigen eines Experten auf dem Gebiet. Diese Einsicht mündet in ein Verfahren, welches es uns ermöglicht, die Qualität individuellen Wissens Objekt für Objekt über den Vergleich von Wissensstrukturkarten zu messen. Auf diesem Verfahren aufbauend wurde ein computergestütztes Intelligentes Tutorielles System (IST) konzipiert, welches den mühsamen Weg des Laien zum Experten hin aufgrund individuell angepasster Lerninhalte erleichtern soll.

Neben diesem praktischen Ertrag dokumentiert die Arbeit auch allgemeinere kognitionspsychologische Einsichten über das Wechselspiel assimilativer und akkomodativer Prozesse beim Lernen von deklarativem Faktenwissen sowie über die Annahme von Hierarchien in der kognitiven Organisation objektbasierten deklarativen Wissens. Diese Einsichten wurden im Zusammenspiel mit den experimentellen Daten und unter Verwendung von innovativen dimensionsorientierten und klassifikatorischen Analysemethoden gewonnen und eröffnen so auch methodologisch einen neuen Horizont für die Expertise-Forschung.

Inhaltsverzeichnis

1. Überblick über die Arbeit und Einführung in den messtheoretischen Rahmen	1
2. Die Modellierung von Klassifikatorischem Expertenwissen am Fallbeispiel der Ornithologie	15
3. Die Rolle von Gewichtungsfaktoren bei Parallelen Sortierdaten im Hinblick auf spätere NMDS Lösungen	43
4. Kognitive Adäquatheit von Wissensstrukturkarten – eine explorative Studie	53
5. Lernen mit Wissensstrukturkarten	71
6. DIVE IN – adaptiver Unterricht mit hierarchisch ineinandergreifenden Wissensstrukturkarten	81
7. Die Veränderung von Wissenskarten beim Merkmalslernen	93
8. Der NMDS-Stresswert als Maß für Assimilation und Akkomodation beim Lernen von Wissensstrukturen	119
9. A Structure-Based Expert Model of the ICD-10 Mental Disorders	143
10. Das Zürcher Wissensstrukturmodell für Psychopathologie (ZWP) – eine Pilotstudie	157

1.

**Überblick über die Arbeit und
Einführung in den messtheoretischen Rahmen**

1.1. Veränderung von Mentalen Modellen beim Lernen

In der Wissenspsychologie geht man davon aus, dass Menschen ihr Wissen bezüglich verschiedenartigster Phänomene in Form von mentalen Modellen organisieren. Seel (2000, Glossar) definiert die mentalen Modelle als „... kognitive Konstruktionen, mittels derer eine Person ihre Erfahrung oder ihr Denken derart organisiert, dass sie eine systematische Repräsentation ihres Wissens erreicht, um subjektive Plausibilität zu erzeugen oder spezifische Vorgänge der Objekt- und Ereigniswelt in der Vorstellung zu simulieren.“ Der Mensch unterscheidet, differenziert und klassifiziert Gegenstände, Ideen, Emotionen, Einstellungen oder Konstrukte, um nur einige Beispiele zu nennen. Dabei haben durchaus verschiedene Ordnungssysteme nebeneinander ihre Existenzberechtigung. Verantwortlich für das Entstehen solcher Ordnungssysteme sind Lernprozesse. Lernen, so Piaget (1976), ist ein Wechselspiel von assimilativen und akkomodativen Prozessen. Assimilation bezeichnet das Phänomen, dass das Individuum neue Elemente in sein Wissen integriert, ohne dabei sein bestehendes Wissen strukturell in Frage stellen zu müssen. Die Wissensstrukturen können dabei freilich in ihrem Umfang erweitert oder strukturell verfeinert werden. Akkomodation bezeichnet im Gegensatz zur Assimilation die Veränderung bestehender Wissensstrukturen im Sinne einer mehr oder weniger starken Umstrukturierung. Ausgelöst wird sie durch neue Informationen, deren Integration in das bestehende Wissenssystem zu unüberbrückbaren logischen Widersprüchen führen würde.¹

NMDS-Stresswert als Mass für Assimilation und Akkomodation beim Lernen von Wissensstrukturen. In Kapitel 8 (S.119) wird ein Verfahren vorgestellt, mit dem sich Prozesse der Akkomodation und der Assimilation beim Lernen von deklarativem Faktenwissen messen lassen. Die Verfahrenslogik dieser Untersuchung basiert auf der Erkenntnis, dass sich deklaratives Faktenwissen verschiedenen Inhalts mit dem Strukturverfahren der NMDS in kognitiven Karten darstellen lässt (Läge, 2001; Marx & Läge, 1995; Marx & Hejj, 1989; Gigerenzer 1981) und nimmt an, dass sich der Stresswert als Mass für die Konsistenz einer Karte auf das mentale Modell des Lerners als eine Art Entropiemass übertragen lässt. Zwei Lernexperimente, welche diese Annahme stützen, werden berichtet.

Veränderung von Wissenskarten beim Merkmalslernen. In Kapitel 7 (S.93) steht die Art struktureller Veränderungen in Wissensorganisationen, im Blickpunkt des Interesses. Mit fortschreitendem Lernen sollten sich individuelle Karten systematisch verbessern. Das Lernziel ist die Expertise (Annäherung an ein Expertenmodell). Zwei Lernexperimente zeigen den angenommenen Zusammenhang zwischen Merkmalswissen und Güte der Wissenskarte auf. Die zunehmende Konvergenz individueller Wissenskarten untereinander gibt einen ersten Hinweis darauf, wie ähnlich individuelle Wissenskarten einer Zielstruktur werden können. In beiden Experimenten lassen sich die systematischen Verbesserungen der individuellen Wissenskarten belegen.

¹ In der Psychologie kennt man verschiedene Konzepte, die zur Bezeichnung der akkomodativen Veränderung von Wissen verwendet werden: Vosniadou & Brewer (1987) sprechen von einer „radikalen Restrukturierung“, Chi (1992) von einer „radikalen konzeptuellen Veränderung“ und Thagard (1992) beschreibt mit der „konzeptuellen Revolution“ den Prozess bei dem ein konzeptuelles System durch ein anderes ersetzt wird. Die Vielfalt an Auffassungen, was denn die konzeptuelle Veränderung sei, macht die Forderung von Slotta et. al. (1995), nach einer hinreichend präzisen Definition des Konzeptes „konzeptuelle Veränderung“ verständlich. Verschiedene Autoren (z.B. Siegler, 1996) argumentieren, dass Wissen graduell verändert wird. Die konzeptuelle Veränderung setze viele Transformationsschritte voraus, die eine zunehmende Differenzierung kognitiver Strukturen bewirken. Grundlegende Veränderungen konzeptueller Strukturen liefen nur langsam ab und es existierten häufig verschiedenartige Konzeptionen nebeneinander, bis sich eine davon behauptet und durchsetzt. (aus Seel, 200, S. 247)

Wissensstrukturmodell für die Psychopathologie. Eine Replikation dieser Ergebnisse im Feld werden in Kapitel 10 (S.157) vorgestellt.

1.2. Verfahren für die Ähnlichkeitsklassifikation

Im Rahmen der psychologischen Skalierungsverfahren² liegen eine Vielzahl von Vorschlägen zur Ähnlichkeitsklassifikation vor. Das Spektrum reicht dabei von der unmittelbaren paarweisen Ähnlichkeitseinschätzung über kategoriale Sortiervverfahren und dimensionale Rangreihen bis hin zu indirekten Methoden, wie Freien Assoziationen. All diesen Befragungstechniken ist gemeinsam, dass sie in einer Matrix paarweiser Ähnlichkeitswerte zwischen einer zuvor ausgewählten Anzahl von Objekten münden. Diese Matrix wird in der Regel durch Multidimensionale Skalierung so in eine niedrigdimensionale, leicht zu interpretierende geometrische Struktur überführt, dass die Distanzen zwischen den Objekten deren Ähnlichkeitsrelationen bestmöglichst abbilden.

Dass sich merkmalsorientiertes Sachwissen einer Person mit sogenannten „Kognitiven Karten“ darstellen lässt, ist bereits vielfach belegt worden (Läge, 2001; Marx & Hejj, 1989; Marx & Läge, 1995). Als dimensionale Gebilde zeigen sie die Relationen, die eine Person zwischen den Objekten sieht, in Form von kleineren oder grösseren Distanzen auf und lassen als Gesamtbild auch Rückschlüsse auf die strukturierenden (bzw. klassifizierenden) Merkmale zu.

Im Folgenden werden mehrere Verfahren zur Ähnlichkeitsklassifikation vorgestellt. Aus kognitionspsychologischer Perspektive unterscheiden wir dabei zwischen „direkten“ und „indirekten“ Vergleichsoperatoren.

1.2.1. Ähnlichkeitsklassifikation mit Hilfe von „direkten“ Vergleichsoperatoren

1.2.1.1. Direkte Paar-Ähnlichkeitseinschätzung

Das Ähnlichkeitsmass wird in diesem Verfahren durch eine direkte Frage ermittelt. Der Person werden nacheinander Objektpaare mit der Bitte vorgegeben, sie jeweils auf einer mehrstufigen Skala einzuschätzen. An unserem Institut verwenden wir dafür eine Ähnlichkeitsskala von 1-9 (Abbildung 1). Bei einem Satz von n Objekten werden einer Person $(n \cdot (n-1)) / 2$ Paare zur Einschätzung vorgelegt.

1.2.1.2. Paralleles Sortieren

Beim Parallelen Sortieren (Abbildung 2) soll die Versuchsperson, einen Gesamtpool von Objekten in so viele Kategorien aufteilen, wie ihr nach Ähnlichkeitsgesichtspunkten sinnvoll erscheinen. Die Auswahl der Kriterien und die Entscheidung über die Grösse der gebildeten Kategorien obliegen dabei vollständig der befragten Person.

Als Mass für die Ähnlichkeit zwischen zwei Objekten kann die Inzidenz herangezogen werden, d.h. die Frage, ob zwei Objekte von einer Versuchsperson in dieselbe Kategorie sortiert werden oder nicht. Zwei Objekte haben eine Ähnlichkeit 1, wenn sie derselben Kategorie zugeordnet werden und eine Ähnlichkeit 0, wenn sie nicht derselben Kategorie angehören. Auf diese Weise errechnet sich eine Binärmatrix für alle sortierten Objektpaare.

² Borg, 1983; Gigerenzer, 1981; Shepard, Romney & Nerlove, 1972; Sixtl, 1982

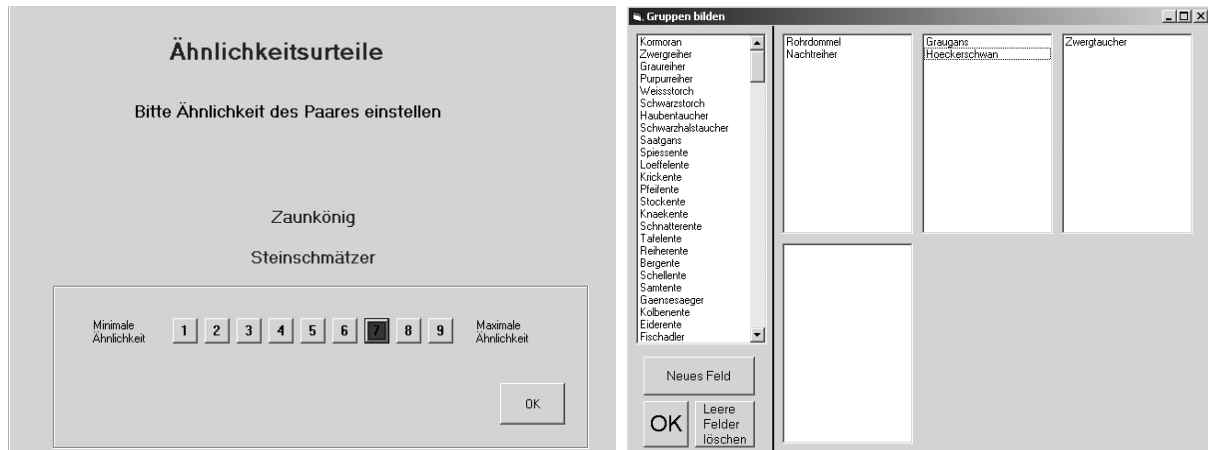


Abbildung 1: Technische Umsetzung für die Direkte Paar-Ähnlichkeitseinschätzung am Computer
Abbildung 2: Erhebungstool für das Parallele Sortieren

1.2.1.3. Hierarchieannahme

Bei der Modellierung von kognitiven Karten (Wissensstrukturkarten) auf Expertenniveau ist folgendes anzufügen: Da wir von hierarchischen Wissenssystemen ausgehen, kann beim Parallelen Sortieren nur ein Zusammenfügen mehrerer individueller Datenmatrizen diesem Hierarchiedanken gerecht werden, da sich Datensätze auf unterschiedlichen Ebenen der Merkmalsdifferenzierung sortieren lassen. Ein einzelner binärskalierter Datensatz hingegen kann Repräsentant, aber nicht Abbild eines hierarchischen Systems sein.

1.2.2. Ähnlichkeitsklassifikation mit Hilfe von Beschreibungsvariablen („indirekte“ Vergleichsoperatoren)

Eine dem behavioristischen Denken verpflichtete Grundidee besagt, dass die Bedeutung eines Begriffs durch die verbalen Reaktionen, die dieser auslöst, gemessen werden kann. In der psychologischen Forschung haben denn auch Listen Freier Assoziationen denn auch eine lange Tradition. Assoziationen lassen sich von verschiedenen Personen zum selben Reizwort erheben und zu einer Assoziationsverteilung zusammenstellen. „Man kann sich [...] eine [...] Assoziations-Verteilung vorstellen als hierarchisch strukturiertes Gebilde etwa im Sinne einer Hullschen Habit-family-Hierarchie. Das Organisationsprinzip dieser Hierarchie ist [dabei] die relative Häufigkeit der einzelnen Elemente.“ Marx (1976). Bei n-befragten Personen kann eine Assoziation maximal n-mal genannt werden. Je häufiger eine Assoziation zu einem Reizwort genannt wird, desto größer ihr Beschreibungswert.

Objektbeschreibungen in Sachbüchern sind in gewisser Weise strukturell mit Assoziationsdaten vergleichbar. Der Autor zählt die ein Objekt beschreibenden Merkmale auf. Zwei dokumentierte Objekte lassen sich über ihre Listen von Eigenschaftsausprägungen im Prinzip genau so vergleichen wie Assoziationsleistungen. Für die Berechnung von Ähnlichkeiten zwischen Mengen von Beschreibungsvariablen gibt es eine ganze Reihe von Vorschlägen.

1.2.2.1. Ähnlichkeitsmasse im Information Retrieval

Für die Berechnung von Ähnlichkeiten zwischen Mengen von Beschreibungsvariablen gibt es eine ganze Reihe von Vorschlägen. Im *Information Retrieval* arbeitet man bei der Bestimmung von Ähnlichkeiten vor allem mit Vektorraummodellen. Dokumente, aber auch Anfrage-terme, lassen sich als Vektoren darstellen. Dazu steht eine Vielzahl von Algorithmen zur Verfügung. Die Ähnlichkeit zwischen zwei Dokumenten (oder einem Dokument und einer An-

frage) lässt sich über das *Skalarprodukt* der zwei Dokumentenvektoren d_i und d_j bestimmen (wobei d_{ki} und d_{kj} der k -te Term des Vektors d_i bzw. d_j darstellen). Eine andere Möglichkeit ist es, den Winkel zwischen den beiden Vektoren als Mass für die Ähnlichkeit heranzuziehen (*Cosinusmass*). Zur Verbesserung der Retrievalleistung werden zusätzlich verschiedene Merkmalsgewichtungen, auf die wir hier nicht näher eingehen wollen, vorgeschlagen (Salton & Buckley, 1988; Salton & Allan, 1993). Dabei wird jedem Repräsentationsmerkmal im Dokumentenvektor (oder Anfragevektor) ein Gewicht zugeordnet. Dabei repräsentiert w_{ki} die Gewichtung des k -ten Terms (t_k) im i -ten Dokument (d_i). Im einfachsten Fall besteht die Gewichtung eines Merkmals k in seiner Auftretenshäufigkeit f in Dokument i . In diesem Falle gilt: $w_{ki} = f_{ki}$.

Formale Definitionen:

$D = \{d_1, \dots, d_i, d_j, \dots, d_n\}$	mit $1 \leq i \leq n$	eine Menge von Dokumenten
$T = \{t_1, \dots, t_k, \dots, t_m\}$	mit $1 \leq k \leq m$	eine Menge von Termen
1		Häufigkeit (frequency) des Terms k in Dokument i , absolute Häufigkeit.
w_{ki}		Gewichtung (weight) des Terms k (t_k) in Dokument i (d_i). Im einfachsten Fall besteht die Gewichtung eines Merkmals k in seiner Auftretenshäufigkeit f in Dokument i , sprich $w_{ki} = f_{ki}$.
$\sum_{k=1}^m f_{ki}$		Anzahl Terme im Dokument i
$p_{ki} = \frac{f_{ki}}{\sum_{k=1}^m f_{ki}}$		relative Häufigkeit (possibility) des Terms k in Dokument i
$f_{kij} = \min(f_{ki}, f_{kj})$		gemeinsame, absolute Häufigkeit (joint frequency) des Terms k in den beiden Dokumenten i und j
$p_{kij} = \min(p_{ki}, p_{kj})$		gemeinsame relative Häufigkeit (joint possibility) des Terms k in den beiden Dokumenten i und j

Ergebnisse mittels *Skalarprodukt* weisen eine starke Abhängigkeit von der Anzahl und Grösse der einzelnen Werte in den Vektoren auf, was gleichbedeutend mit der Länge der Dokumente ist. Würde man (unabhängig von der Gewichtung der einzelnen Attribute) den Text eines Dokumentes einfach verdoppeln, so würde sich aufgrund der Linearität auch die Ähnlichkeit der Dokumente verdoppeln, obwohl kein zusätzlicher Informationsgewinn vorliegt.

Das *Cosinusmass* ist unabhängig von der Länge. Das Mass gibt lediglich die Richtung der Vektoren an. Zwei Dokumente sind sich am ähnlichsten, wenn ihre Vektoren in dieselbe Richtung zeigen.

Skalarprodukt

$$|X \cap Y| \quad \text{oder} \quad s_{ij} = \sum_{k=1}^m w_{ki} * w_{kj}$$

Cosinusmass

$$\frac{|X \cap Y|}{|X|^{\frac{1}{2}} * |Y|^{\frac{1}{2}}} \quad \text{oder} \quad s_{ij} = \frac{\sum_{k=1}^m w_{ki} * w_{kj}}{\sqrt{\sum_{k=1}^m w_{ki}^2} * \sqrt{\sum_{k=1}^m w_{kj}^2}}$$

Der Vollständigkeit halber sind hier drei weitere gebräuchliche Masse wie das Dice-, das Jaccard- und das Overlap-Mass zu erwähnen.

Dice³

$$\frac{2|X \cap Y|}{|X| + |Y|} \quad \text{oder} \quad s_{ij} = \frac{2 * \sum_{k=1}^m w_{ki} * w_{kj}}{\sum_{k=1}^m w_{ki} + \sum_{k=1}^m w_{kj}}$$

Jaccard⁴

$$\frac{|X \cap Y|}{|X \cup Y|} \quad \text{oder} \quad s_{ij} = \frac{\sum_{k=1}^m w_{ki} * w_{kj}}{\sum_{k=1}^m w_{ki} + \sum_{k=1}^m w_{kj} - \sum_{k=1}^m w_{ki} * w_{kj}}$$

Overlap

$$\frac{|X \cap Y|}{\min(|X|, |Y|)} \quad \text{oder} \quad s_{ij} = \frac{\sum_{k=1}^m \min(w_{ki}, w_{kj})}{\min\left(\sum_{k=1}^m w_{ki}, \sum_{k=1}^m w_{kj}\right)}$$

³ Dice (1945), ursprünglich entwickelt für binäre Datensätze.

⁴ Achtung Fehler bei Bortz (1999): Jaccard (1908) \neq Rogers & Tanimoto (1960)

Sowohl bei Dice als auch bei Jaccard variiert die Ähnlichkeit in Abhängigkeit von der Relation der Länge der Vektoren. Für eine genauere Beschreibung der Vor- und Nachteile dieser Masse sei jedoch auf Ferber (2003) verwiesen.

1.2.2.2. Überlappungskoeffizient entwickelt für Assoziationsdaten

Marx (1976a) hat als quantitatives Mass für den Vergleich von Assoziationsverteilungen den Überlappungskoeffizienten⁵ vorgeschlagen.

Enthalten zwei Assoziationsverteilungen keinerlei gemeinsame Assoziationen, so wird der ÜK = 0. Sind beide Verteilungen identisch, beträgt der Wert ÜK = 1. Relative Häufigkeiten einer jeden Assoziation j innerhalb von A und B werden hier separat gerechnet. Dies hat die beiden grossen Vorteile, dass das Ähnlichkeitsmass auch dann anwendbar ist, wenn beide Verteilungen a) unterschiedlich umfangreich sind und b) nicht-gemeinsame Assoziationen enthalten.

Verteilung A	Verteilung B	gemeinsame rel. Häufigkeit (p_{kij})
a .00	a .26	.00
b .28	b .00	.00
c .42	c .37	.37
d .05	d .07	.05
e .09	e .19	.09
f .09	f .04	.04
g .07	g .07	.07
ÜK = $\sum p_{kij}$		= 0.62

Tabelle 1: Beispiel für die Berechnung des Überlappungskoeffizienten (ÜK), aus Marx (1976a)⁶

Überlappungskoeffizient⁷

$$s_{ij} = \sum_{k=1}^m p_{kij} = \sum_{k=1}^m \min(p_{ki}, p_{kj})$$

Erste Expertenmodelle. In Kapitel 9 (S.143) wird ein Expertenmodell für das Wissensgebiet der Psychopathologie vorgestellt. Dieses Modell wurde in Zusammenarbeit mit Psychiatrischen Fachkräften über Direkte Paar-Vergleiche erhoben und zeigt die Möglichkeit auf, streng kategorialen Wissenssystemen aus der Psychopathologie, wie die ICD-10 oder die DSM-IV, ein Modell gegenüberstellen zu können, welches neben klassifikatorischen auch dimensionale Aspekte psychischer Erkrankungen zu berücksichtigen und abzubilden vermag.

Rolle von Gewichtungsfaktoren bei Parallelen Sortierdaten im Hinblick auf spätere NMDS Lösungen. In Kapitel 3 (S.71) zeigen wir, dass sich ein Gewichten individueller Binärmatrizen aus dem Parallelen Sortieren, und zwar proportional zur gebildeten Kategorienzahl, einen po-

⁵ Achtung: Nicht zu verwechseln mit dem oben aufgeführten Overlap-Mass.

⁶ In der Begriffsverwendung angepasste Formel. Beispiel aus Marx (1976a), in korrigierter Fassung.

⁷ Aus der Tradition der 50er und 60 Jahre existieren eine Reihe weiterer Ähnlichkeitskoeffizienten für Assoziationsverteilungen (Bousfield, Whitmarsh & Danick, 1958; Deese 1959, Rothkopf 1960; Marshall & Cofer, 1963; Garskof & Houston, 1963; Garskof 1965) und für Primärassoziationen (Jenkins & Cofer, 1957), die jedoch alle wesentliche Schwachpunkte aufweisen. Die Koeffizienten sind bei Marx (1976b) ausführlich diskutiert. Der Überlappungskoeffizienten wurde dabei gegenüber diesen Koeffizienten statistisch abgesichert.

sitiven Effekt auf Teilstrukturen eines grösseren ganzheitlicheren Expertensystems haben kann. Die Überprüfung dieses Effekts beruht nicht zuletzt auf der Annahme kognitiver Parallelität derjenigen Experten, welche am Aufbau der Struktur beteiligt waren. Die Annahme, Personen mit gleichem Expertisegrad würden, angesprochen darauf eine Reihe von Objekten nach einem bestimmten Merkmal zu sortieren, auf ein und dasselbe Wissenssystem, wenn auch auf unterschiedlichen Beschreibungsebenen, zugreifen, darf mit einer hohen Wahrscheinlichkeit angenommen werden (= teilweise Bestätigung der Hierarchieannahme).

Modellierung von Klassifikatorischem Expertenwissen. In Kapitel 2 (S.15) werden drei psychologische Erhebungsverfahren für die Ähnlichkeitsklassifikation einander gegenübergestellt. Hauptanliegen dieser Studie war es, eine effizientere Methode der Expertenbefragung denn die direkter Paarvergleiche zu finden, denn der Erhebungsaufwand steigt bei der „Direkten Paar-Ähnlichkeitseinschätzung“ mit zunehmender Objektmenge logarithmisch. Sowohl das Verfahren „Paralleles Sortieren“, als auch die „Ähnlichkeitsklassifikation mit Hilfe von Beschreibungsvariablen“ sind, mit Einschränkungen, für die Erhebung von Wissensstrukturkarten durchaus vertretbare Alternativen. Sie zeigen untereinander ausserordentlich hohe strukturelle Übereinstimmung.

Kognitive Adäquatheit von Wissensstrukturkarten. In Kapitel 4 (S.43) interessieren wir uns für die Organisiertheit von Wissensstrukturkarten im Allgemeinen und im Zusammenhang mit der Möglichkeit, Wissensstrukturkarten als Hilfsmittel bei der Vermittlung von Merkmalswissen einzusetzen. Wissensstrukturkarten können mehr oder weniger prägnant sein. Prägnant ist eine Wissensstrukturkarte dann, wenn sie in ihrer strukturellen Komplexität der Auffassungsgabe eines Menschen angepasst ist, wenn dieser sie also in ihren Grundzügen lesen und deuten kann. Diese Hinordnung von Prägnanz auf das individuell Intelligible bezeichnen wir als „Kognitive Adäquatheit“. Und diese ist abhängig von allgemeiner Intelligenz und fachspezifischer Expertise. Für einen Laien, dies die Überlegung, müsste eine Karte umso funktionaler und also umso prägnanter sein, je einfacher sie strukturiert ist. Daher wurde ein Vorschlag zur Vereinfachung von schwach klassifikatorischen Systemen gemacht. Exemplarisch wurden ornithologische Wissensstrukturkarten auf ihre Funktionalität untersucht. Die klassifikatorischsten Karten waren für unsere ornithologischen Laien dabei nur selten auch die prägnantesten. Die Prägnanz einer Karte, zumindest dies wird deutlich, lässt sich letzten Endes immer nur im Experiment bestimmen.

1.3. Wissensstrukturmodelle & IT-Systeme

Aus pädagogischer Sicht stellt sich die Frage, mit welchen didaktischen Möglichkeiten der Lerner in der Ausbildung seines eigenen Wissensmodells unterstützt werden kann. Der Lernfortschritt, so die These, ist dann am besten, wenn der zu lernende Stoff einen bestimmten Neuheitswert hat (dadurch spannend ist) und nicht überfordert. Um diesem Anspruch genüge leisten zu können, muss ein Tutor oder ein Tutorsystem mit dem momentanen Wissensstand des Lerners genau vertraut sein. Wissenslücken oder fehlerhaftes Wissen müssen bekannt sein. Nur auf dieser Informationsgrundlage kann die Vermittlung von Wissensinhalten genau dort ansetzen, wo dem Lerner am meisten geholfen ist. In letzter Konsequenz wird hier der Wunsch formuliert, den lästigen und mühsamen Werdegang des Laien zum Experten zu vereinfachen und effizient zu gestalten.

Intelligente Tutorielle Systeme⁸ (ITS) bilden einen Ansatz zur Gewährleistung, der an den Wissensstand des Lerners angepassten schulischen Bildung, wobei die Idee von Lehrmaschinen nicht neu ist. Leutner (1992a) stellt in seinem Buch „Adaptive Lehrsysteme“ einen historischen Abriss über die Entwicklung solcher Unterrichtsformen dar, welche, angefangen mit ersten Ansätzen in den zwanziger Jahren des letzten Jahrhunderts, mit Maschinen zur „Intelligenzprüfung und Informationsvermittlung“ (vgl. Pressey, 1926) und deren Weiterentwicklung durch Skinner (1958) und Crowder (1959) in den sechziger Jahren in den ersten computerisierten Systemen mündeten (CAI: computer assisted instruction).

Mit dem Aufkommen der Künstlichen Intelligenz wurden an die CAI's neue Anforderungen gestellt. Die nun als Intelligente Tutorielle Systeme bezeichneten Programme sollen sich dem Instruktionsverhalten einer menschlichen Lehrkraft angleichen (vgl. Goldstein, 1982), d.h. sie sollen über Lehrstoffexpertise verfügen. Erst der Vergleich zwischen dem IST-Zustand des Lerners und dem SOLL-Zustand des Expertenmodells führen zu „intelligentem Verhalten“ des Systems und bilden die Basis für einen ökonomischen und individuell angepassten Lernvorgang (vgl. zu den Komponenten eines ITS: Van Merriënboer, Van Den Berg & Maaswinkel, 1989).

Ein derartiger „adaptiver Unterricht“ muss zwei Zielen genügen: Zum einen der Fokussierung auf die Gestaltung didaktisch optimal aufbereiteter, multimedialer Unterrichtsinhalte, zum anderen der überaus wichtigen, verlässlichen und effizienten Statusdiagnose als Grundlage für die Erstellung der kommenden Lektion.

Eine lernprozessbegleitende Diagnose des Wissens führt jedoch zum sogenannten „Testlängendilemma“, auf das bereits Leutner (1992b, 1993) hingewiesen hat: Auf der einen Seite soll eine möglichst umfassende und genaue Diagnose erstellt werden, um bei der Zusammenstellung der folgenden Lektionen so adaptiv wie möglich sein zu können. Auf der anderen Seite will man aber nicht zuviel Zeit und kognitive Ressourcen der Lernenden für die Diagnose beanspruchen, weil beides von der verbleibenden Lernzeit abgeht. Die Adaptivität im Inhalt gewinnt ihre Relevanz dadurch, dass zum einen Mehrfachlernen weitgehend vermieden und zum anderen falsches Wissen (welches über einen selbstgesteuerten Lernprozess nur schwerlich detektiert wird, vgl. Charney, Reder & Kusbit, 1990 und Strohschneider, 1990) ausgemerzt werden kann.

Das Effizienzproblem in der Wissensüberprüfung wird bei der Diagnostik von merkmalsbezogenem Sachwissen besonders deutlich. Geht es um die unzähligen, für die Qualität von Expertenentscheidungen aber zwingend notwendigen Fakten, müsste für eine zuverlässige Diagnose der gesamte Wissensstand innerhalb eines Fachgebietes mit Hunderten von Fragen abgetestet werden. Nur so wäre man in der Lage, Lücken und Falschwissen einer Person zu detektieren, was in der Praxis ein nicht durchführbares Unterfangen darstellt.

Klein (2000) resümiert in diesem Zusammenhang die Schwierigkeiten einer solchen Adaptivität. Er führt aus, dass die Adaptierbarkeit (d.h. die Anpassbarkeit des Systems durch den Nutzer, wie z.B. die Wahl unterschiedlicher Bildschirmarrangements oder das An- und Ausschalten von Hilfefunktionen) technisch wie konzeptionell wesentlich leichter zu lösen ist als die inhaltliche Adaptivität des Lernmaterials (vgl. auch Leutner, 1992a, 1995). In der Tat be-

⁸ Auf definitorische Unterscheidungen zwischen den unzähligen und zum Teil gleichbedeutenden Bezeichnungen für (virtuelle) Lehrsysteme, wie bspw. ITS (Intelligente Tutorielle Systeme), ATS (Adaptive Tutorielle Systeme) CBT (Computer Based Training) usw. wird hier nicht weiter eingegangen. In der vorliegenden Arbeit wird der etwas allgemeinere und geläufigere Terminus ITS verwendet und meint allgemein, dass sich das Lehrsystem an die individuellen Bedürfnisse des Lernenden anpassen kann. Im Speziellen bedeutet dies für diese Arbeit, dass das System auf individuelle Wissensdefizite entsprechend didaktisch und inhaltlich reagiert.

schränken sich die heutigen ITS-Systeme weitgehend auf die Vermittlung regelbasiertem Wissen, Wissen also, welches sich mit Wenn-Dann-Regeln charakterisieren lässt und vielfach durch einen kleinen Satz an Entscheidungs- und Transformationsregeln abgedeckt werden kann (zu Umsetzungs-Beispiele für Mathematik- und Physikwissen vgl. Albert & Lukas, 1999 und Siemer & Angelides, 1998; für einen erweiterten Überblick über bestehende ITS-Systeme vgl. Daub, 2001).

1.3.1. Wissensstrukturkarten und Wissensdiagnostik

Die Forschergruppe „Angewandte Kognitionspsychologie“ der Universität Zürich, hat es sich zum Ziel gesetzt, Wissensstrukturkarten bei der Diagnostik von Merkmalsbasiertem Sachwissen einzusetzen. Im Rahmen des von 2001 bis 2004 vom Schweizerischen Nationalfonds geförderten Projektes MODAL (Merkmalsorientierte Diagnostik in adaptiven Lernsystemen) wurde das folgende Modell vorgeschlagen (Abbildung 3):

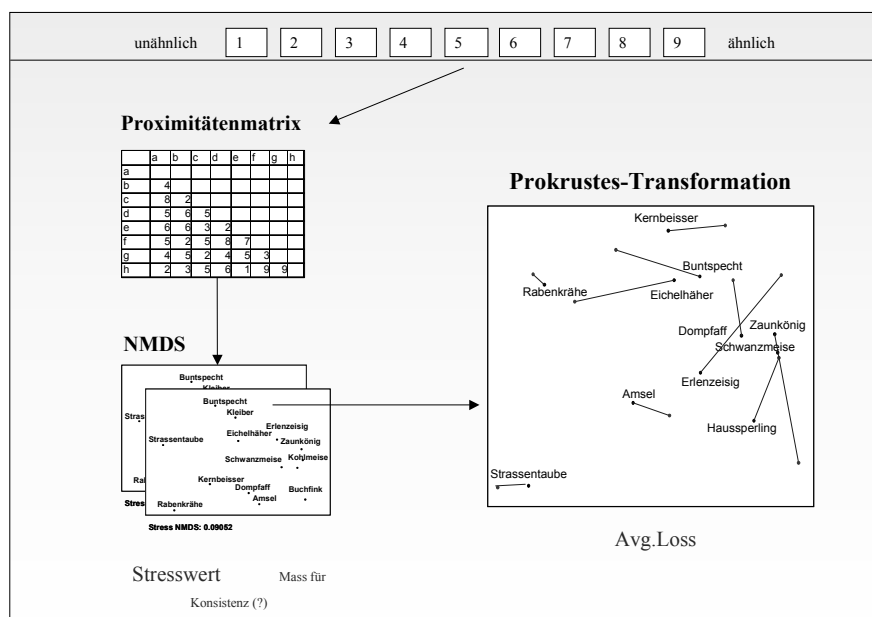


Abbildung 3: Verfahrenslogik des Zürcher-Modells: Über „Direkte Ähnlichkeitseinschätzung“ wird das Lernerwissen erhoben. Mit dem Skalierungsverfahren der NMDS werden zweidimensionale Karten berechnet, die das individuelle Lernermodell abbilden. Über einen Vergleich mit dem Systeminternen Expertenmodell werden fehlplatzierte Objekte (= falsches oder ungenaues Merkmalswissen) erkannt.

Die Merkmalsbasierte Wissensdiagnostik des Zürcher-Modells basiert auf der Idee, individuelle Lernerkarten mit systeminternen Expertenstrukturen zu vergleichen. Lernergenerierte Paarähnlichkeiten zu einem bestimmten Objektsatz werden mittels der Nonmetrischen Multidimensionalen Skalierung RobuScal⁹ zweidimensional skaliert (zum Verfahren der NMDS vgl. Borg & Groenen, 1997). Die zu diagnostizierenden Objekte sind dabei als Punkte in der Karte abgebildet. Ähnliche Objekte liegen nahe beieinander, währenddem unähnliche Objekte weit auseinander zu liegen kommen. Mit dem Expertenmodell wird das Lernermodell nun über eine Prokrustes-Transformation verglichen (Gower & Dijksternhuis, 2004). Das Lernermodell wird so lange gedreht und gestreckt, bis es optimal über das Expertenmodell passt. Die

⁹ Der NMDS-Algorithmus RobuScal erlaubt die Interpretation von NMDS-Lösungen auf individueller Ebene, denn mathematische Korrekturverfahren sorgen für stabile und gegen Ausreisser in den Daten robuste Konfigurationen (Läge, Daub, Bosia, Ryf & Jäger, in review).

gefundenen Abweichungen zwischen IST und SOLL geben dann Hinweise auf die Wissensdefizite im Kenntnisstand des Lernalers (quantifiziert wird über den so genannten *Average Loss* (AvgLoss) = Gesamtabweichung oder über den *Object Loss* (ObjLoss) = Abweichung jeden einzelnen Objektes (mehr zu AvgLoss/ObjLoss in Läge, 2001)).

Die erläuterte Verfahrenslogik wird in der Dissertationsarbeit von Streule (2005) weiterführend behandelt. In seiner Arbeit werden Grenzen und Möglichkeiten des Verfahrens ausgelotet, und Erweiterungen und Spezifizierungen des Diagnostikkonzepts vorgeschlagen.

1.3.2. Lernen mit Wissensstrukturkarten

Wissen in der Form von Wissensstrukturkarten darzustellen, ist bestimmt keine Neuigkeit. MindMaps sind die wohl bekanntesten Beispiele dafür, Wissen in geordneter Form darzustellen. Gerade im Zeitalter von Internet ist man an der Visualisierung von Wissenszusammenhängen, also am Sichtbar machen von Struktur, sehr interessiert. Von der Visualisierung der Datenflut in einem geeigneten Format erhofft man sich zum Beispiel eine Verbesserung der Benutzerfreundlichkeit im Dokumenten- oder generell, im Information Retrieval. Ein eindrückliches Beispiel einer solchen Informationsvisualisierung kann mit DocMINER¹⁰ gegeben werden. DocMINER arbeitet mit der Metapher von „Dokumentenlandkarten“. Technisch integriert DocMINER dabei „statistisch-linguistische“ Verfahren¹¹, leistungsfähige Dimensionsreduktionsverfahren und eine Visualisierungstechnologie auf der Basis von neuronalen Netzen. DocMiner ermöglicht es dem Nutzer erstmals, eine Dokumentendatenbank in seinem ganzen Umfang zu explorieren, ohne dabei den Überblick zu verlieren. Die einzelnen Dokumente werden dem Nutzer als Punkte in einer zweidimensionalen Karte präsentiert. Dokumente ähnlichen semantischen Inhalts liegen nahe beieinander, unterschiedliche Dokumente weit auseinander. Die Struktur repräsentiert ein Klassifikationssystem semantischer Ähnlichkeiten. Beim normalen Retrieval besteht das Problem, dass der Nutzer auch bei einer erfolgreichen Anfrage immer nur einen winzigen Ausschnitt präsentiert bekommt und nie weiss, inwieweit die Gesamtmenge an Dokumenten seinem tatsächlichen Bedürfnis entspricht. Anstatt dem Benutzer nur auf Anfrage Information in kleinen Häppchen zu geben, wird bei DocMINER als Ausgangspunkt die graphische Darstellung inhaltlicher Zusammenhänge direkt präsentiert. Er kann nun die Daten zunächst als Ganzes aus verschiedenen Blickwinkeln betrachten und dann durch Selektion von Teilbereichen auf bestimmte Ausschnitte der Daten fokussieren. Auf diese Weise ist eine explorative Analyse der Daten möglich.

Da Wissenskarten Beziehungen, Zusammenhänge und Abhängigkeiten von Sachverhalten sichtbar machen, sind sie aus pädagogischer Sicht äusserst interessant.

Lernen mit Wissensstrukturkarten. Kognitive Karten sollten sich besonders gut dazu eignen um den Wissensstrukturaufbau zu unterstützen und die Integration von Faktenwissen zu beschleunigen. Für die Vermittlung von ornithologischem Merkmalswissen wurden in einem Experiment (Kapitel 5, S.71) Wissensstrukturkarten eingesetzt. Die Ergebnisse zeigen, dass es sich tatsächlich lohnt, dem Lerner Kognitive Karten (Wissenskarten) zu präsentieren. Dabei ist zu vermuten, dass die assimilativen Effekte um so grösser sind, je mehr der Lerner bereits

¹⁰ **DocMINER** ist ein Kooperationsprojekt der folgenden Institute: Fraunhofer Institut für Medienkommunikation, Abteilung für künstliche Intelligenz Universität Dortmund & Fraunhoferinstitut für Arbeitswirtschaft und Organisation. Kurze Projektbeschreibung, Becks & Host (2000a). Weitere vergleichbare Ansätze zum Dokumenten Retrieval, welche auf die Metapher von Dokumentenlandkarten zurückgreifen, sind in Becks (2001, S. 51) ausführlich beschrieben.

¹¹ Zur autonomen maschinellen Erstellung von Wissenskarten aufgrund von semantischer Beziehung zwischen Text-Dokumenten kann hier auf die Methoden des Dokumenten-Clusterings der semantischen Netzwerke von Lin et. al. (1991), Kohonen et al. (2000), Becks et. al. (2000b), Becks (2001) verwiesen werden.

weiss, je mehr er also die klassifikatorischen Möglichkeiten eines solchen Kognitiven Modells ausnutzen kann. Die Ergebnisse deuten stark darauf hin, dass sich Kognitive Karten in die Reihe der lernfördernden Kognitiven Modelle einreihen lassen.

DIVE IN – adaptiver Unterricht mit hierarchisch ineinandergreifenden Wissensstrukturkarten. Mit DIVE IN (Kapitel 6, S.81) wird ein Verfahren vorgeschlagen, welches die hierarchisierten Realisierungen im kartenbezogenen Information Retrieval (z.B. Doc Miner) mit dem Anspruch adaptiver Lernsysteme verknüpft. Sie verbindet didaktische Forderungen nach aktivem explorativem Lernen mit knallhartem Feedback hinsichtlich des gezeigten Verhaltens.

1.4. Literatur

- Alber, D. & Lukas, J. (Eds.), (1999). *Knowledge spaces. Theories, empirical research and applications*. Mahwah: Erlbaum.
- Becks, A. & Host, M. (2000a). *Visuell gestütztes Retrieval mit Dokumentenlandkarten*. www.doculine.com/news/2000/0700/VisuellRetrieval.htm (zuletzt besucht Juni 2004).
- Becks, A., Sklorz, S., Jarke, M.: A (2000b). *Modular Approach for Exploring the Semantic Structure of Technical Document Collections*, ACM Proc. Of AVI 2000, Palermo.
- Becks, A. (2001). *Visual Knowledge Management with Adaptable Maps*. Dissertationsschrift an der Technischen Hochschule Aachen.
- Borg, I. (1983). Scaling: A review of the German scaling literature of the last fifteen years. *The German Journal of Psychology*, 7, 63-79.
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling – Theory and applications*. New York: Springer.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. Auflage). Berlin: Springer.
- Bousfield, W.A., Whitmarsh, G.A. & Dannick, J.J. (1958). Partial response identity in verbal generalizations. *Psychological Report*, 4, 703-713.
- Charney, D., Reder, L. & Kusbit, G.W. (1990). Goal setting and procedure selection in acquiring computer skills. A comparison of tutorials, problem solving and learner exploration. *Cognition and Instruction*, 7, 323-342.
- Chi, M.T.H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In Giere, R.N. (Ed.), *Minnesota studies in the philosophy of science. Vol. XV. Cognitive models of science*. University of Minnesota Press, Minneapolis, 129-186.
- Crowder, N.A. (1959). Automating tutoring by means of intrinsic programming. In E. Ganter (Ed.), *Automatic teaching. The state of the art* (pp.109-116). New York: John Wiley and Sons.
- Daub, S. (2001). *Similarity Simulation – Wie man den Code des Globalurteils knackt*. Dissertationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Report*, 5, 305-312.
- Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26, 297-302.
- Ferber, Reginald (2003). *Information Retrieval. Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web*. Heidelberg: dpunkt-Verlag.
- Garskof, B.E. 1965). Relation between single word association and continued association response hierarchies. *Psychological Reports*, 16, 307-309.
- Garskof, B.E. & Houston, J.P. (1963). Measurement of verbal relatedness: An idiographic approach. *Psychological Review*, 70, 277-288.

- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhard.
- Goldstein, I.P. (1982). The genetic graph: A representation for the evolution of procedural knowledge. In D. Sleeman & J.S. Brown (Eds.), *Intelligent tutorial systems* (pp. 51-78). New York: Academic Press.
- Gower, J.C. & Dijksterhuis, G.B. (2004). *Procrustes Problems*. Oxford: Oxford University Press.
- Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Büll. Soc. Vaud. Sci. nat.* 44, 223-270.
- Jenkins, P.M. & cofer, C.N. (1957). An exploratory study of discrete free association to compound verbal stimuli. *Psychological Reports*, 3, 599-602.
- Klein, B. (2000). *Didaktisches Design hypermedialer Lernumgebungen. Die adaptive Lernumgebung „incops“ zur Einführung in die Kognitionspsychologie*. Marburg: Tectum.
- Kohonen, T., Kaski, S. et al. (2000). Self Organization of a Massive Document Collection. *IEEE Transactions on Neural Networks*, Vol. 11, No. 3.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophische Fakultät der Universität Zürich.
- Läge, D., Daub, S., Bosia, L., Ryf, S. & Jäger, C. (in review). *Robustness in nonmetric mulidimensional scaling*.
- Leutner, D. (1992a). *Adaptive Lehrsysteme. Instruktionspsychologische Grundlagen und experimentelle Analysen*. Weinheim: Psychologie Verlags Union.
- Leutner, D. (1992b). Das Testlängendilemma in der lernprozess-begleitenden Wissensdiagnostik. *Zeitschrift für Pädagogische Psychologie*, 24, 232-248.
- Leutner, D. (1993). Das gleitende Testfenster als Lösung des Testlängendilemmas: Eine Robustheitsstudie. *Zeitschrift für Psychologie*, 7, 33-45.
- Leutner, D. (1995). Adaptivität und Adaptierbarkeit multimedialer Lehr- und Informationssysteme. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia*. Weinheim: Psychologie Verlags Union. S. 140-149.
- Lin, X., Soergel, D., Marchionini, G., (1991): *A self-organizing semantic map for information retrieval*. In Proc. Of 14th ACM/SIGIR conf. Research and Development in Information Retrieval.
- Marshall, G.R. & Cofer, C.N. (1963). Associative indices as measures of word relatedness: A summary and comparison of ten methods. *Journal of verbal Learning and verbal Behaviour*, 1, 408-421.
- Marx, W. (1976a). Die Messung der Assoziativen Bedeutungsähnlichkeit. *Zeitschrift für experimentelle und angewandte Psychologie*, 1, 62-76.
- Marx, W. (1976b). Die statistische Sicherung des Überlappungs-Koeffizienten. *Zeitschrift für experimentelle und angewandte Psychologie*, 23, 267-270.
- Marx, W. & Hejj, A. (1989). *Subjektive Strukturen*. Göttingen: Hogrefe.
- Marx, W. & Läge, D. (1995). *Der ideologische Ring*. Göttingen: Hogrefe.
- Pressey, S.L. (1926). A simple apparatus which gives tests and scores and teaches. In (reprint) A. A. Lumsdaine & R. Glaser (Eds.), (1969). *Teaching machines and programmed learning* (pp 35-41). Washington D.C.: National Education Association of the United States.
- Rogers, D.J. & Tanimoto, T.T. (1960). A computer programm for classifying plants. *Science* 132, 1115-1118.
- Rothkopf, E.Z. (1960). Two predictors of stimulus equivalence iin paired associate learning. *Psychological Reports*, 7, 241-250.
- Salton, G. & Allan, J. (1993). Selective Text Utilization and Text Traversal. In *The fifth ACM Conference on Hypertext Proceedings* (pp. 131-144). New York: Academic Press.

- Salton, G. & Buckley, C. (1988). Term weighting approaches in Automatic Text Retrieval. *Information Processing and Management*, 24, 513-523.
- Seel N.M. (2000). *Psychologie des Lernens*. München: Ernst Reinhard.
- Shepard, R.N., Romney, K.A. & Netherlove, S.B. (Eds.), (1972). *Multidimensional Scaling: Theory and applications in the Behavioural Science*. New York: Seminar Press.
- Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- Sixtl, F. (1982). *Messmethoden der Psychologie*. Weinheim: Beltz.
- Skinner, B.F. (1958). Teaching machines. In (reprint) A. A. Lumsdaine & R. Glaser (Eds.), (1960). *Teaching machines and programmed learning* (pp. 137-158). Washington D.C.: National Education Association of the United States.
- Slotta, J.D., Chi, M.T.H. Joram, E. (1995). Assessing students misclassifications of physics concepts. An ontological basis for conceptual change. *Cognition and Instruction*, 13, 373-400.
- Streule, R. (2006). *Sag mir, was ich noch nicht weiss - Individualisierung des Wissenserwerbs über adaptive Auswahl der Lerninhalte mit Hilfe der Nonmetrischen Multidimensionalen Skalierung*. Dissertationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Strohschneider, S. (1990). *Wissenserwerb und Handlungsregulation*. Wiesbaden: Deutscher Universitätsverlag.
- Thagart, P. (1992). *Conceptual revolutions*. Princeton University Press, Princeton NJ.
- Van Merriëboer, J.J.G., VanDen Berg, K.G. & Maaswinkel, D.M. (1989). Some experiences with two intelligent tutoring systems for teaching computer programming: PROUST and LISP-tutor. In J. M. Pieters (Ed.), *Intelligent tutorial systems and instruction* (pp. 7-44). Enschede: OTG Onderwijsleerprocessen.
- Vosniadou, S. & Brewer, W.F. (1987). Theories of knowledge restructuring in development. *Review of educational Research*, 54, 51-67.

2.

**Die Modellierung von
Klassifikatorischem Expertenwissen am Fallbeispiel
der Ornithologie**

2.1. Einleitung

2.1.1. Einführung ins Thema

Der Mensch unterscheidet und differenziert Gegenstände, Ideen, Emotionen, Einstellungen, u.s.w. Der Mensch organisiert und klassifiziert die Merkmale seiner Umwelt.

Die aristotelische *naturalistische Einstellung*, die von Autoren wie Simpson (1967) oder Mayr (1975) vertreten wird, nimmt an, dass Kategorien reale Einheiten der Welt darstellen und wir Menschen unser Bestes tun, um sie und die Gegenstände, die unter sie fallen, zu erfassen. Die kantianisch geprägte *handlungstheoretische Einstellung* geht hingegen von der Annahme aus, dass Kategorien Schöpfungen unseres Denkens sind (Janich, 1992; Janich & Weingarten, 1999; Herkber, 1999). Wir gestalten die Wirklichkeit so, wie wir sie erfahren, und erfahren sie so, wie sie unseren Kategorien entspricht. Der Kantianer zieht aus dem Fehlen von Konformität zwischen unseren geschlossenen Begriffssystemen einerseits und der offenen, verschwommenen Welt andererseits die Konsequenz, dass wir die Welt selbst (die Welt „an sich“) mit unseren Begriffen nie erfassen können. Das heisst, sobald wir beim Erfassen eines Dinges sind, wird das Ding, das wir zu erfassen versuchen, quasi mit einem erkenntnistheoretischen Schleier versehen. Das Einzige, das wir erkennen können, sind begrifflich geformte Erkenntnisobjekte. Die Gegenstände selbst bleiben immer hinter dem von uns erzeugten Schleier.

Erkenntnisse aus Laborstudien über Farbbezeichnungen von Berlin und Kay (1991) unterstützen die Kantianische Hypothese. Drei von ihnen beschriebene Eigenschaften von Kategorien (zumindest die Farbkategorien) sprechen gegen die traditionelle Auffassung. 1. Farbkategorien zeigen keine deutlichen Grenzen, sondern haben unscharfe Ränder. 2. Nicht alle Referenzobjekte einer Kategorie besitzen denselben Status, sondern es gibt privilegierte (typische) Exemplare und 3. Kategorien bestehen nicht a priori, sondern sind durch die Konstitution des perzeptiven und kognitiven Apparates des Menschen bedingt – beispielsweise durch den Umstand, dass die Farbrezeptoren des Auges bei bestimmten Wellenlängen des Lichts ein Maximum an Stimulation erfahren.

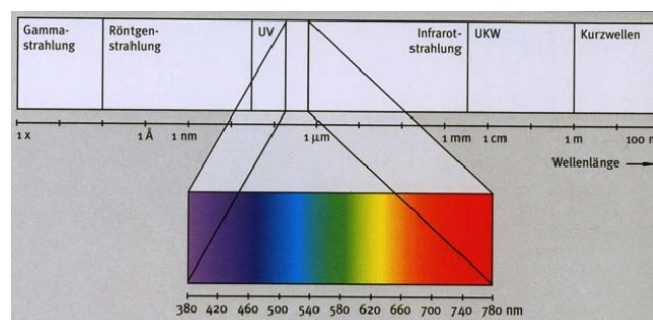


Abbildung 4: Beispiel aus Smith (2002). Die Darstellung zeigt den Bereich des sichtbaren Lichtes auf der Skala der elektromagnetischen Wellen. Wir Menschen unterscheiden zwischen blau, grün, gelb, rot u.s.w. Sie sind, wenn man so will, digitale Einheiten. Die Welt da Draussen ist jedoch ein Kontinuum.

Es ist also ein Bestreben des Menschen, in jede Vielfalt eine Ordnung hineinzubringen. Das Verfahrensprinzip für das Errichten von Ordnungssystemen ist die systematische *Klassifikati-*

on der Sinneseindrücke¹². Informationsklassen oder –cluster sind dabei das Instrument praktischer Anwendung und unterliegen unter anderem dem Faktor der Ökonomie. Unsere Kategorien funktionieren meist auch dann, wenn die Wirklichkeit voll von Grenzfällen und kontinuierlichen Übergängen ist. Das muss nicht heissen, dass wir uns der fließenden Übergänge nicht trotz allem bewusst sind. In verschiedenen Experimenten die im Rahmen der Prototypentheorie durchgeführt wurden, liess sich z.B. zeigen, dass der Mensch durchaus zwischen einem Kern von prototypischen Fällen und einem Randbereich von Grenzfällen unterscheidet (Rosch, 1973, S.130-134; Rosch, 1975a). Zum Beispiel sind für viele Sprecher einer Sprachgruppe bestimmte Elemente einer Kategorie privilegiert als andere, oder wie Rosch (1973, S.111) es ausdrückt: „Some colors to which English speakers apply the word ‚red‘ are ‚redder‘ than others.“

2.1.2. Ähnlichkeit als Grundprinzip für Klassifikation

Es stellt sich die Frage, nach welchen Informationsmustern der Mensch bestimmt, zu welcher Kategorie ein Element gehört. Nach Rosch und Mervis (1975b) geschieht dies nach einem aus der Spätphilosophie Ludwig Wittgenstein's entlehnten Prinzips der Familienähnlichkeit¹³. Wittgenstein hält fest, dass sich beispielsweise für die Kategorie „Spiel“ keine Merkmale angeben lassen, die für jede Art von Spiel Gültigkeit besitzen. Also schliesst er, die Kategorie „Spiel“ setze sich aus einer Vielzahl teilweise inhomogener Elemente zusammen, von denen zwar viele einige Eigenschaften teilen, bei weitem aber nicht alle. So ist theoretisch der Extremfall denkbar, dass Element 1 die Eigenschaft A und B besitzt, Element 2 die Eigenschaft B und C und Element 3 die Eigenschaften C und D. Die Elemente 1 und 3 befinden sich nun in der gleichen Kategorie, und zwar nicht weil sie dieselben Eigenschaften aufweisen, sondern weil jedem von ihnen eine Eigenschaft zukommt, die sie mit dem gleichfalls zur Kategorie gehörigen Element 2 teilen. Dieses Prinzip der Familienähnlichkeit lässt sich wie folgt illustrieren:

The network of overlapping similarities constituting the concept [...] is compared with the various resemblances that hold between members of a family. These may be of very different kinds: resemblances in build, facial features, colour of eyes or hair, gait, temperament, manner of speaking, attitude, or manners. [...] Although we can make such respects of resemblances precise, it is not in virtue of their all having some set of common properties that we group together members of an extended family; no property is sufficient for membership in the group, not is any one necessary (Baker & Hacker, 1980).

2.1.3. Hierarchieannahme

Wissenssysteme, davon gehen wir aus, sind hierarchisch organisiert (vgl. dazu Kapitel 3, S.43). Die Ausbildung von Kategorien und die Verwendung von hierarchischen Begriffssystemen im Umgang mit Wissen wird schon seit Aristoteles diskutiert, der in verschiedenen ontologischen Kreisen als der erste Ontologe der Weltgeschichte gilt. Die Lehre von hierarchischen Klassifizierungen, die er in seiner Kategorienlehre (vgl. Rolfes, 1974) dargelegt hat, liefert nicht nur für die verschiedenen Taxonomisierungen der Welt von Organismen (seit Lin-

¹² Dabei kann eine Klassifikation nur selten als definitiv abgeschlossen gelten, denn jede Informationsaufnahme stellt uns auch immer wieder vor Probleme der Einordnung neuer Erkenntnisse.

¹³ *Philosophische Untersuchungen*, Wittgenstein (1989 [Werkausgabe])

né¹⁴ in der *Biologie*) immer noch „die“ Grundlage, sondern wird auch noch für heutige Katalogisierungs- und Datenbanksysteme verwendet.¹⁵

2.1.4. Nützlichkeit oder naturwissenschaftliche Plausibilität – zwei Aspekte

Nicht jede Systematik eignet sich gleichermassen dazu, Wissen effizient weiterzugeben. Merkmalsorientiertes Sachwissen lässt sich nach den verschiedensten Kriterien systematisieren. In der Biologie hat sich in den letzten Jahrzehnten die Phylogenetik¹⁶ als ordnendes Prinzip für die Systematik durchgesetzt. Die aus den rekonstruierten Phylogenien abgeleiteten Klassen und ihre Zusammenhänge nennt man *Phylogenetische Systematik*. In der Phylogenetischen Systematik ist die Verwandtschaftlichkeit der Naturgegenstände durch den Verlauf der Evolution bestimmt. Hier kann es demnach nur „ein“ richtiges System geben. Die *Phänetik*¹⁷ (auch numerische Taxonomie) ordnet die Lebewesen nach ihrer Ähnlichkeit. Die Methoden der Phänetik unterscheiden sich wenig von den Methoden, wie sie bereits von Darwin benutzt wurden. Die Verwandtschaftlichkeit wird durch übereinstimmende Merkmale bestimmt. Der Vergleich mit anderen Naturgegenständen ist die Voraussetzung für das Erkennen der Merkmale, welche für den Naturgegenstand charakteristisch (zu den Methoden der Phänetik vgl. Rieppel, 1988 oder Harvey & Pagel, 1991). Eine dritte, von der darwinistischen Evolutionstheorie inspirierte Schule, ist die *evolutionäre Klassifikation*. Im Gegensatz zur phylogenetischen Systematik erlaubt sie, dass die subjektive Meinung des Wissenschaftlers in das System einfließt und toleriert, dass verschiedene Systeme Gültigkeit haben.

Den meisten Bestimmungsbüchern für Pflanzen oder Tiere dient die stammesgeschichtliche (phylogenetische) Verwandtschaftlichkeit als Ordnungsprinzip. Dies, obwohl die stammesgeschichtliche Abstammung der Naturgegenstände (Tiere, Pflanzen) oft noch weitgehend unbekannt sind. So kommt es, dass realistisch gesehen, die heute von Taxonomen verwendeten Systeme oft nicht wesentlich anders als zu Beginn des letzten Jahrhunderts sind.

Für die menschliche Wirklichkeit ist indes nicht immer nützlich, was naturwissenschaftlich richtig erscheint. Die Verunsicherung beim bestimmen von Pilzen, um nur ein Beispiel zu nennen, wäre kleiner, würden die Pilze primär ihrer auffälligsten Merkmale entsprechend und nicht nach der „vermeintlich“ stammesgeschichtlichen Verwandtschaft aufgeführt. Giftige oder ungeniessbare Pilze werden oft mit essbaren Pilzen verwechselt, denn auffällige Merkmale wie Hutfarbe, Grösse, Fussform, u.s.w. sind selten familiendefinierende Merkmale. In Bestimmungsbüchern von Pilzen wird aus diesem Grund oft mit Querverweisen auf giftige Doppelgänger in anderen Familien hingewiesen (z.B. Laux, 2001). Die geltenden Unterscheidungskriterien biologischer Systematiken sind dem Laien selten zugänglich, denn sie erfordern bereits einen hohen Grad an Sachwissen bezüglich der verwendeten Begrifflichkeit. Wer erinnert sich nicht an den Botanikunterricht im Gymnasium, wo man gezwungen war, Begriffe wie einkeimblättrig, 1-geschlechtlich, Rhizom, dorsiventral oder unterständig erst verstehen zu lernen, bevor man dann eine Butterblume über einen Bestimmungsschlüssel (in der

¹⁴ **Linnésches System:** Erstes international akzeptiertes System der Ordnung von Naturgegenständen in der zweiten Hälfte des 18. Jahrhunderts. Linné führte auch die wissenschaftliche Bezeichnung der Naturgegenstände mit Gattungs- und Artnamen ein, z.B. *Anas platyrhynchos* (=plattenschnäblige Ente, die Stockente).

¹⁵ Die **aristotelischen Realisten** halten hierarchische Klassifikationen für unser Denken gerade deshalb für nützlich, weil sie die Welt so wie sie ist darstellen und deshalb auch der natürlichen Art wie wir denken entsprechen. Begriffliche Hierarchien widerspiegeln die Welt. Der Kantianer hingegen favorisiert Begriffsbäume als bloss begriffliche Strukturen die der Mensch selbst kreiert.

¹⁶ **Phylogenetik:** Den begriff der Phylogenie wurde von ERNST HAECKEL, einer der ersten Vertreter von Darwins Thesen geprägt. Er entwickelte in Deutschland als erster konkrete Rekonstruktionen der Stammesgeschichte von Tieren.

¹⁷ **Phän:** deutlich in Erscheinung tretendes Merkmal eines Lebewesens, das mit anderen zusammen den Phänotypus bildet.

Schweiz gebräuchlich ist von Dr. August Binz 1920 (Binz/Heitz, 1990) entwickeltes System) zu bestimmen in der Lage war.

Doch gibt es auch Versuche, laienfreundlichere Klassifikationen geltend zu machen. Ein sehr beliebtes Bestimmungsbuch für Wildwachsende Blütenpflanzen in Mitteleuropa ist beim Kosmos Verlag erhältlich. In „Was blüht denn da?“ (Aichele & Golte-Bechtle, 1997) werden die Pflanzen primär nach der Farbe ihrer Blüten klassifiziert. Diese Art der Informationspräsentation erweist sich als sehr nützlich. Die Blütenpflanze lässt sich so auch von einem Laien in erstaunlich kurzer Zeit bestimmen. Das auffällige Unterscheidungsmerkmal „Farbe“ bestimmt die oberste Hierarchieebene dieses Wissenssystems. Es war der Schwede Carl von Linné, der 1735 in seinem Werk „Systema naturae“, als erster eine umfassende Klassifikation der Lebewesen unternahm. Da Linné vom Gedanken evolutionärer Entwicklung noch nichts wusste, betrachtete er seine taxonomischen Einheiten als ganz und gar künstlich. Dabei richtete er sein Hauptaugenmerk rein auf die Praktikabilität von Klassifikationssystemen. Im Vordergrund stand die schnelle Bestimmung der Objekte.

2.2. Ziel- und Problemstellung

Ziel dieser Arbeit ist die Visualisierung von Klassifikationssystemen (Wissensmodellen, Systematiken) auf Expertenniveau für Objektbezogenes Merkmalswissen. Kategorisierungsleistungen lassen sich mit der Methode der NMDS (Nonmetrische Multidimensionale Skalierung) darstellen.¹⁸ Die Nonmetrische Multidimensionale Skalierung ist ein Verfahren aus der Familie der multivariaten Analysemethoden. Objektbasiertes Merkmalswissen kann über die NMDS in einem mehrdimensionalen Raum dargestellt werden, indem sehr ähnliche Elemente (bzw. Objekte) nahe beieinander und unähnliche weit voneinander weg positioniert sind (zur Verfahrenslogik vgl. Borg & Groenen, 1997).

Im Rahmen der psychologischen Skalierungsverfahren liegen eine Vielzahl von Vorschlägen zur Ähnlichkeitsklassifikation vor (vgl. Kapitel 1.2, S.3).

2.2.1. Probleme der „Direkten Paar-Ähnlichkeitseinschätzung“

Im Rahmen des Projektes MODAL (vgl. S.10) wurden mit dem Verfahren „Direkte Paar-Ähnlichkeitseinschätzung“ bereits die verschiedensten Wissensstrukturkarten erstellt (vgl. Kapitel 7, 8 und 9; Expertenkarten für die Ornithologie, die Geographie und für Psychopathologische Störungsbilder). Das Verfahren zeichnete sich dabei als valides und sogar auf individueller Ebene als weitgehend zuverlässiges Verfahren aus (vgl. vor allem Kapitel 7). Die erwähnten Modelle gehen jedoch kaum über die Modellgrösse von 20 Objekte hinaus, was den folgenden Grund hat:

Beim Verfahren der direkten Paarähnlichkeitseinschätzung in der klassischen Form müssen alle Paarvergleiche zwischen den interessierenden Elementen des Wissensgebietes erhoben werden. Die resultierende Anzahl ergibt sich über die Formel $n*(n-1)/2$. Bei 20 Objekten müssen bspw. 190, bei 40 Objekten bereits 780 Ähnlichkeiten beurteilt werden. Nur vereinzelt wird in der empirischen Forschung der Aufwand betrieben, über die Methode des direkten Paarvergleichs, grössere Modelle zu erheben, wie das beispielsweise Lee, Byatt & Rhodes (2000) mit ihren Probanden für die grosse Urteilsmenge von 1176 Ähnlichkeitseinschätzungen bei 49 Objekten zur Beurteilung von Gesichtern durchgeführt haben. Für experimentelle

¹⁸ Zum Einsatz kommt hier RobuScal, ein Verfahren der Nonmetrischen Multidimensionalen Skalierung (Läge, Daub, Bosia, Ryf & Jäger, in review)

Laborstudien mit einer klar formulierten Fragestellung mag dieser Aufwand gerechtfertigt sein, für die Entwicklung von Expertenmodellen in der Grössenordnung von 200 Objekten ist dies jedoch ein nicht durchführbares Unterfangen. 200 Objekte bedeuten 19900 Paarvergleiche (Arbeitsaufwand ca. 35h).

Streule (2005), hat den Versuch unternommen, das Problem der hohen Urteilsanzahl durch eine gezielte Reduktion der Urteile zu senken. Eine ausgesuchte Menge an Objekten bilden dabei die Grundstruktur. Weitere Objekte werden nun über einige wenige Urteile in diese Struktur eingepasst. Wie die Ergebnisse zeigen, ist eine Reduktion der Urteile gemessen an der Veränderung der Kartenstruktur durchaus möglich. Der Erhebungsaufwand wächst damit linear statt exponentiell, bleibt jedoch bei einer Objektmenge von 200 nach wie vor unverhältnismässig hoch. Das Hauptproblem stellt dabei die Auswahl der Objekte für die Grundstruktur dar. Sie ist ganz und gar subjektiv. Die Objekte werden dabei nach dem Kriterium „das Gesamtsystem muss in seiner Ganzheit repräsentiert sein“ ausgesucht. Dieser Punkt scheint mir problematisch. Denn die Gesamtheit aller Objekte bestimmt die Ordnung des Systems oder um es in den Worten der Gestaltpsychologen auszudrücken: „Das Ganze ist mehr als die Summe seiner Teile.“

2.2.2. Auf der Suche nach Alternativen

Viele kognitionspsychologisch interessante Fragestellungen lassen sich problemlos auf der Grundlage kleinerer Modelle untersuchen. Im Hinblick auf die Entwicklung eines tutoriellen Systems muss der Objektumfang jedoch eindeutig ausgeweitet werden. Wir müssen uns also der Frage stellen, ob sich nicht vielleicht andere Verfahren der Ähnlichkeitsklassifikation besser dazu eignen, um klassifikatorisches Expertenwissen *effizient* und dennoch *inhaltlich korrekt* im grossen Stil zu erheben.

Die oben angeführten Bedenken bezüglich mangelnder Effizienz und Validität der Ergebnisse für grössere Modelle, haben uns dazu bewogen zwei weitere, in der psychologischen Tradition ebenfalls gebräuchliche Skalierungsverfahren, zu untersuchen: 1. „*Paralleles Sortieren*“, 2. „*Ähnlichkeitsklassifikation mit Hilfe von Beschreibungsvariablen*“ (Kapitel 1.2).

2.3. Methode und Vorgehen

Exemplarisch und explorativ wollen wir über diese beiden Methoden klassifikatorisches Expertenwissen für das Wissensgebiet der Ornithologie erheben.

Weltweit kennt man ungefähr 10 000 Vogelarten. Ungefähr 300 Vogelarten sind in Mitteleuropa häufig oder können zumindest regelmässig beobachtet werden. Die Vogelwarte Sempach (Schweiz) stellt davon im Internet ca. 200 Arten vor (www.vogelwarte.ch, zuletzt besucht Februar 2005). Datenbank auch auf CD-Rom mit dem Namen „Bubo“ erhältlich). 200 Arten sollen denn auch die für ein Expertenmodell angestrebte Objektanzahl sein.

Zwei Expertenmodelle wollen wir erheben: 1. Ein Expertenmodell für den Merkmalsbereich „*Nahrung*“, 2. Expertenmodell für den Merkmalsbereich „*Habitat/Lebensraum*“. (Diese Merkmale haben sich für Experten als besonders wichtig erwiesen, vgl. Daub, 2001).

Um Missverständnisse zu vermeiden, muss hier hervorgehoben werden, dass diese Wissensmodelle, mit der biologischen Vorstellung der Systematisierung von Naturgegenständen eher wenig gemeinsam haben. Die „Künstlichkeit“ dieser beiden Modelle (vgl. Kapitel 2.1.4) muss klar hervorgehoben werden. Was wir hier zu visualisieren versuchen, sind die von Experten verwendeten Ähnlichkeitskriterien und –prinzipien in einzelnen Merkmalsbereichen, welche sich für die praktische Anwendung besonders eignen.

2.3.1. Paralleles Sortieren

Versuchspersonen. 19 Ornithologen nahmen an der Erhebung „Paralleles Sortieren“ teil: 11 Vogelexperten für die Erstellung des Modells „Nahrung“ und 8 für das Modell „Habitat“. Die Ornithologen waren zwischen 27 und 55 Jahre alt, haben alle eine Ausbildung zum Exkursionsleiter absolviert und führen aktiv Vogelexkursionen durch.

Erhebungsmethode. Das „Parallele Sortieren“ wurde am Computer durchgeführt. Jedem Ornithologen wurde je eine Liste von 221 Vogelarten plus eines der beiden Kategorisierungskriterien („Nahrung“ oder „Habitat“) vorgegeben. Als erstes mussten die Experten ungefähr angeben, wie viele Kategorien sie zu bilden gedenken. Entsprechend der Anzahl genannter Kategorien wurden die Vogelarten mit „drag and drop“ einander zugeordnet. Die Kategorienanzahl durfte jederzeit beliebig verändert werden. Bei Unsicherheiten bezüglich den Merkmalsausprägungen eines Vogels stand es dem Experten frei, ein Bestimmungsbuch zu Rate zu ziehen. Nachdem alle Vogelarten einer Kategorie zugeteilt waren, mussten die Ornithologen jede Kategorie mit einem Kategorienlabel versehen. Für das Merkmal „Nahrung“ (n=11 Experten) wurden von den Experten insgesamt: 5, 6, 8, 8, 9, 9, 13, 28, 28, 32 und 39 Kategorien und für das Merkmal „Habitat“ (n=8 Experten): 5, 6, 10, 12, 12, 12, 15 und 21 Kategorien bestimmt (vgl. Tabelle 2).

Zweidimensionale Wissensstrukturmodelle. Jede Sortieraufgabe liefert pro Experte eine individuelle one-way-one-mode 01-Matrix (1: zwei Objekte gehören derselben Kategorie an; 0: zwei Objekte wurden unterschiedlichen Kategorien zugeteilt). Die 11 Matrizen zu Merkmal „Nahrung“ und die 8 zu Merkmal „Habitat“ wurden jeweils separat gemittelt. Die beiden Matrizen wurden dann mittels NMDS-Algorithmus RobuScal (Läge, Daub, Bosia, Ryf & Jäger, in review) zweidimensional skaliert. Auf diese Weise erhalten wir für beide Merkmalsbereiche je eine Wissensstrukturkarte (vgl. Abbildung 6 und Abbildung 10).¹⁹

¹⁹ Die beiden Wissensstrukturen „Nahrung“ (Abbildung 6) und „Habitat“ (Abbildung 10) aus dem Parallelen Sortieren sind, was die Objektauswahl angeht, zwecks Vergleichbarkeit, den beiden Wissensstrukturkarten „Nahrung“ und „Habitat“ aus dem Vergleich von Merkmalstexten (Abbildung 7 und Abbildung 11) angepasst. Für den Merkmalsbereich „Nahrung“ wurden die 221 Vögel auf 187, für den Merkmalsbereich „Habitat“ auf 183 Arten reduziert.

Experte A	Experte B	Experte C	Experte D
Allesfresser	Carnivoren	Aasfresser	Allesfresser
Fisch oder Schalentiere	Fischfresser	Allesfresser	auf dem Boden suchende Insektenfresser
Fleischfresser	Früchte-Beeren	Fischfresser	Blattinsekten fressende Vögel
Insektenfresser	Herbivoren	Fleischfresser	Fischfresser
Samen-oder Pflanzenfresser	Insektenfresser	Grünnahrung	Fließwasserinsekten fressende Vögel
	Körnerfresser	Grünnahrung und Wassergetier	Fluginsekten jagende Vögel
	Omnivoren	Insekten-Beerenfresser	giftig behaarte Raupenfresser
	Weichtierfresser	Insekten-Fleischfresser	Greifvögel auch vogelfressend
		Insekten-Körnerfresser	Hautflüglerfressende Vögel
		reine Insektenfresser	hochspezialisierte Samenfresser
		Samen-Beeren	Hühner: Blätter, Beeren, Nadeln
		Samenfresser	Insektenfresser (kleine Insekten)
		Wassergetier	Insektenfresser (Beute aufspiessend)
			Kleinsäuger jagende nachtaktive Vögel
			Maulwurfsgrillenfressende Vögel
			Mäuse und Amphibien
			Murmeltierfressende Greifvögel
			Saatgutfresser
			Samenfressende Hühner, weniger spezialisiert als Rauhfusshühner
			Samen
			Sumpfvögel: Insektenfresser
			Vegetarische Nahrung wird im Kropf zu Kropfmilch verarbeitet
			Waldinsektenfressende Vögel
			Wasservögel: Muschelfresser
			Wasservögel: Vegetarisch
			Watvögel in Riedgebieten nach Insekten suchend
			Watvögel Nahrung im Schlick suchend

Tabelle 2: Die Tabelle zeigt, die von vier Experten gebildeten Kategorien zum Merkmal Nahrung bei 221 Vogelarten. Die Anzahl der gebildeten Kategorien reicht dabei von 5 bis 29 Kategorien.

2.3.2. Ähnlichkeitsklassifikation über Textvergleich

Vorgehen. Für jede Vogelart wurden je zwei Merkmalsverteilungen erstellt, einmal für den Merkmalsbereich „Nahrung“ und einmal für den Merkmalsbereich „Habitat“. 5 Sachbücher dienten als Quelle für die Extraktion der beschreibenden Terme: *Bubo* (Schweizerische Vogelwarte, CD-Rom, auch unter www.vogelwarte.ch), *Vogelatlas 2.0*. (United soft Media, CD-Rom), *Der Kosmos-Vogelatlas* (Delin & Svensson, 1989), *Kompendium der Vögel Mitteleuropas* (Bezzel, 1993a, 1993b), *Die Vögel Mitteleuropas* (Singer, 2000).

Die Ähnlichkeit zwischen je zwei Merkmalsverteilungen wurde mit dem Überlappungskoeffizienten nach Marx (1976) bestimmt (zur Verfahrenslogik vgl. Kapitel 1.2.2).

Trigramm-Kodierung. Die Merkmalslisten wurden vor der Ähnlichkeitsberechnung allerdings trigrammiert²⁰. Die Trigrammierung wurde deshalb eingeführt, weil sie die Möglichkeit bietet, die Bedeutungsähnlichkeit zwischen zwei Indextermen wie „Fisch“ und „kleine Fische“ zu erfassen. Ohne Einsatz dieses Verfahren (es handelt sich hierbei um eine Art „primitiven Stemming“) würden Nahrungsmerkmale wie „Fisch“ und „kleine Fische“ als unterschiedliche Merkmalsausprägungen aufgefasst, und zwei fischfressende Vogelarten würden in der Ähnlichkeitsberechnung nicht als zusammengehörend erkannt. Die hier gewählte Form nennen wir „3-space-Verfahren. Dabei werden die Merkmalslisten wie folgt nochmals zergliedert: Trigrammieren wir „kleine Fische“ mit dem „3-space-Verfahren“, so erhalten wir die folgenden 13 Therme: „_kl kle lei ein ine ne_ e_f _fi fis isc sch che he_“. Space steht

²⁰ Die Trigramm-Kodierung basiert auf einem einfachen Repräsentationsalgorithmus, der Verknüpfungen zwischen Informationen herstellen kann. Brünken (1998) dokumentiert, dass es sich dabei um ein kognitiv plausibles, auf beliebige textuelle Dokumente anwendbares Verfahren handelt (die Dokumentenähnlichkeiten berechnete Brünken mit dem Cosinuskoeffizienten). Die n-gramm-Kodierung wird seit langem in verschiedenen Gebieten der Informatik verwendet. Sie wird normalerweise zur Textkomprimierung, zur fehlertoleranten Worterkennung, zur Spracherkennung, zur String-Suche und zum Information-Retrieval eingesetzt. Ein Überblick über die verschiedenen Verfahren finden sich bei Cohen (1995) und Adams (1992).

also für „Lehrschlag“. Wir vergleichen die Merkmalstexte unserer Vögel über Trigramm-Listen (vgl. auch Brünken, 1998, S.61).

Vorgesehen war für beide Merkmalskarten eine Grundgesamtheit von 200 Vogelarten. Doch waren die Vogelarten leider nicht in allen Quellen gleichermassen vertreten. Zum Teil fehlten einzelne Merkmalsbeschreibungen. Der Abgleich zwischen allen Quellen führte dazu, dass auf einige Vogelarten verzichtet werden musste. Die Objektanzahl für die Nahrungskarte betrug schlussendlich 187, für die Habitatskarte 183 Vogelarten.

Zweidimensionale Wissensstrukturmodelle. Mit dem NMDS-Algorithmus RobuScal wurden die Merkmalsähnlichkeiten für die beiden Merkmalsbereiche „Nahrung“ und „Habitat/Lebensraum“ zweidimensional skaliert.

2.3.3. Kodierung der Merkmale

Um die Wissensstrukturen (2x„Nahrung“; 2x„Habitat/Lebensraum“) inhaltlich beurteilen zu können, müssen die Merkmalsausprägungen in den Karten sichtbar gemacht werden.

In der Datenbank der Vogelwarte Sempach (www.vogelwarte.ch; bzw. CD-Rom „Bubo“), sind die einzelnen Vogelarten bezüglich einzelner Merkmalsbereiche mit ein, zwei bis höchstens drei Merkmalsausprägungen beschrieben (Abbildung 5). Da mehrere Vogelarten dieselben Merkmale bzw. Merkmalskombinationen aufweisen, klassifiziert diese Beschreibung sehr hoch und so fassten wir auch jedes in der Datenbank vorkommende Beschreibungsmerkmal (bei Nahrung sind dies: Insekten, Spinnen, Allesfresser, Pflanzen, Beeren, Würmer, Früchte, Aas, Knochen, Vögel, Sämereien, Muscheln, Fische, Knospen, Frösche, Mäuse, Kleinsäuger, Abfall; bei Habitat: Seen, Flüsse, Schilf, Gebirge, Felsen, Wald, Obstgärten, Siedlungen, Wiesen, Äcker, Kulturland, Buschwald, Hecken, Waldrand, Laubwald) und jede vorkommende Merkmalskombination als eine Merkmalskategorie auf und codieren sie mit einer

Zahl. Das Birkhuhn zum Beispiel frisst Pflanzen und Beeren. Somit gehört das Birkhuhn zur Merkmalskategorie „Pflanzen/Beeren“ an. Die Kategorie „Pflanzen/Beeren“ hat den Code 29.

Für die Beschreibung der Nahrungskarte erhielten wir so 34, für die Beschreibung der Habitatskarte 47 Merkmalskategorien.

Da 187 bzw. 183 Vogelarten in einer zweidimensionalen Karte aus Platzmängeln nicht mit ihrem Artnamen gekennzeichnet werden können, beschränken wir uns in den folgenden Darstellungen darauf, die Vogelarten über ihre Merkmalskategorien zu vertreten. Für die Beurteilung der Karte wird der Artnamen denn auch als zweitrangig eingestuft, da wir uns primär für die Merkmalsdimensionen im Klassifikationssystem interessieren.

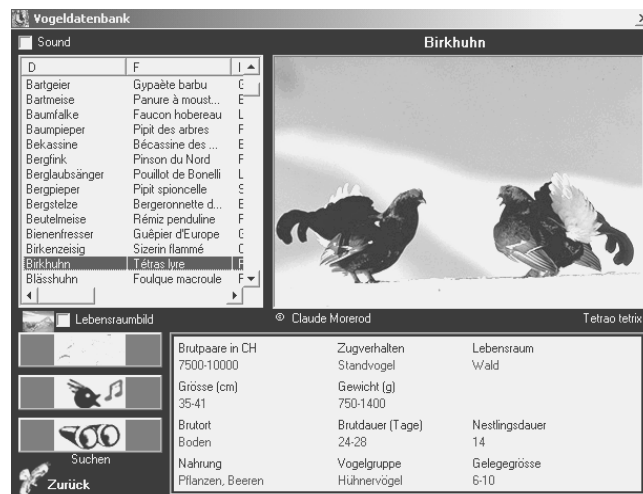


Abbildung 5: Merkmalsbeschreibung in „Bubo“. 200 Vogelarten der Schweizer Vogelwelt sind auf dieser CD-Rom erfasst. Dabei haben wir jede Merkmalskombination als eigene Merkmalskategorie definiert. Das Birkhuhn gehört somit zur Nahrungskategorie „Pflanzen/Beeren“ (Zahlencode 29) und zur Habitatskategorie „Wald“ (Zahlencode 10).

2.4. Resultate

2.4.1. Wissensstrukturkarten für den Merkmalsbereich „Nahrung“

Abbildung 6 zeigt die Wissensstrukturkarte für das Merkmal „Nahrung“ aus dem Parallelen Sortieren. Drei grosse nominale Klassen lassen sich aus der Karte herauslesen: 1. Fisch-Fleischfresser (oben), 2. Insektenfresser (rechts) und 3. die Pflanzenfresser (links). Am auffälligsten ist sicherlich die grosse Gruppe von Insektenfressern Mitte rechts. Tatsächlich sind mehr als die Hälfte aller Vogelarten in diesem Bereich der Karte platziert. Insekten sind z.B. Käfer, Läuse, Mücken, aber auch Insektenlarven wie z.B. Raupen. Im grossen Stil verspeist werden hier unter anderem auch Spinnen (Codierung 1 = Insekten/Spinnen). Hier finden sich vor allem kleinere Vogelarten wie Grasmücken, Pieper, Schwalbenarten, Merlen, Segler, Braunellen, Stelzen, Zaunkönige u.s.w., vorwiegend also Singvögel. Bewegen wir uns auf der Karte etwas von den Einsen weg und weiter nach unten in Richtung untere linke Ecke der Gesamtstruktur, so finden wir hier insektenfressende Vogelarten, welche immer mehr auch Bienen, zum Teil Würmer und im Winter aber auch Körner zu sich nehmen (Codierungen wie 3, 6, 12). Es sind dies z.B. Vogelarten wie Meisen, Spechte und Drosseln. Je weiter wir uns jetzt nach links bewegen, desto vegetarischer ernähren sich die Vögel (ab Codierung 31, 29). Typische Samenfresser (Codierung 31) sind z.B. viele Finkenarten, aber auch die Sperlinge sind hier anzutreffen und auch Tauben. Unter 29 finden wir viele Hühnervögel, typische Pflanzenfresser (Sprossen, Knospen, Sämereien, ...). Tatsächlich gibt es in der Vogelwelt wahrscheinlich eher wenige Arten, welche sich konsequent und ausschliesslich nur vegetarisch ernähren. Vor allem während der Jungenaufzucht füttern auch konsequentere Vegetarier ihrem Nachwuchs gerne Insekten. Und auch unter den Insektenfressern gibt es viele Arten, welche im Winter auf pflanzliche Nahrung umstellen müssen, sofern sie denn nicht gen Süden ziehen. Bewegen wir uns von den Vegetariern weg nach oben, so gelangen wir über pflanzenfressende Entenarten (unter Codierung 30) zu den Muschelfressern und schliesslich zu den Fischfressern. Typische Fischfresser sind z.B. Entenvögel wie der Gänsesäger. Auch diverse Reiherarten fressen Fisch, typisch auch der Eisvogel oder der Fischadler, und auch der Schwarzmilan mag gerne toten Fisch. Der Graureiher hingegen frisst nicht nur Fisch, sondern auch Frösche, und besonders gerne jagt er auf dem Feld Mäuse (Codierung 27) womit wir zu den konsequenteren Fleischfressern gelangen, wie etwa Greifvögel oder auch die Eulenartigen (Codierung 17 bis 23). Hinter der Codierung 15 und auch 16 verstecken sich die Schnepfenvögel, von denen viele Arten im seichten Wasser von Gewässern, z.B. auch an Land, nach Insektenlarven oder Würmern stochern.

Die grossen Nahrungskategorien, wie wir sie der Karte finden, sind natürlich eine grobe Vereinfachung der angetroffenen Wirklichkeit und beschreiben gerade deshalb sehr eindrücklich die geistige Fähigkeit des Menschen, grosse komplizierte Wissenswelten in einfache, kognitiv-ökonomisch sinnvolle Wissenssysteme zu verwandeln.

Die Organisation der Punkte in der Karte orientiert sich an einem 3-Simplex. Ein reguläres 3-Simplex beschreibt ein gleichseitiges Dreieck und ist in der NMDS dann zu erwarten, wenn die Ähnlichkeit der Objekte innerhalb einer Äquivalenzklasse maximal und die Ähnlichkeit von Objekten verschiedener Klassen minimal ist und nur diese beiden Ähnlichkeitswerte auftreten. Die Merkmalskarte „Nahrung“ aus dem Parallelen Sortieren entspricht in den Grundzügen dieser Organisation, ist jedoch nicht streng kategorial (die Übergänge zwischen den drei Eckpunkten sind eher fließend). Die Struktur wird aber als prägnant beurteilt. (vgl. *Identifikation von nominalen Klassen*, Gigerenzer, 1981, S.364).

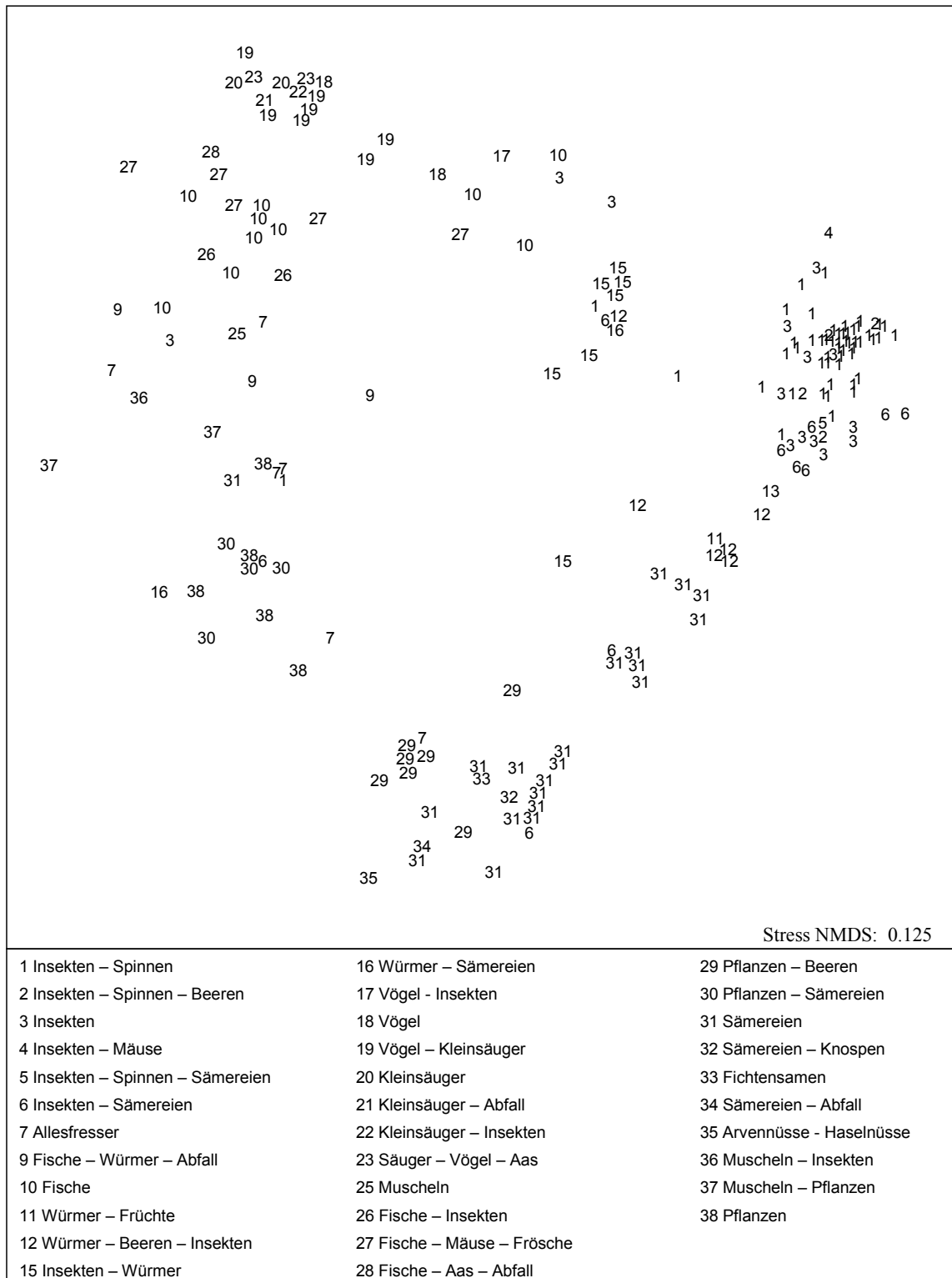


Abbildung 6: Wissensstrukturkarte „Nahrung“. Ähnlichkeitswerte aus dem Parallelen Sortieren. Die Vogelarten beschreiben drei grosse Nahrungsklassen: Fleisch-Fischfresser, Vegetarier und Insektenfresser wobei die Übergänge fließend sind.

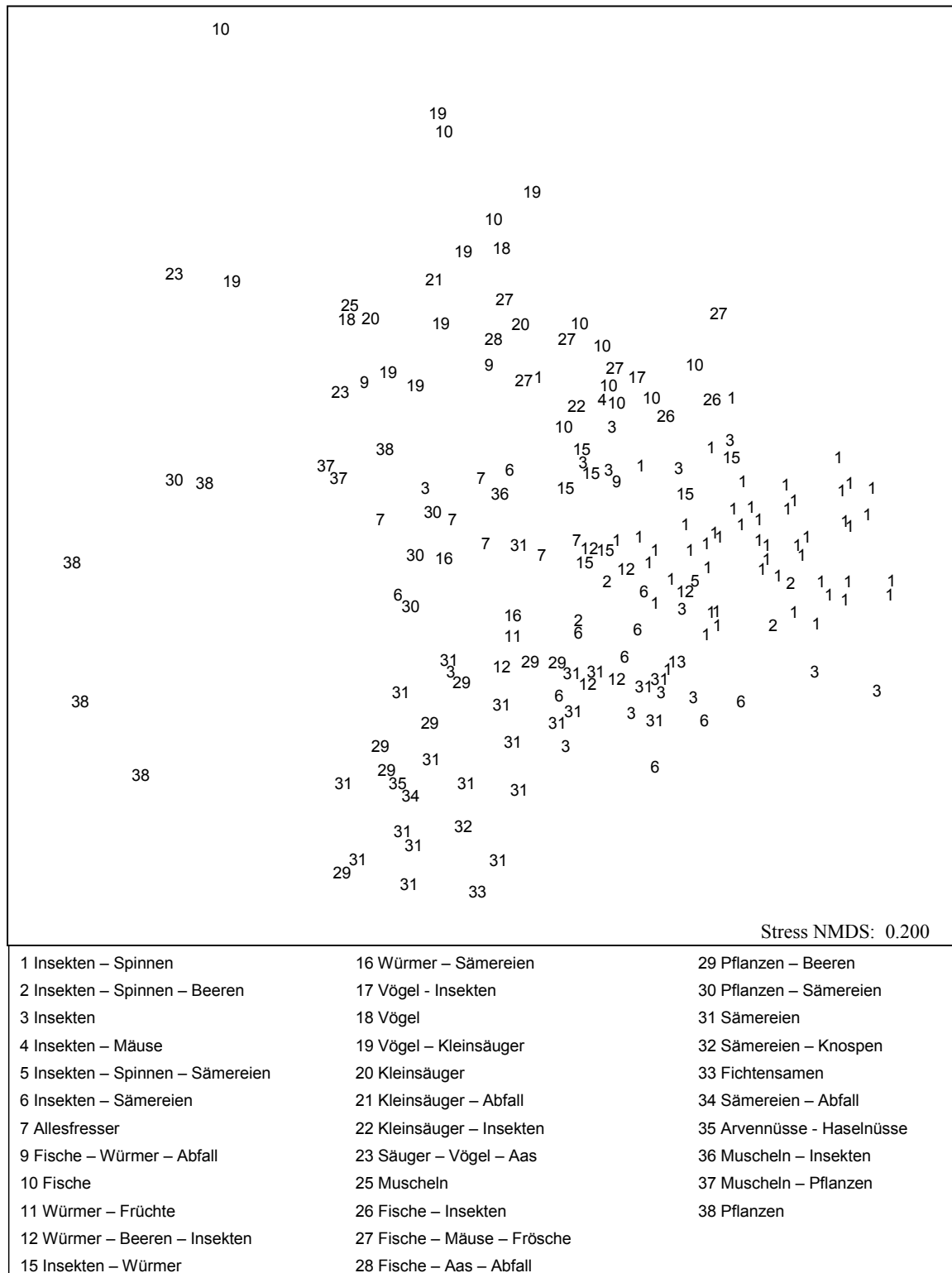


Abbildung 7: Wissensstrukturkarte „Nahrung“. Ähnlichkeiten berechnet über den Vergleich von Merkmalslisten (siehe Text). Die drei grossen Klassen Fleisch-Fischfresser, Insektenfresser und Vegetarier von Abbildung 6 überlappen hier räumlich sehr stark.

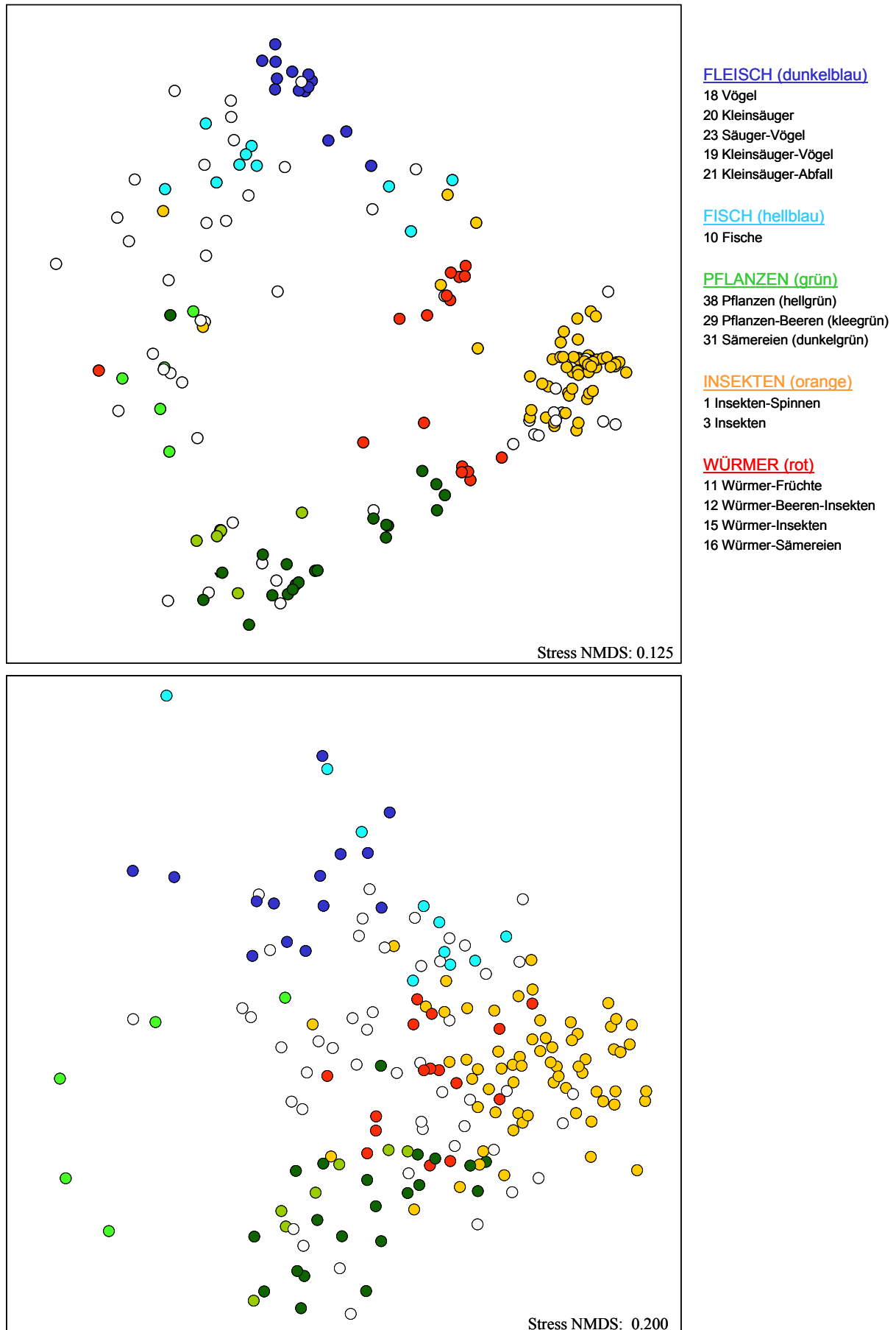


Abbildung 8a = Abbildung 6 in Farbe (Nahrungskarte „Paralleles Sortieren“)
 Abbildung 8b = Abbildung 7 in Farbe (Nahrungskarte „Textvergleich“)

Abbildung 7 zeigt die Wissensstrukturkarte „Nahrung“ aus den über Merkmalslisten berechneten Objektähnlichkeiten. Diese „künstlich“ berechnete Karte zeigt entgegen der eher klassifikatorischen Nahrungskarte keine deutlichen Objektcluster. Die Objekte streuen über den gesamten Raum. Zwei Merkmalsdimensionen scheinen hier in einem Winkel von ca. 55% aufeinander zu stehen (oben Fischfresser, unten Pflanzenfresser und rechts Mitte Insektenfresser). Die Übergänge zwischen den Extrempolen sind fließen.

Einige der grösseren „zahlencodierten“ Fresskategorien (grösser im Sinne von: mehrere Vogelarten zeigen dasselbe Fressverhalten) wurden für die Abbildungen 8a und 8b zu grösseren Klassen zusammengefasst und farblich hervorgehoben. Dabei kennzeichnet dunkelblau die Fleisch-, hellblau die Fischfresser (10), hell-, klee- und dunkelgrün die Pflanzenfresser (38, 29, 31), orange die Insektenfresser (1, 3) und rot die Würmerfresser (11, 12, 15, 16).

Wie ähnlich oder unähnlich die beiden Strukturen tatsächlich sind sehen wir, wenn wir sie mittels Prokrustes-Transformation²¹ übereinander legen (Abbildung 9). Ein AvgLoss von 0.41 ist ein mittelgrosser Ähnlichkeitswert. Er besagt, dass die beiden Strukturen einander in der Gesamtstruktur sehr ähnlich, in den Detailstrukturen eher unterschiedlich sind. Die Karte aus dem parallelen Sortieren ist viel stärker geclustert als die „künstlich“ berechnete Karte. Einige Parallelelemente in der künstlich berechneten Karte (links aussen) werden in der Darstellung deshalb nach aussen geschoben.

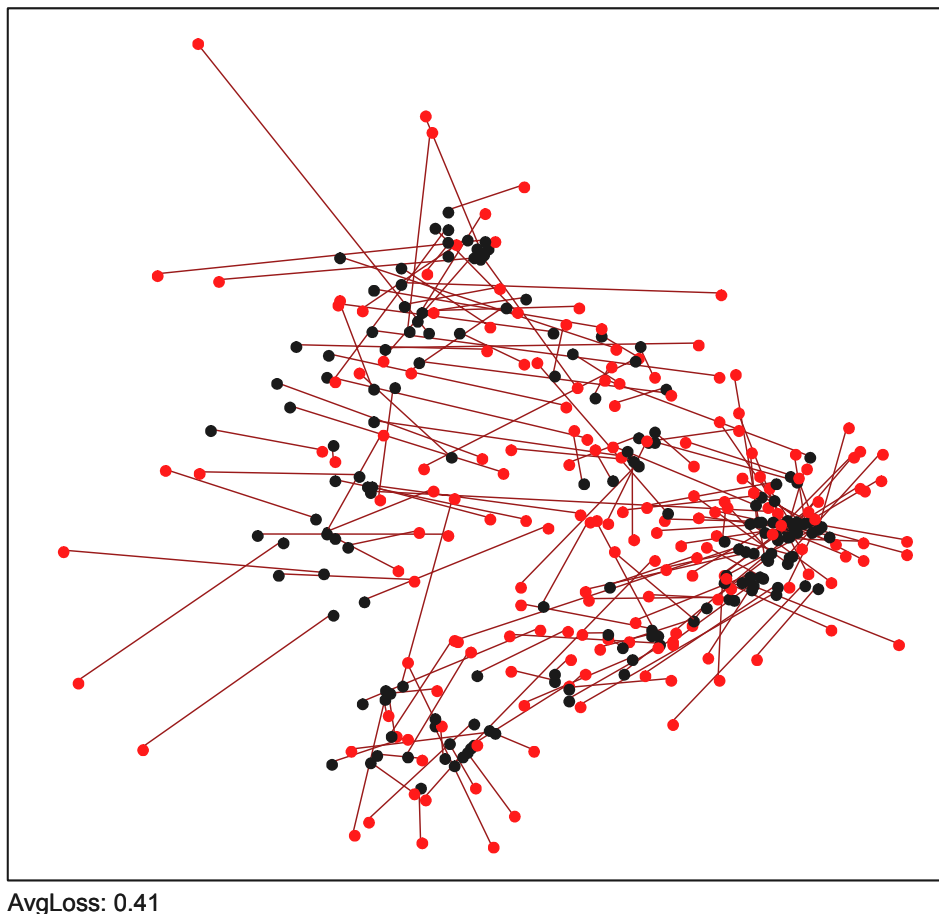


Abbildung 9: Prokrustes-Transformation der beiden Nahrungskarten (Paralleles Sortieren = schwarze Punkte)

²¹ Zur Verfahrenslogik der Prokrustes-Transformation ganz allgemein vgl. Gower & Dijksterhuis (2004). Das in dieser Arbeit verwendete Verfahren ist bei Läge (2001, S.139) beschrieben.

2.4.2. Wissensstrukturkarten für den Merkmalsbereich „Habitat/Lebensraum“

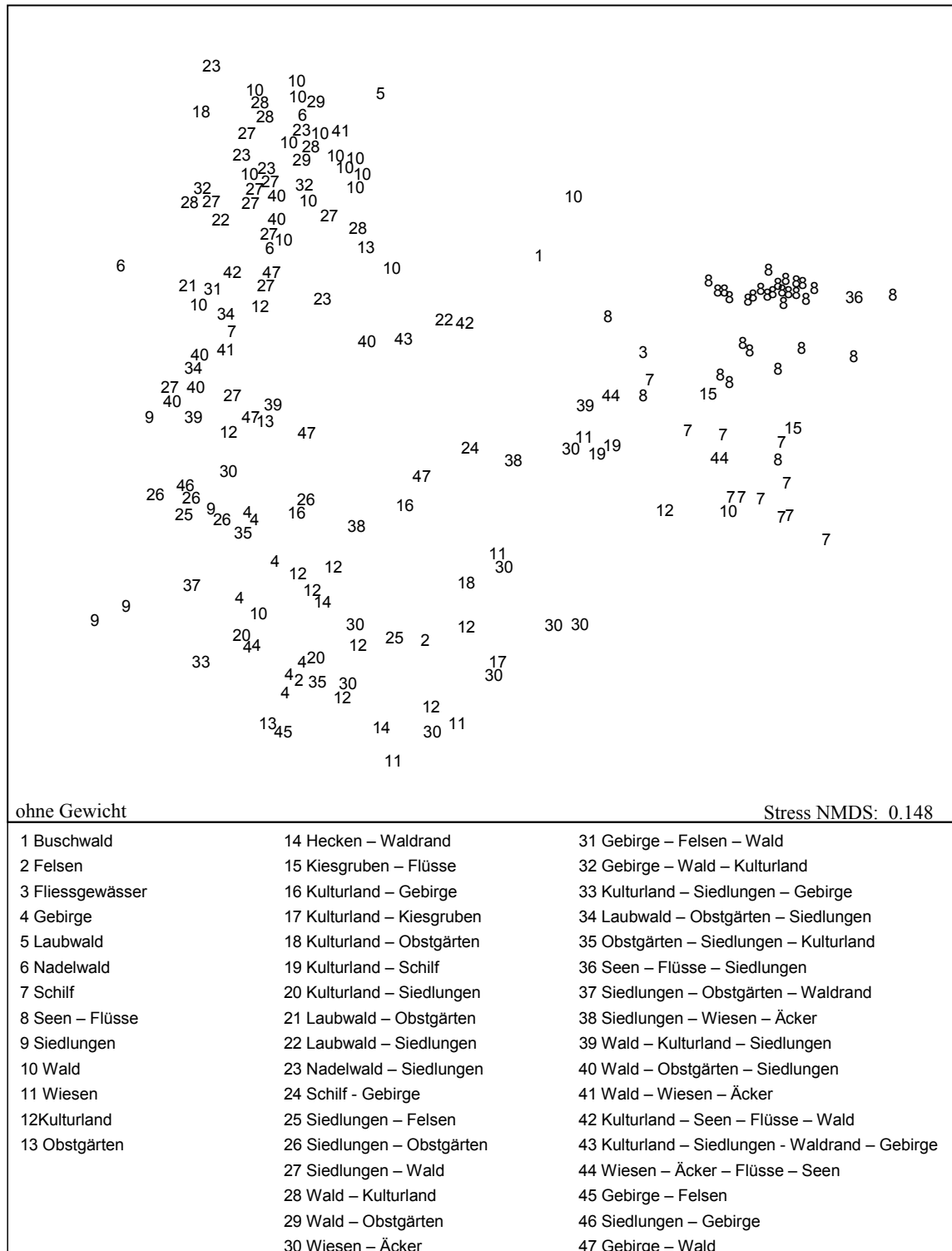


Abbildung 10: Wissensstrukturkarte „Habitat“ aus dem Parallelen Sortieren.

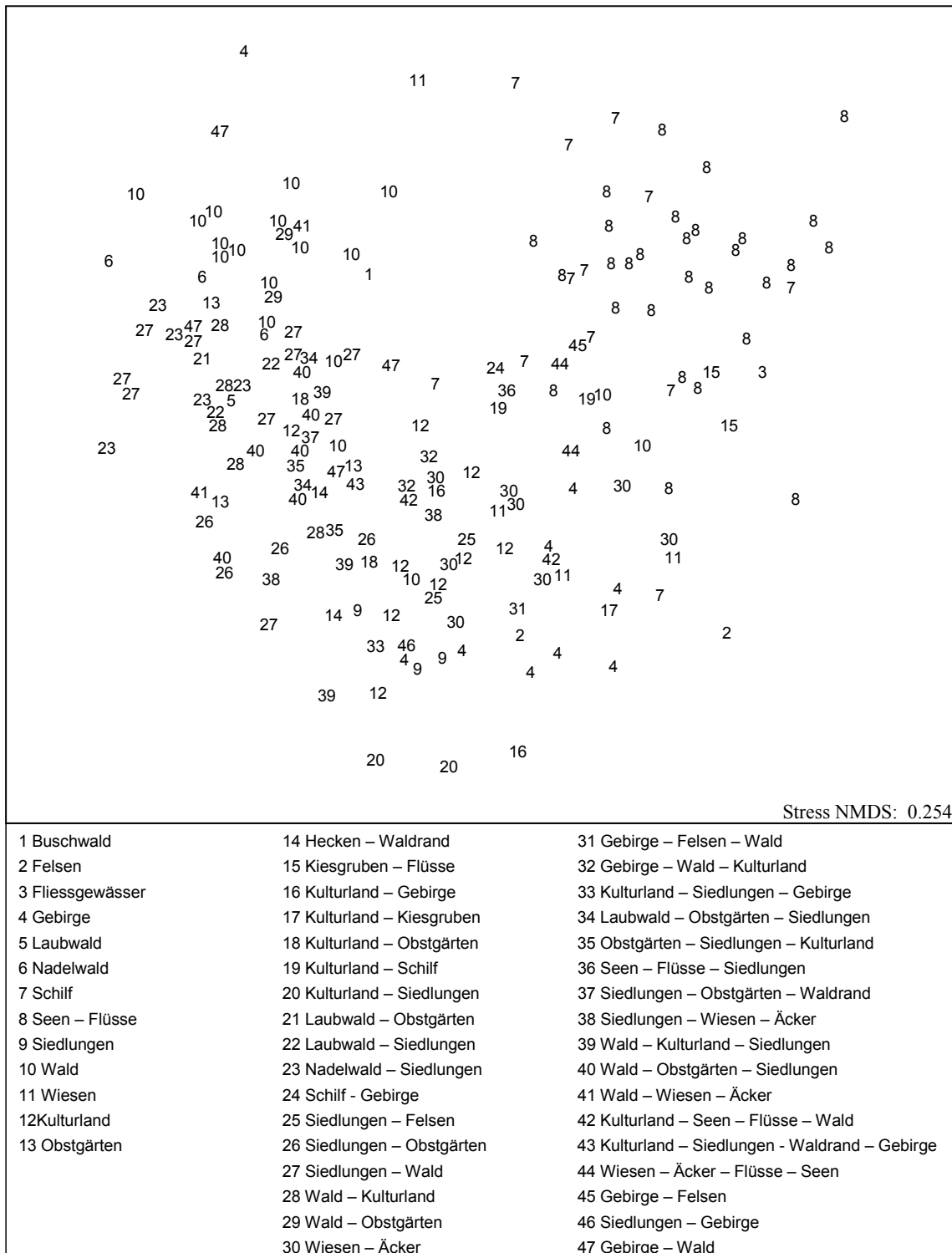
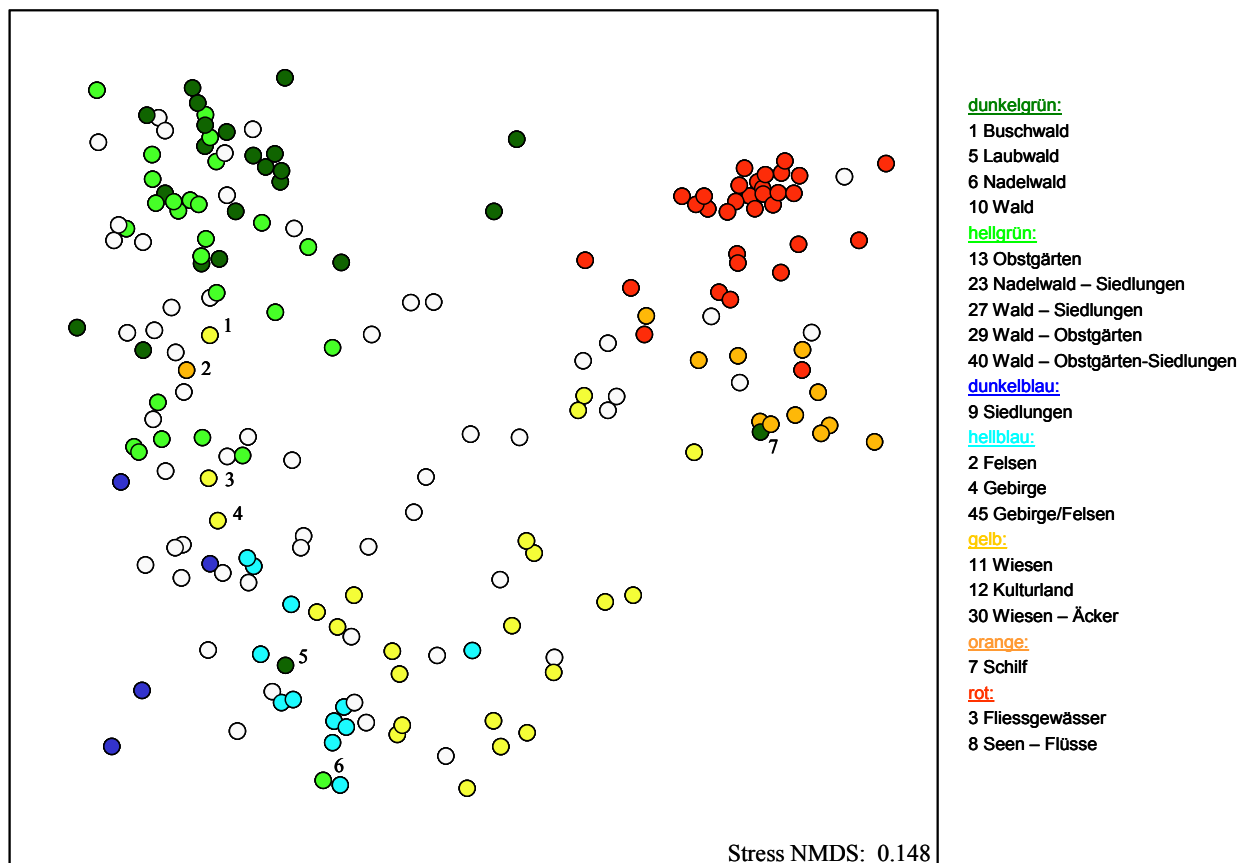


Abbildung 11: Wissensstrukturkarte „Habitat“ berechnet über Merkmalsbeschreibungen aus Büchern.



- 1 Rotmilan: Hügelige Gebiete. Kulturland mit Wald.
- 2 Karmingimpel: Auch feuchte Laubwälder mit dichtem Unterholz, Gärten, Kulturland und Weinberge.
- 3 Steinkauz: Kulturland mit alten Obstwiesen, Wäldchen, Hecken; Parks, Halbwüsten und Steinbrüche.
- 4 Gartengrasmücke: Gehölze und Wälder mit dichtem Unterholz, auch Parks und Gärten.
- 5 Birkhuhn: Vor allem im Gebirge, obere Waldgrenze. Waldübergangszone.
- 6 Rotkopfwürger: **Expertenfehler**. Gehört in die Kategorie Obstgärten, Parks, Siedlungen ...
- 7 Beutelmeise: Lebt nicht nur im Wald, sondern auch in Sumpfbereichen und Verlandungszonen.

Abbildung 12: Wissensstrukturkarte „Habitat“ aus dem Parallelen Sortieren. Einige Habitatskategorien sind hier zusammengefasst und farblich dargestellt.²²

²² Von den insgesamt 183 Vogelarten in Abbildung 12 scheinen 7 Vogelarten, was ihre Farbcodierung angeht, fehlplatziert. Eine genauere Analyse der Merkmalsausprägungen zeigt allerdings: Lediglich der Rotkopfwürger mit dem Kategorie 13 (Obstgärten) ist bei den Gebirgs- und Felsenbewohner in falscher Gesellschaft. Die anderen 6 Vogelarten sind, wenn wir uns an den etwas genaueren Merkmalsbeschreibungen in Singer (2000) orientieren, durchaus am richtigen Platz. Zum Beispiel ist das Habitat des Rotmilans (Kategorie 12, Kulturland). Beschreibung bei Singer: Hügelige Gebiete, Kulturland mit Wald. Für unsere Experten bestimmt der Ort wo der Rotmilan seine Jungen aufzieht oder wohin er sich nach der Jagt oder zum Schlafen zurückzieht (Wald) die Kategorienzugehörigkeit, währenddessen bei „Bubo“ das Jagdgebiet des Rotmilans den Ausschlag für die Kategorienzugehörigkeit gibt (hügeliges Kulturland). Ein anderes Beispiel ist der Karmingimpel. Habitat in „Bubo“: Schilf. Die Beschreibung bei Singer passt besser mit der Positionierung dieses Vogels in der Karte überein (Schilf, auch feuchte Laubwälder mit dichtem Unterholz, Gärten, Kulturland und Weinberge). Ähnlich bei Steinkauz, Gartengrasmücke, Birkhuhn und Beutelmeise. Das Habitat der Beutelmeise wird bei Bubo falsch beschrieben.

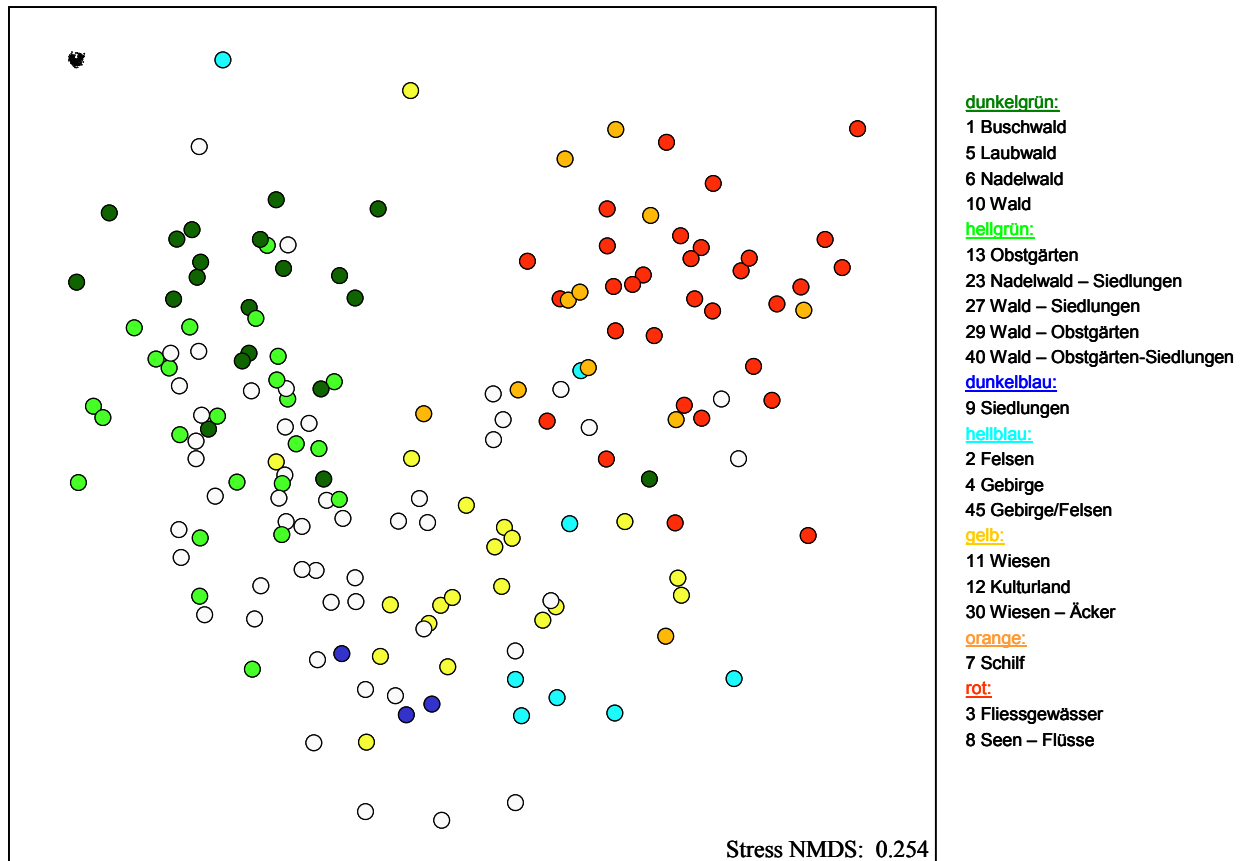


Abbildung 13: Wissensstrukturkarte „Habitat“, Ähnlichkeiten wurden über Merkmalsbeschreibungen in Büchern (Textähnlichkeiten) berechnet. Verschiedene Habitatskategorien sind hier farblich dargestellt.

Beide Habitatskarten sind U-förmig organisiert. (vgl. Abbildung 10 und Abbildung 11). Ausserdem zeigen sie nur wenige bis gar keine erkennbaren Objektcluster. Bei der Karte aus dem Parallelen Sortieren finden wir im oberen Bereich der Karte sowohl im linken als auch im rechten Bereich der Struktur eine erhöhte Objektdichte. Es stehen hier die Wald- den Seebewohnern gegenüber.

Beide Habitatskarten zeigen die Organisation einer Simplex (vgl. *Identifikation von Simplex Organisationen*, Gigerenzer, 1981, S.360). Die Simplexstruktur gilt als eine einfache Form der kognitiven Organisation. Die Bedeutungsträger werden entlang einer einzigen Bedeutungskomponente unterschieden. Man kann hier annehmen, dass es sich bei den Wald- und Seebewohnern um zwei entgegengesetzte Pole einer Merkmalsdimension handelt. Zwischen diese beiden Habitatpole fügen sich, von links nach rechts gelesen, die Habitatskategorien wie „Pärke“, „Siedlungen“, „Felsen/Gebirge“, „Kulturland“ (gemeint ist offenes Kulturland wie Wiesen, Brachland, u.s.w.) und „Schilf“ ein.

Viele Vogelarten bewohnen gleichzeitig oder im Wechsel der Jahreszeiten mehrere Habitate (Beispiel: 16, 24, 38, 42 oder 47, Abbildung 10). Manche an Felswänden brütende Vogelarten fühlen sich auch an hohen Hauswänden, in und an hohen Türmen der Städte oder in Siedlungen sehr wohl (Mauersegler, Felsenschwalbe, Haustaube (ehemals Felsentaube), Turmfalke, ...). Eine klare Einteilung ist deshalb nicht immer ganz einfach und fiel wohl auch unseren Experten nicht leicht.

Um die Ähnlichkeit der beiden Karten sichtbar hervorheben zu können, wurden einzelne grössere „zahlencodierten“ Habitatskategorien (grösser im Sinne von, mehrere Vogelarten be-

siedeln denselben Lebensraum) für die Abbildung 12 und Abbildung 13 zu grösseren Klassen zusammengefasst und farblich hervorgehoben. Vogelarten des Waldes (dunkelgrün, 1, 5, 6, 10), Obstgarten-, Park- und Siedlungsbewohner (hellgrün, 13, 23, 27, 29, 40), Siedlungsbewohner (blau, 9), Gebirgs- und Felsenbewohner (hellblau, 2, 4, 45), Bewohnern von offenem Kulturland (gelb, 11, 12, 30), Schilf- (orange, 7) und Vogelarten auf Flüssen und Seen (rot, 3, 8).

Wenn wir die beiden Habitatskarten mittels der Prokrustes-Transformation übereinander legen, zeigt sich sehr schnell, dass die beiden Strukturen nicht ganz so identisch sind, wie die eingefärbten Abbildungen es vielleicht vermuten lassen. Deutlich bleibt lediglich die bananenförmige Verteilung der Vogelarten erhalten. Wirkliche Ausreisser sind jedoch eher selten (z.B. der Wachtelkönig²³).

Auch hier gilt: Die beiden Strukturen sind einander in der Gesamtstruktur sehr ähnlich, in den Detailstrukturen aber eher unähnlich. Die Karte aus dem parallelen Sortieren ist etwas stärker geclustert als die Karte aus dem Merkmalsvergleich.

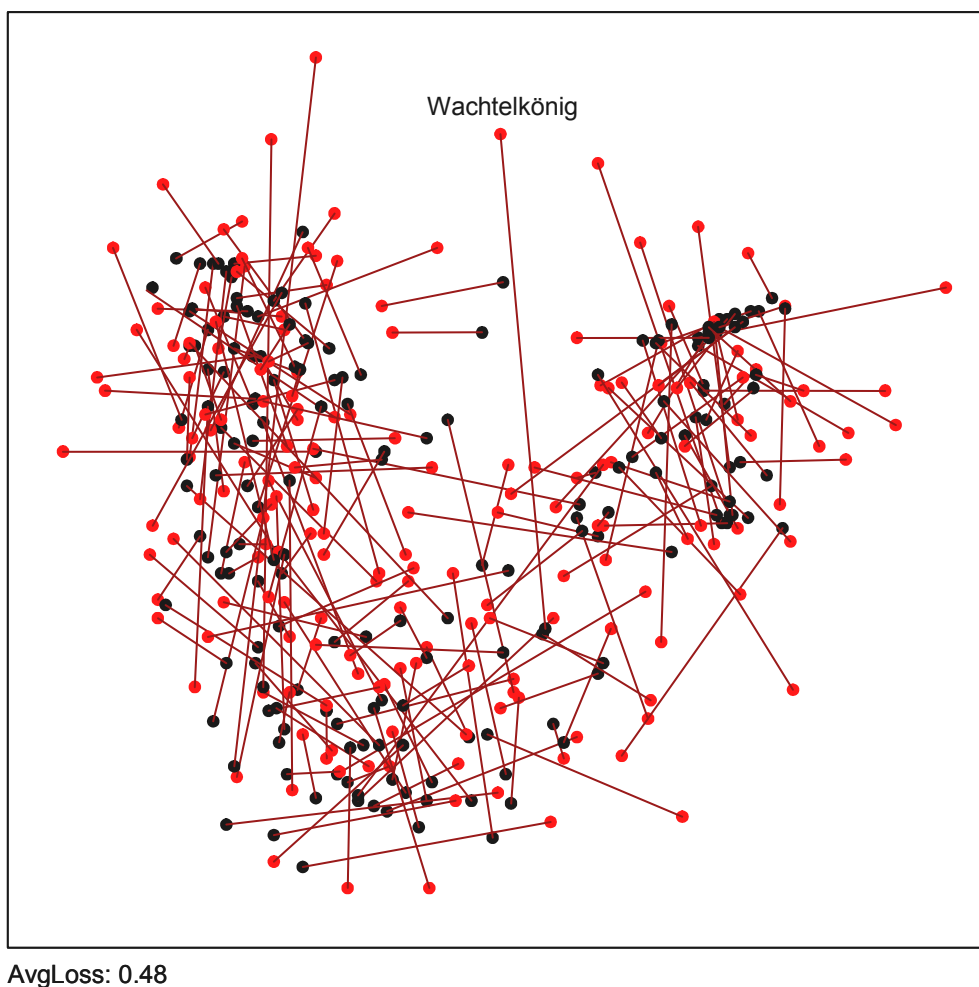


Abbildung 14: Prokrustes-Transformation der beiden Habitatskarten (Paralleles Sortieren = schwarze Punkte)

²³ Der Wachtelkönig findet in der künstlichen Karte wie wir sehen, seinen Platz zwischen See und Wald. Beim parallelen Sortieren eher in der Gegend Kulturland, was nach Ansicht der Autorin die sinnvollere Lösung ist.

2.4.3. „Paralleles Sortieren“ und „Ähnlichkeitsklassifikation mit Beschreibungsvariablen = Textvergleich“ oder „Direkte Paar-Ähnlichkeitseinschätzung“?

Im Rahmen des MODAL-Projektes wurden Paarähnlichkeiten bisher ausschliesslich mit dem Verfahren „Direkte Paar-Ähnlichkeitseinschätzung“ erhoben.

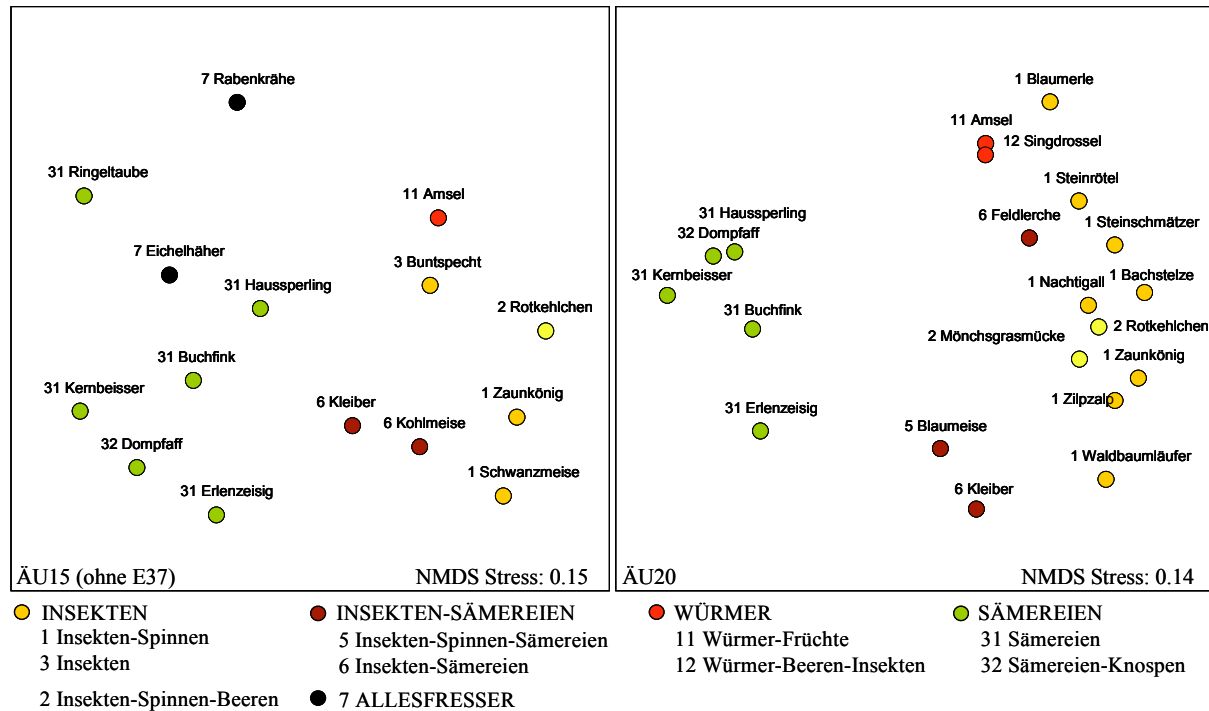


Abbildung 15: Expertenmodelle für den Merkmalsbereich „Nahrung“ mit der Objektzahl 15 bzw. 20 Vogelarten. Die beiden Karten wurden mittels paarweiser Ähnlichkeitseinschätzung erhoben.

Abbildung 15 zeigt zwei Nahrungsmodelle dieser Art. Je sieben Experten nahmen an der Erhebung der beiden Modelle teil.²⁴ Die Objektziehung in den beiden Modellen beträgt 15 bzw. 20 Vogelarten. Farblich hervorgehoben sind hier die folgenden Nahrungskategorien: Würmerfresser (rot), Insektenfresser (gelb/orange), Insekten-Samenfresser (braun), Samenfresser (grün) und Allesfresser (schwarz). Die 20iger-Karte (ÄU20) ist die kategorialere der beiden Strukturen. Links Aussen finden wir hier die Samenfresser (verschiedene Finkenarten und der Haussperling), im rechten äusseren Spektrum der Karte die Insekten-Spinnenfresser. Die Mischköstler (Feldlerche, Blaumeise und Kleiber) rücken von den reinen Insektenfressern in Richtung Samenfresser, ebenso die beiden Drosselartigen, Amsel und Singdrossel (in Rot).²⁵ Die beiden Nahrungspole, links die Vegetarier und rechts die Insektenfresser, finden sich auch in der 15er Karte (ÄU15) wieder. Allerdings zeigt diese NMDS keinerlei Prägnanz.

²⁴ Eine der individuellen Expertenmatrizen mit der Objektziehung 15 Vogelarten (E37) musste aus der Stichprobe entfernt werden, da sich die Karte dieses Experten von den individuellen Karten der anderen 6 Experten systematisch unterschied.

²⁵ Es ist bei der Klassifizierung nach bestimmten Merkmalsbereichen (hier Nahrung) wahrscheinlich nicht zu verhindern, dass immer auch „gelernte“ rein phylogenetische Annahmen in die Klassifikation mit einfließen. Wahrscheinlich auch dann, wenn die phylogenetische Klassifikation zum interessierenden Merkmalsbereich objektiv betrachtet wenig Parallelen aufweist (vgl. dazu Kapitel 2.5). Dies hier jedoch nur als Nebenbemerkung.

Die beiden Erhebungsverfahren („Paralleles Sortieren“ und „Textvergleich“) liefern, wie die folgende Darstellung (Abbildung 16) zeigt, durchaus vergleichbare Resultate. Die Proximitätsmatrizen der vier Nahrungs- und Habitatskarten (Abbildungen 6, 7, 10 und 11), wurden dabei auf eine Objektziehung von 15 und 20 Vogelarten reduziert und erneut zweidimensional skaliert. Die Abbildungen 17, 18, 19 und 20 zeigen die Prokrustes-Transformationen dieser vier Modelle mit den beiden Modellen direkter Paarvergleiche in Abbildung 15. Dabei steht das Kürzel ÄU für Ähnlichkeitsurteil, PS für Paralleles Sortieren und TV für Textvergleich.

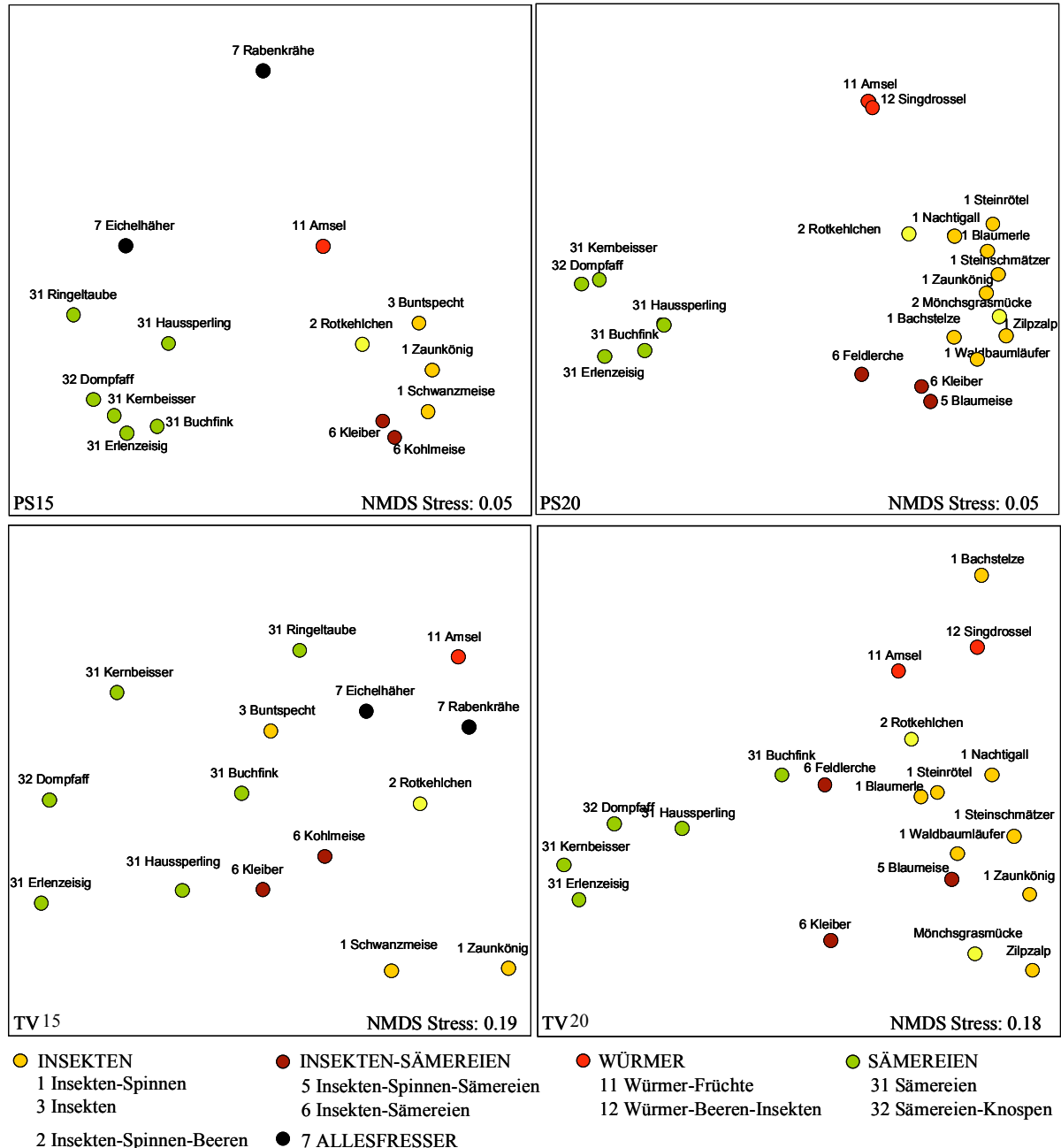


Abbildung 16: PS = Paralleles Sortieren, TV = Textvergleich (zu den Verfahren vgl. Kapitel 2.3.)

In PS15 (Abbildung 16 links oben) sind die Vogelarten räumlich in zwei Gruppen aufgeteilt, in Pflanzenfresser links und Insektenfresser rechts. Diese Zweiteilung der Karte findet sich bei TV15 (Abbildung 16 links unten) nur mehr in den Randbereichen bestätigt. Die Karte

zeigt im Mittelfeld grosses Rauschen (keine Objektkluster, sondern fließende Übergänge). Auch ist der Stresswert dieser Karte sehr viel höher (eine Replikation der Beobachtungen in der Betrachtung der beiden umfassenderen Wissensstrukturen Nahrung in Abbildung 6 und Abbildung 7).

Mit ÄU15 teilt PS15 eine ausserordentlich hohe strukturelle Ähnlichkeit. Ein AvgLoss von 2.9 steht nach empirischer Erfahrung schon beinahe für strukturelle Identität zweier Karten (Abbildung 17). Auch wenn der AvgLoss von 0.42 zwischen ÄU15 und TV15 nicht gleichermassen sensationell ist, so ist auch diese Übereinstimmung der Strukturen beachtlich. Er besagt, dass die beiden Strukturen zumindest global gesehen eine hohe Übereinstimmung zeigen, wenn die Anordnungen der Objekte im Detail auch etwas auseinandergehen.

Von allen drei 20iger Karten zeigt PS20 die kategorialste Organisation seiner Objekte (Abbildung 16 rechts oben). Wir finden hier keine Zwei-, sondern eine Dreiteilung der Objekte. Links aussen die „Pflanzenfresser“, rechts aussen die „Insektenfresser“ und oben die „Würmerfresser“. In ÄU20 finden wir die Dreiteilung „Insektenfresser“, „Pflanzenfresser“ und „Mischköstler“ (Abbildung 15, rechts). Die Struktur von TV20 wiederum ist ganz und gar unklassifikatorisch. Die Übergänge in der horizontalen Dimension, zwischen links aussen „Pflanzenfressern“ und rechts aussen „Insektenfressern“, sind auch hier fließend. Auch ist der Stresswert der Karte TV20 gegenüber der Karte PS20 vergleichsweise hoch (Abbildung 16, rechts unten).

Die beiden Prokrustes transformationen ÄU20/PS20 (Abbildung 19) und ÄU20/TV20 (Abbildung 20) sind gewissermassen eine Replikation der Resultate von oben: grosse strukturelle Übereinstimmung zwischen der PS und der ÄU-Karte selbst in Detailstrukturen (AvgLoss 0.32) und ein AvgLoss von 0.40 zwischen der ÄU- und der künstlich berechneten Karte. Also auch hier liegt eine mittelgrosse Ähnlichkeit der beiden Karten vor welche besagt, dass die beiden Konfigurationen einander in der Gesamtstruktur sehr ähnlich, in den Detailstrukturen jedoch verschieden sind.

Fazit. Die beiden empirischen Verfahren, „Direkte Paar-Ähnlichkeitseinschätzung“ und „Paralleles Sortieren“ liefern sehr ähnliche Ergebnisse. Beide sind sie äusserst klassifikatorisch, was beim Parallelen Sortieren wenig erstaunt, denn schliesslich wird die Kategorisierung von den Experten explizit verlangt. Den künstlich errechneten Ähnlichkeitsdaten fehlt es hingegen deutlich an struktureller Prägnanz. Die Karten sind kaum klassifikatorisch, die Übergänge zwischen den Merkmalspolen fließend (grosse Streuung). An den Eckpunkten zeigen auch die künstlich errechneten Karten mit den beiden anderen Verfahrenslösungen hohe strukturelle Übereinstimmung. Die Karten sind einander, was die Grundkriterien betrifft, sehr ähnlich.

Vorschläge für die Vereinfachung von schwach klassifikatorischen Wissensstrukturen finden sich in Kapitel 4!

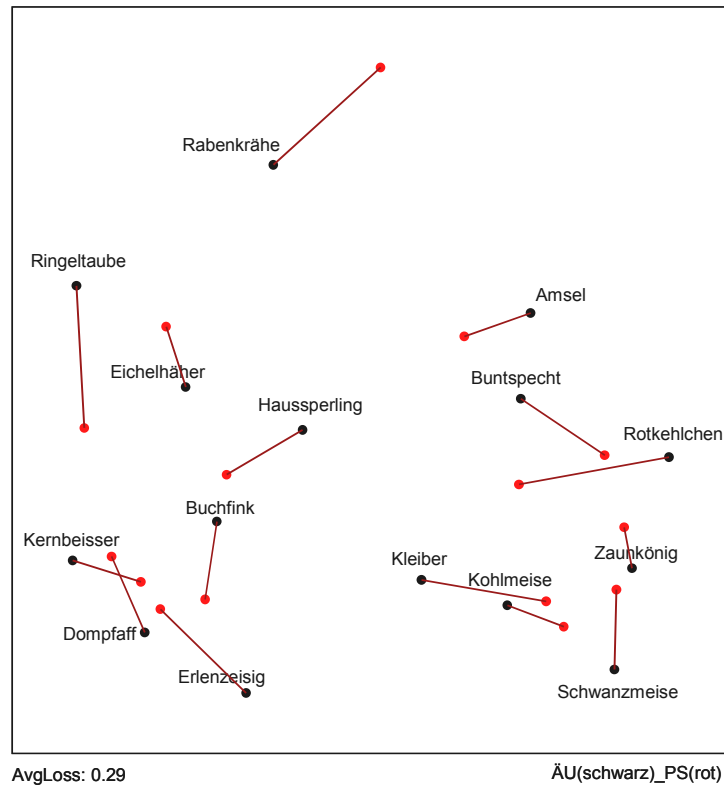


Abbildung 17: Prokrustestransformation ÄU15 (schwarze Punkte mit Label) und PS15

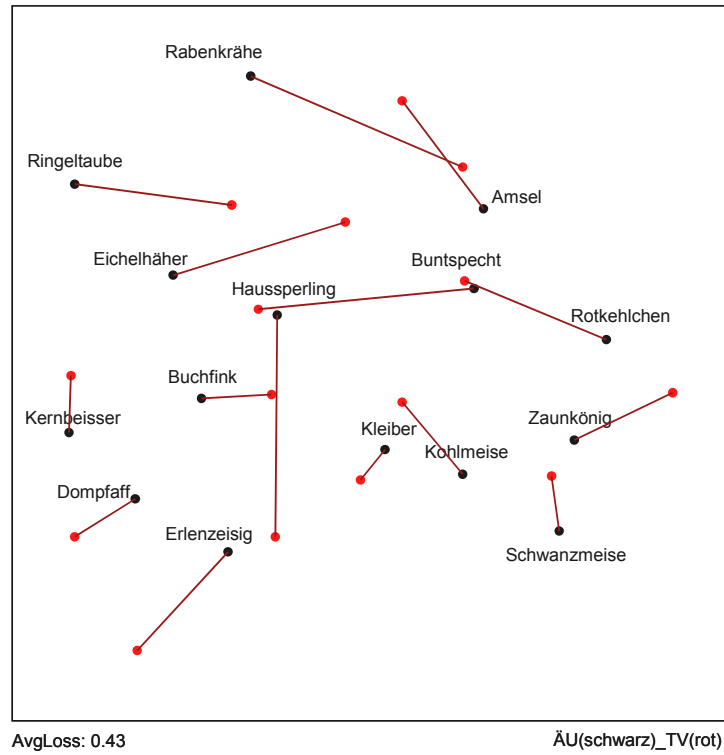


Abbildung 18: Prokrustestransformation ÄU15 (schwarze Punkte mit Label) mit TV15

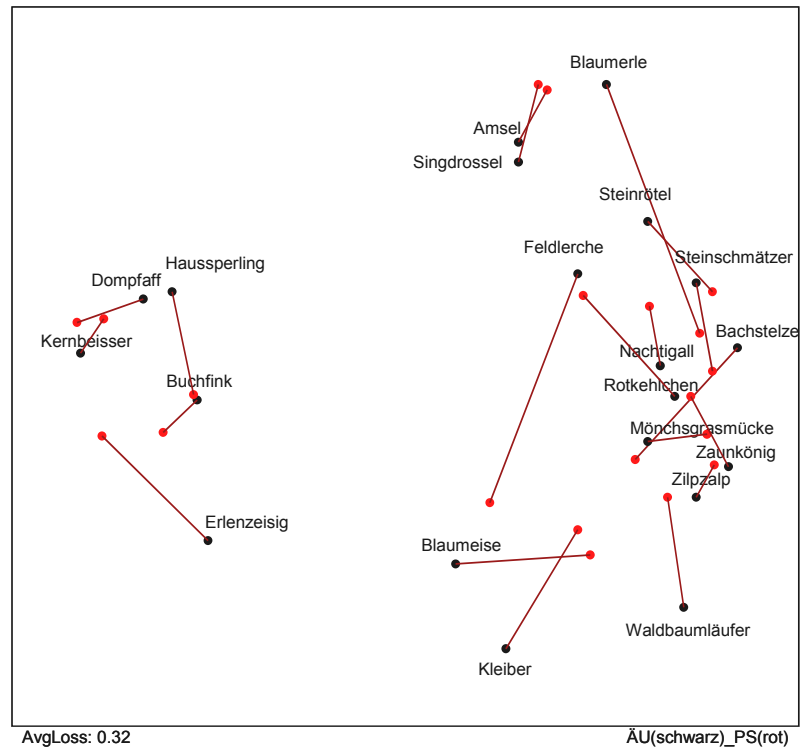


Abbildung 19: Prokrustestransformation ÄU_{20} (schwarze Punkte mit Label) mit PS20

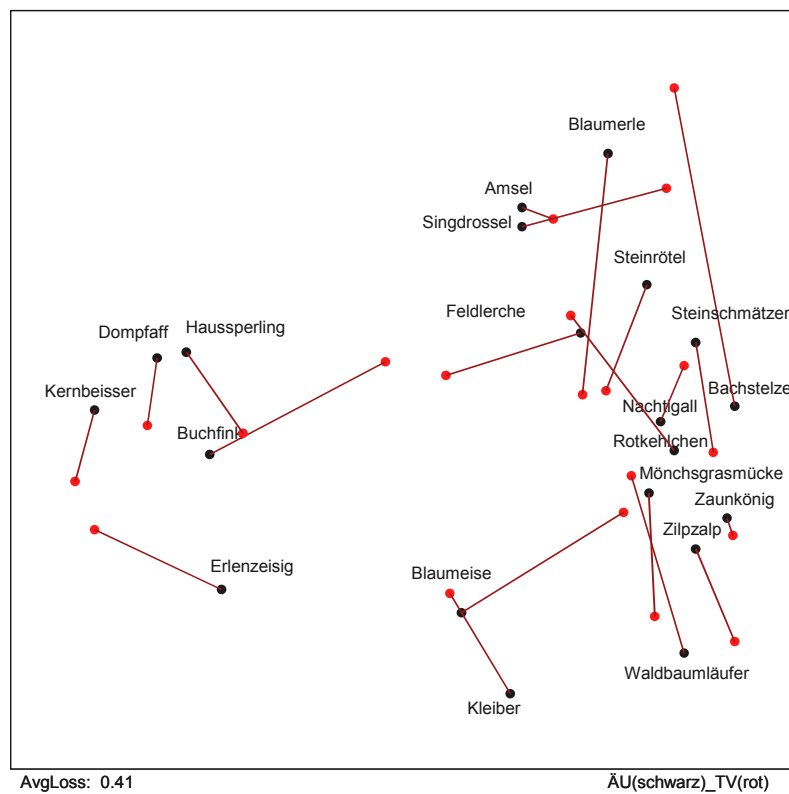


Abbildung 20: Prokrustestransformation ÄU_{20} (schwarze Punkte mit Label) mit TV20

2.5. Exkurs

In einer Nebenbemerkung (Fussnote 25) haben wir die Vermutung geäußert, dass bei der Klassifizierung der Vogelarten bezüglich „Nahrung“ oder „Habitat“ wahrscheinlich immer auch „gelernte“ phylogenetische Aspekte in die Ähnlichkeitsklassifikation mit einfließen und zwar auch dann, wenn der gewählte Merkmalsbereich objektiv betrachtet (ein Extrembeispiel wäre zum Beispiel die Einteilung von Vogelarten nach Farbe oder Grösse), mit der nach rein phylogenetischen Kriterien erstellten Klassifikation keine Übereinstimmung zeigt.

Der Einfluss dieser phylogenetischen Kategorisierung im Bereich, der von uns systematisierten Merkmale, ist jedoch wahrscheinlich weit weniger dramatisch als vermutet, wie das folgende Beispiel zeigt.

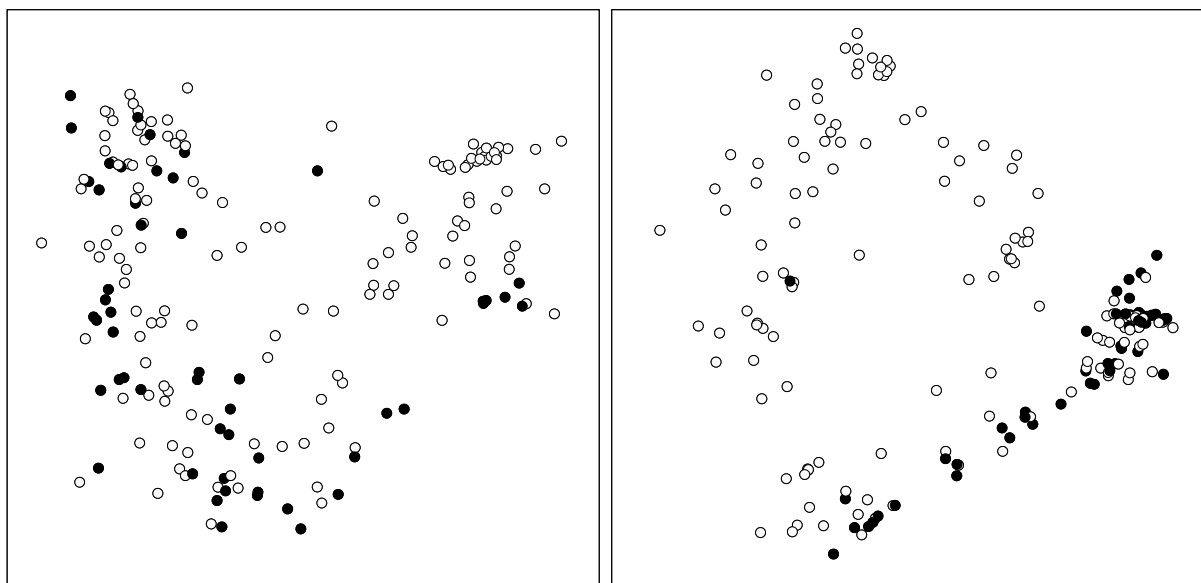


Abbildung 21: Expertenstrukturen „Habitat/Lebensraum“ mit 183 und „Nahrung“ mit 187 Vogelarten, erhoben mit dem Verfahren „Paralleles Sortieren“. Die schwarzen Punkte markieren 51 Singvögel.

Die in Abbildung 21 dargestellten Wissensstrukturen sind uns aus dem vorangehenden Kapitel bereits gut bekannt (vgl. Habitatskarte, Abbildung 6 und Nahrungskarte, Abbildung 10). Beide wurden sie mit dem empirischen Verfahren „Paralleles Sortieren“ erhoben. In der Ornithologie werden die Vogelarten in Singvögel und Nichtsingvögel unterteilt. Von den insgesamt 183 Vogelarten der Habitatskarte bzw. den 187 Vogelarten der Nahrungskarte, gehören gerade mal 81 Vogelarten zu den Singvögeln. Es sind dies Vogelfamilien wie (Lerchen, Schwalben, Stelzen, Seidenschwänze, Wasserramseln, Zaunkönige, Braunellen, Sänger, Grasmücken, Goldhähnchen, Rohrmeisen, Schwanzmeisen, Meisen, Kleiber, Mauerläufer, Baumläufer, Beutelmeisen, Pirole, Würger, Rabenvögel, Stare, Sperlinge, Finken, Ammern). Für ein Lernexperiment, beschrieben bei Streule (2005), wurden 51 dieser 81 Vogelarten ausgewählt. Die beiden Karten in Abbildung 21 entstanden im Rahmen dieses Projektes. Sie zeigen die Positionen dieser 51 Vogelarten in den beiden Karten mittels schwarzer Punkte an. Interessant sind die hohe Konvergenz des Nahrungs- und die sehr niedrige Konvergenz der Habitatskarte mit der stammesgeschichtlich vorgegebenen Zweiteilung der Vogelarten in Sing- und Nicht-Singvögel. Offensichtlich ist der Zusammenhang zwischen der phylogenetischen Verwandtschaftlichkeit und dem Nahrungsverhalten sehr hoch, bei Habitat hingegen sehr niedrig. Wo in der Habitatskarte die Singvögel über die gesamte Karte verstreut sind, liegen

sie in der Nahrungskarte nahe beieinander. Das Beispiel zeigt, dass unsere Experten beim „Parallelen Sortieren“ der Objekte durchaus in der Lage waren, die Merkmalsähnlichkeiten mehr oder weniger unabhängig von phylogenetischen Lehrsätzen einzuschätzen.

2.6. Diskussion

Da der Erhebungsaufwand bei der „Direkten Paar-Ähnlichkeitseinschätzung“ bei zunehmender Objektmenge exponentiell zunimmt, nehmen wir erfreut zur Kenntnis, dass sich sowohl über das Verfahren „Paralleles Sortieren“ als auch über den „Vergleich von Beschreibungsvariablen“, inhaltlich vertretbare Expertenmodelle erheben lassen. Die Möglichkeit, dokumentiertes Expertenwissen in Wissensstrukturkarten umzusetzen, bietet die Chance, Objektziehungen beliebiger Grössenordnung zu bearbeiten und nicht auf die Mitarbeit von Experten angewiesen zu sein.

Die klassifikatorischsten Strukturen erhalten wir über Paralleles Sortieren. Darüber sind wir allerdings wenig erstaunt. Schliesslich zwingen wir hier die Experten gerade dazu ihr Wissen klassifikatorisch zu organisieren. Die Wissensstrukturkarte zeigt für das Merkmal „Nahrung“ eine an die 3-Simplex angenäherte Organisation, die „Habitatskarte“ ein nach oben geöffnetes U-Simplex (vgl. Abbildung 22).

Im Vergleich dazu scheinen die Wissensstrukturkarten aus dem „Textvergleich“ wenig zu clustern. Zwar zeigt auch hier die Habitatskarte eine U-förmige Organisation, die entsprechende Nahrungskarte aus diesem Verfahren hingegen, zeigt eine hohe Objektstreuung über

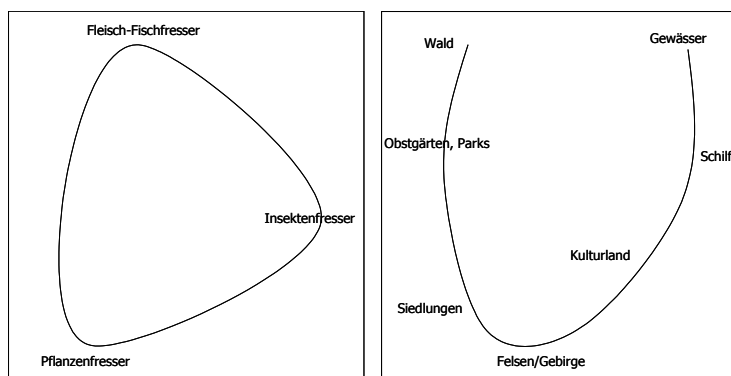


Abbildung 22: Schematische Darstellungen der beiden Expertenkarten „Nahrung“ (Abbildung 6) und „Habitat“ (Abbildung 10). Beide Karten wurden in Zusammenarbeit mit Experten und dem empirischen Verfahren „Paralleles Sortieren“ erhoben.

die ganze Karte. Offensichtlich fehlt es dem verwendeten Textmaterial zum Merkmalsbereich Nahrung an klassifikatorischer Prägnanz. Das mag darin begründet sein, dass Vögel eine hohe Generalisierungstendenz in ihrem Nahrungsverhalten zeigen: Auch die Vegetarier unter den Vögeln fressen ab und an auch schon mal ein Insekt. In den Habitatstexten hingegen sind die Habitatspräferenzen der einzelnen Vogelarten vermutlich eindeutiger. See-

vögel leben nicht im Wald und Bodenbrüter brüten niemals auf Bäumen. Die verwendete Begrifflichkeit in den Merkmalsbeschreibungen zum Merkmalsgebiet Habitat ist auf allen Ebenen der Objektbeschreibung klassifikatorisch, auch auf differenzierterem Niveau der Merkmalsbeschreibung.

In Kapitel 4 wird deshalb der Vorschlag gemacht, schwach klassifikatorische Systeme über den Umweg einer hierarchischen Clusteranalyse zu vereinfachen, denn die hierarchische Clusteranalyse transformiert die Unähnlichkeiten in Distanzen, welche nicht der Euklidik, sondern der ultrametrischen Ungleichung genüge leisten.

Methodenrevision. Es gibt jedoch auch andere Wege zur klassifikatorischen Karte. Jedes Buch, so haben wir festgestellt, repräsentiert das Wissen „eines“ Experten. Die Bücher waren

so ausgewählt worden, dass sie die Objekte auf unterschiedlichen Ebenen der Merkmalsdifferenzierung beschreiben. Wir vermuten nun, dass wir mit demselben Prozedere der Berechnung von Paarähnlichkeiten wie bisher, angewandt nun aber auf jede einzelne Textquelle, unterschiedlich klassifikatorische Systeme erhalten hätten. Bei einem Zusammenfügen dieser individuellen Ähnlichkeitsmatrizen hätten hochklassifikatorische Systeme einen viel höheren Einfluss auf die Organisation der Gesamtstruktur. Es wäre interessant zu überprüfen, ob nicht vielleicht auf diese Art und Weise eine Angleichung der Nahrungskarte aus dem „Textvergleich“, an die 3-Simplex der Nahrungskarte aus dem „Parallelen Sortieren“, möglich wäre.

Trotz aller Kritik zeigen die beiden Wissensstrukturkarten aus dem „Textvergleich“ mit den Strukturen aus dem „Parallelen Sortieren“ dennoch hohe strukturelle Übereinstimmung. In Detailstrukturen unterscheiden sie sich zwar etwas, doch was die Merkmalsdimensionen angeht, so liefern beide Verfahren ähnliche Ergebnisse.

2.7. Literatur

- Adams, E. S. (1992). *A study of Trigrams and their feasibility as index terms in a full text Information-Retrieval system*. Unveröff. Dissertation, The Faculty of Engineering and Applied Science, the George Washington University, Washington.
- Aichele, D. & Golte-Bechtle, M. (1997). *Was blüht denn da?*. Stuttgart: Franckh-Kosmos.
- Baker, G.P. & Hacker, P.M.S. (1980). *Wittgenstein. Understanding and Meaning. An Analytical Commentary on the Philosophical Investigations*. Oxford: Blackwell.
- Berlin, B. & Kay, P. (1991). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.
- Bezzel, E. (1993a). *Kompendium der Vögel Mitteleuropas. Passeres*. Wiesbaden: AULA Verlag.
- Bezzel, E. (1993b). *Kompendium der Vögel Mitteleuropas. Nonpasseriformes*. Wiesbaden: AULA Verlag.
- Binz, A. & Heitz, Ch. (1990). *Schul- und Exkursionsflora für die Schweiz*. (10. Auflage, überarbeitet von Heitz (Hrsg.)). Basel: Schwabe & Co.
- Borg, I. (1993). Scaling: A review of the German scaling literature of the last fifteen years. *The German Journal of Psychology*, 7, 63-79.
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling – Theory and applications*. New York: Springer.
- Brünken, R. (1998). *Automatische Rekonstruktion von Inhaltsbeziehungen zwischen Dokumenten. Benutzeradaptives Information-Retrieval in Wissensbasen*. Aachen: Shaker.
- Cohen, J. D. (1995). Highlights: language- and domain-independent automatic indexing terms for abstracting. *Information Science* 46, 162-174.
- Delin, H. & Svensson, L. (1989). *Der Kosmos-Vogelatlas*. Stuttgart: Kosmos.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhard.
- Gower, J.C. & Dijksterhuis, G.B. (2004). *Prokrustes Problems*. Oxford: Oxford University Press.
- Harvey, Ph. & Pagel, Md. (1991). The comparative method in evolutionary biology. *Oxford Series in ecology and evolution*. Oxford: Oxford University Press.
- Herkner, B. (1999). Über die evolutionäre Entstehung des tetrapoden Lokomotionsapparates der Landwirbeltiere. Ein konstruktionsmorphologisches Transformationsmodell auf evolutionstheoretischer Grundlage. *Carolinea, Beiheft 13*, 1-353.
- Janich, P. (1992). *Grenzen der Naturwissenschaft*. C.H. München: Beck.

- Janich, P. & Weingarten, M. (1999). *Wissenschaftstheorie der Biologie*. München: W. Fink Verlag.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophische Fakultät der Universität Zürich.
- Läge, D., Daub, S., Bosia, L., Ryf, S. & Jäger, C. (in review). *Robustness in nonmetric mulidimensional scaling*.
- Laux, H.E. (2001). *Der grosse Kosmos Pilzführer – Alle Speisepilze mit ihren giftigen Doppelgängern*. Stuttgart: Franckh-Kosmos Verlags-GmbH & Co.
- Lee, K., Byatt, G. & Rodes G. (2000). Caricature effects, distinctiveness, and indentification: testing the face-space framework. *Psychological Science*, 11, 379-385.
- Marx, W. (1976). Die statistische Sicherung des Überlappungs-Koeffizienten. *Zeitschrift für experimentelle und angewandte Psychologie*, 23, 267-270.
- Mayr, E. (1975). *Grundlagen der zoologischen Klassifikation*. Hamburg: Parey.
- Rieppel, O. (1988). *Fundamental of comparative biology*. Basel: Birkhäuser.
- Rolfes, E. (Hrsg.), (1974). *Aristoteles Kategorien / Aristoteles*. Hamburg: Meiner.
- Rosch, E. (1973). "On the Internal Structure of Perceptual and Semantic Categories". In T. E. Moore (Ed.). *Cognitive Development and the Acquisition of Language*. New York: Academic Press, 111-144.
- Rosch, E. (1975a). "Cognitive Representations of Semantic Categories". *Journal of Experimental Psychology: General*, 104 (3), 192-233.
- Rosch, E. & Mervis, C. (1975b). Family Resemblances: Studies in the International Structure of Categories. *Cognitive Psychology*, 8, 573-605.
- Schweizerische Vogelwarte. *Bubo* (3.Auflage, CD-Rom). Sempach. (Datenbank auch unter www.vogelwarte.ch, zuletzt besucht Februar 2005)
- Shepard, R.N., Romney, K.A. & Netherlove, S.B. (Eds.), (1972). *Multidimensional Scaling: Theory and applications in the Behavioural Science*. New York: Seminar Press.
- Simpson, G.G. (1967). *Principles of animal taxonomy*. New York: Columbia Univ. Press.
- Singer, D. (2000). *Die Vögel Mitteleuropas*. Stuttgart: Franckh-Kosmos.
- Sixtl, F. (1982). *Messmethoden der Psychologie* (2.Auflage). Weinheim: Beltz.
- Smith, B. (2002). Aristoteles, Kant und die Quantenphysik. In R. Hagengruber (Ed.), *Philosophie und Wissenschaft* (S.79-97). Würzburg: Königshausen und Neumann.
- Streule, R. (2006). *Sag mir, was ich noch nicht weiss. Individualisierung des Wissenserwerbs über adaptive Auswahl der Lerninhalte mit Hilfe der Nonmetrischen Multidimensionalen Skalierung*. Dissertationsschrift an der Philosophischen Fakultät der Universität Zürich.
- United soft Media, *Vogelatlas 2.0*. [CD-Rom]. Stuttgart: Kosmos.
- Wittgenstein, L. (1989). *Tractus logico-philosophicus*. Tagebücher 1914-1916. Tagebücher 1914-1916. *Philosophische Untersuchungen* [Werkausgabe]. Frankfurt: Suhrkamp.

3.

Die Rolle von Gewichtungsfaktoren bei Parallelen Sortierdaten im Hinblick auf spätere NMDS Lösungen

Zusammenfassung

Beim Parallelen Sortieren wird eine Person gebeten, eine Objektmenge auf frei zu wählende Weise in Teilgruppen zu zerlegen. Mit diesem in der Wissenspsychologie beliebten Verfahren erhält man Koinzidenzmatrizen als Grundlage zum Abbilden der Wissensstruktur einer befragten Stichprobe (z.B. mit Clusteranalyse oder NMDS). Die Varianz in einer solchen gemittelten Wissensstruktur kann entweder durch Heterogenität in den Sortierkriterien der einzelnen Befragten geprägt sein oder durch hierarchisch unterschiedlich hohe Schnittlinien eines Wissensbaumes (letzteres wird mit dem Konzept der Kognitiven Parallelität ausgedrückt). Ein einfacher Gewichtungsalgorithmus ermöglicht die Überprüfung dieser Alternativen. Er wird anhand ornithologischer Expertendaten durchexerziert und zeigt auf, dass im untersuchten Fall eine hohe Kognitive Parallelität besteht.

3.1. Paralleles Sortieren und die Hierarchieannahme

Ähnlichkeitsklassifikation in der Psychologie. Im Rahmen der psychologischen Skalierungsverfahren liegen eine Vielzahl von Vorschlägen zur Ähnlichkeitsklassifikation vor (z. B. Borg, 1983; Gigerenzer, 1981; Shepard, Romney & Nerlove, 1972; Sixtl, 1982; vgl. Kapitel 1). Das Spektrum reicht dabei von der unmittelbaren Paarweisen Ähnlichkeitseinschätzung über kategoriale Sortierverfahren und dimensionale Rangreihen bis hin zu indirekten Methoden wie dem Freien Assoziieren. All diesen Befragungstechniken ist gemeinsam, dass sie in einer Matrix Paarweiser Ähnlichkeitswerte zwischen einer zuvor ausgewählten Anzahl von Objekten münden. Diese Matrix wird dann durch Multidimensionale Skalierung in eine niedrigdimensionale, leicht zu interpretierende geometrische Struktur überführt, so dass die Distanzen zwischen den Objekten deren Ähnlichkeitsrelationen bestmöglichst abbilden.

Paralleles Sortieren und Koinzidenzmatrix. Beim Parallelen Sortieren (Abbildung 23) soll die Versuchsperson, einen Gesamtpool von Objekten in so viele Kategorien aufteilen, wie ihr nach Ähnlichkeitsgesichtspunkten sinnvoll erscheinen. Die Auswahl der Kriterien und die Entscheidung über die Grösse der gebildeten Kategorien obliegt dabei vollständig der befragten Person.

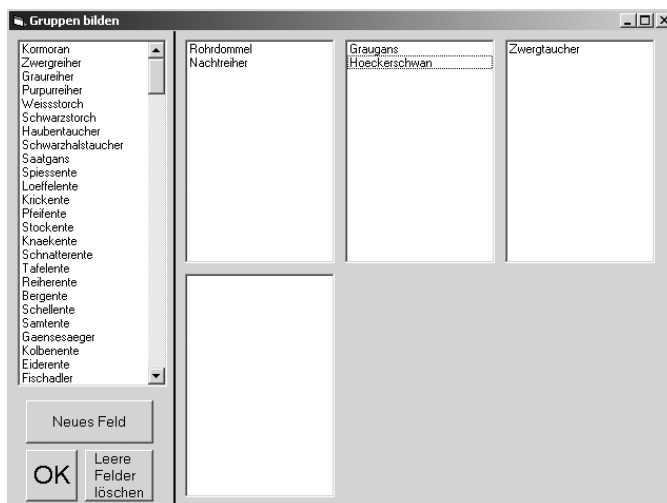


Abbildung 23: „Paralleles Sortieren“ am Computer

Als Mass für die Ähnlichkeit zwischen zwei Objekten kann die Koinzidenz herangezogen werden, d.h. die Frage, ob zwei Objekte von einer Versuchsperson in dieselbe Kategorie sortiert werden oder nicht. Zwei Objekte erhalten dabei den Ähnlichkeitswert 1, wenn sie derselben Kategorie zugeordnet werden, und den Ähnlichkeitswert 0, wenn sie nicht derselben Kategorie angehören. Auf diese Weise errechnet sich pro Sortierlösung eine Binärmatrix für alle Objektpaare.

Hierarchieannahme. Bei der Modellierung von Kognitiven Karten (Wissensstrukturkarten) auf Expertenniveau ist folgender Punkt wichtig: Wie wir be-

reits des öfteren betont haben, gehen wir davon aus, dass der Mensch sein Wissen (insbesondere merkmalsbasiertes Sachwissen) vornehmlich kategorial und hierarchisch organisiert. Wenn wir von solchen hierarchischen Wissenssystemen ausgehen, so kann bei der Methode „Parallelen Sortieren“ nur ein Zusammenfügen mehrerer individueller Datenmatrizen dem Hierarchiegedanken gerecht werden, denn ein einzelner binärskalierter Datensatz kann nicht die Datengrundlage eines hierarchischen Systems sein. Die kognitionspsychologische Annahme ist diejenige, dass jede einzelne Person das hierarchische System beim Parallelen Sortieren auf der Ebene „schneidet“, auf der sie es für am plausibelsten hält. Die Summe über eine Stichprobe befragter Personen mit im Wesentlichen gleichem Wissen, ergibt dann eine Verteilung der Hierarchiestufen und somit einen Einblick in die Organisation eines Wissensbereiches für diese Population von Personen.

3.2. Kognitive Parallelität

Wenn verschiedene Personen (etwa zur Erstellung eines Expertenmodells) einen Objektbereich parallel sortieren, so wird unterstellt, dass ihre unterschiedlichen Sortierlösungen lediglich verschiedene Ebenen ein- und desselben Wissenssystems abbilden.

Theoretische Voraussetzung dafür ist, dass die befragten Experten *1. über ein einigermaßen ähnliches Merkmalswissen verfügen und 2. dass ihre Ähnlichkeitskriterien auf der Grundlage gleicher Handlungskonzepte beruhen*. Wir Menschen sind nämlich in der Wahl von Beurteilungskriterien absolut frei. Die Wahl der Beurteilungskriterien ist Frage der Handlungsanforderungen. Das folgende Beispiel einer Aufteilung aus Borge's ‚chinesischer Enzyklopädie‘ von 1966, macht dies anekdotenhaft deutlich. Es heisst dort, dass ...

[...] die Tiere sich wie folgt gruppieren: a) Tiere, die dem Kaiser gehören, b) einbalsamierte Tiere, c) gezähmte, d) Milchschweine, e) Sirenen, f) Fabeltiere, g) herrenlose Hunde, h) in diese Gruppierung gehörige, i) die sich wie Tolle gebären, k) die mit einem ganz feinen Pinsel aus Kamelhaar gezeichnet sind, l) und so weiter, m) die den Wasserkrug zerbrochen haben, n) die von weitem wie Fliegen aussehen. (Borges, 1966; zit. in Smith 2002).

Die Wahl einer bestimmten Wissensorganisation ist also Sache der Perspektive. Dabei gibt es, so macht das obige Zitat deutlich, nicht nur wissenschaftliche Aufteilungen um Bereiche „unserer“ Wirklichkeit systematisch zu erfassen, sondern auch solche, die uns in unserem Kulturkreis als mehr oder weniger beliebig erscheinen. Die Annahme Kognitiver Parallelität besagt nun, dass innerhalb einer untersuchten Population Einigkeit über diese Perspektive besteht.

3.3. Gewichte bei der Berechnung der Koinzidenzmatrix

Bei der Abbildung einer über eine Stichprobe gemittelten Koinzidenzmatrix durch Nonmetrische Multidimensionale Skalierung, wird vor allem die oberste hierarchische Ebene kategorial sichtbar. Vor allem mit wachsender Zahl der Elemente bestimmt der Einfluss der Proximitäten mit Objekten anderer Cluster zunehmend die Intracuster-Strukturen: Die genaue Position eines Objektes wird dann stärker durch die Relationen zu den vielen weit entfernten Objekten festgelegt als durch die Nachbarschaftsbeziehungen zwischen den wenigen Objekten innerhalb eines Clusters.

Dieses Faktum einer nur groben Kategorisierung wird dadurch verstärkt, dass sich in einer individuellen Koinzidenzmatrix, für die eine befragte Person nur wenige Kategorien gebildet hat, mehr Werte der Ausprägung 1 befinden als in einer binären Koinzidenzmatrix, die aufgrund einer Sortierlösung mit vielen Kategorien entstanden ist. Damit ist der Einfluss einer

sehr groben Sortierlösung, auf die über eine Stichprobe gemittelte Matrix, numerisch grösser als der Einfluss einer feineren Sortierlösung.

Diesem Prozess kann entgegengewirkt werden, indem man bei den individuellen Matrizen die Koinzidenzwerte proportional zur Clusterzahl gewichtet: Bildet eine Person viele Cluster, so fallen die einzelnen Matrixwerte bei ihr höher aus als bei einer Person, deren Matrix auf eine kleine Clusterzahl zurückgreift. (Technisch gesehen handelt es sich hierbei um ein Clusterge-
wichtungsverfahren einfachster Art: In der Binärmatrix mit Nullen und Einsen einer Person werden einfach alle Einsen durch die von der Person gebildete Clusterzahl ersetzt.)

Dieses Gewichtungsverfahren sollte (immer vorausgesetzt, die Annahme Kognitiver Parallelität gilt für den Datensatz) dazu führen, dass die Strukturen innerhalb der Cluster einer resultierenden NMDS-Karte, klarer herausgearbeitet werden und dass aber gleichzeitig die Gesamtstruktur der Karte erhalten bleibt. Diese doppelte Annahme wird im Folgenden anhand einer ex-post-facto-Analyse der Daten aus Kapitel 2 dieser Arbeit überprüft. Zuvor soll aber die gerade genannte Strukturlogik noch etwas ausgefaltet werden:

- Besteht *keine hinreichende Kognitive Parallelität*, so wird durch die veränderte Gewichtung individueller Matrizen der Einfluss der einzelnen eingeflossenen Sortierperspektiven zueinander verschoben. In der Folge führt die veränderte Koinzidenzmatrix zu einer substantiell veränderten NMDS-Lösung (mit einer anderen Grundorganisation der Karte).
- Bei *bestehender Kognitiver Parallelität* ist dies nicht der Fall: Würden mehrere zueinander „perfekt“ kognitiv parallele Expertenmodelle zusammengefügt (durch Mitteln der binären Koinzidenzmatrizen), so verhielte sich die Rangordnung der Ähnlichkeitswerte für Objektpaare unabhängig von der Gewichtung der individuellen Matrizen (Objektpaare, die in der ungewichteten Matrix die höchste Ähnlichkeit besessen haben, würden in einem hierarchisch gebauten System auch nach der Gewichtung immer noch die höchsten Werte aufweisen). Lediglich die Absolutwerte würden ein wenig auseinandergezogen (was für die vornehmlich ordinal orientierte NMDS ja ziemlich belanglos ist). *Strukturstabilität nach Gewichtung ist also ein starker Hinweis darauf, dass die Experten die hierarchischen Verknüpfungsregeln ein- und derselben Wissensperspektive beim Sortieren berücksichtigt haben.*
- Bei *hoher, aber nicht perfekter Kognitiver Parallelität*, kommt es bei den ungewichteten Matrizen zu einer Streuung im mittleren und hohen Detaillierungsgrad (in den Grundfesten der Wissensorganisation sind sich die Befragten ja einig). Diese Streuung verhindert – zusammen mit dem oben benannten Einfluss der Proximitäten aus anderen Clustern – eine detailgetreue Ordnung innerhalb der Cluster. Verstärkt man nun durch höhere Gewichtung die detaillierteren Sortierlösungen, so wird zugleich auch die Intraclusterordnung gestärkt (vorausgesetzt natürlich, es gibt eine solche in der Wissensorganisation der befragten Stichprobe). Diese Annahme lässt sich mit geeigneten Daten empirisch überprüfen.

3.4. Datenbasis und Hypothesen

In Kapitel 2 haben wir mit dem Erhebungsverfahren des Parallelen Sortierens zwei unterschiedliche ornithologische Wissenssysteme auf Expertenniveau erhoben. Es sind dies ein Wissenssystem „*Nahrung*“ mit 187 und ein Wissenssystem „*Habitat*“ mit 183 heimischen Vogelarten. 11 Experten waren gebeten worden, 221 Vogelarten nach dem Merkmal Nahrung, 8 Experten nach dem Merkmal Habitat zu sortieren. Da den befragten Personen in diesen Fällen das Sortierkriterium vorgegeben war, erscheinen die Daten besonders geeignet, die Frage Kognitiver Parallelität an ihnen zu untersuchen (sind Befragte in der Wahl der Hauptkriterien frei, dann drückt die Koinzidenzmatrix der Stichprobe ja das numerische Verhältnis

der grundsätzlichen und möglicherweise völlig unkorrelierten Sortierkriterien zueinander aus).

Bezüglich der Anzahl zu bildender Kategorien wurden den Experten keinerlei Vorgaben gemacht. Das führte zu einer breiten Varianz in den individuellen Sortierlösungen: Für das Merkmal Nahrung wurden dabei 1x5, 1x6, 2x8, 2x9, 1x13, 2x28, 1x32 und 1x39 Objektgruppen gebildet. Beim Merkmal Habitat waren es 1x5, 1x6, 1x10, 3x12, 1x15 und 1x21 verschiedene Objektgruppen. Manche der Experten unterteilen die 221 Vogelarten also in sehr wenige und andere in sehr viele unterschiedliche Gruppen auf. Tabelle 3 zeigt beispielhaft vier solche Aufteilungen zum Merkmal Nahrung.

Daten scheinen also geeignet zu sein. Da in der methodischen Herleitung u.a. auch der Einfluss der Proximitäten mit Objekten aus anderen Clustern auf die Intracusterstruktur betont wurde (und zwar steigend mit steigender Gesamtmatrixgrösse), soll die Überprüfung sowohl mit grossen als auch mit kleineren Objektmengen erfolgen:

Bei der grossen Objektmenge wird dabei lediglich auf die Frage der gesamthaften Strukturähnlichkeit geschaut; es wird also die Vorhersage Kognitiver Parallelität der befragten Experten überprüft (qualitativ anhand der Stabilität der prägnanten Struktur und numerisch im AverageLoss einer Prokrustes-Transformation).

Mit einer kleineren Objektmenge lässt sich zusätzlich dazu noch der Organisationsgrad der einzelnen Cluster überprüfen. Wir werden das am Beispiel des Merkmals Nahrung für zwei Objektauswahlen tun. (Wir wählen hier eine Objektziehung von 15 und eine von 20 Singvögeln, Abbildung 24). Beide Karten enthalten zwei Teilcluster, für die sich überprüfen lässt, ob eine bessere Intracuster-Organisation gelingt:

- Zum einen handelt es sich um ein Körnerfresser-Cluster, welches neben vier Finken auch den Sperling umfasst. In der ungewichteten Karte fallen diese fünf Arten mehr oder minder zusammen. Deswegen kann als Hypothese für die Überprüfung formuliert werden, dass sich der Sperling in den gewichteten Karten jeweils von den vier Finken absetzen sollte.
- Die zweite markante Gruppe in beiden Objektauswahlen besteht aus insektenfressenden Singvögeln. Hier wäre zu erwarten, dass sich diejenigen Arten, die im Winter am konsequentesten auf Körnernahrung umstellen (Kleiber und Blaumeise bzw. Kleiber und Kohlmeise) in den gewichteten Karten jeweils ausdifferenzieren.
- Als drittes kann zusätzlich in der 20er-Karte erwartet werden, dass sich am anderen Ende des Clusters auch der Buntspecht aus dem Insektenfresser-Cluster ausdifferenziert, weil er mit Larven und Maden, die er direkt aus dem Holz herausholt, doch eine separate Quelle für seinen Proteinbedarf anzapft.

Durch eine höhere Gewichtung der Sortierlösungen mit mehr Kategorien, sollten diese feinen Ausdifferenzierungen stärker zum Tragen kommen, ohne dabei allerdings die Gesamtstruktur der Karte massgeblich zu verändern (wir erwarten also einen niedrigen AverageLoss).

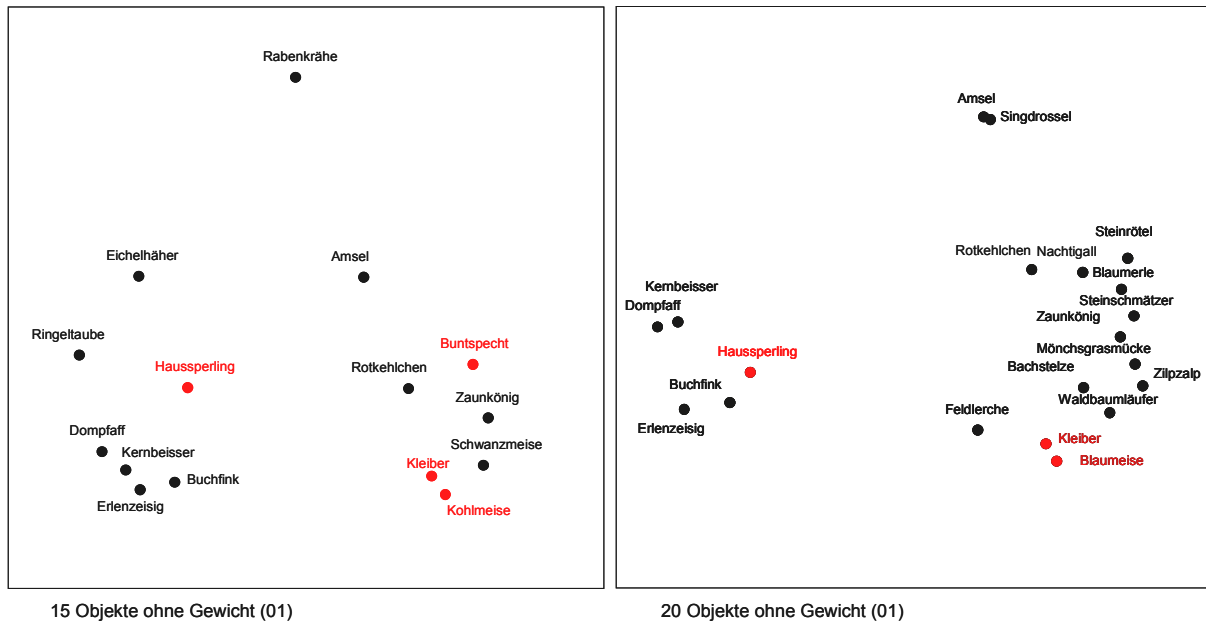


Abbildung 24: Zwei Objektziehungen aus den Sortierdaten „Nahrung“. 15 (links) und 20 (rechts) Singvögel bilden repräsentieren Teilausschnitte aus dem Gesamtpool von insgesamt 221 Sortierten Arten

Experte A	Experte B	Experte C	Experte D
Allesfresser	Carnivoren	Aasfresser	Allesfresser
Fisch oder Schalentiere	Fischfresser	Allesfresser	auf dem Boden suchende Insektenfresser
Fleischfresser	Früchte-Beeren	Fischfresser	Blattinsekten fressende Vögel
Insektenfresser	Herbivoren	Fleischfresser	Fischfresser
Samen-oder Pflanzenfresser	Insektenfresser	Grünnahrung	Fließwasserinsekten fressende Vögel
	Körnerfresser	Grünnahrung und Wassergetier	Fluginsekten jagende Vögel
	Omnivoren	Insekten-Beerenfresser	giftig behaarte Raupenfresser
	Weichtierfresser	Insekten-Fleischfresser	Greifvögel auch vogelfressend
		Insekten-Körnerfresser	Hautflüglerfressende Vögel
		reine Insektenfresser	hochspezialisierte Samenfresser
		Samen-Beeren	Hühner: Blätter, Beeren, Nadeln
		Samenfresser	Insektenfresser (kleine Insekten)
		Wassergetier	Insektenfresser (Beute aufspiessend)
			Kleinsäuger jagende nachaktive Vögel
			Maulwurfsgrillenfressende Vögel
			Mäuse und Amphibien
			Murmeltierfressende Greifvögel
			Saatgutfresser
			Samenfressende Hühner, weniger spezialisiert als Rauhfusshühner
			Samen
			Sumpfvögel: Insektenfresser
			Vegetarische Nahrung wird im Kropf zu Kropfmilch verarbeitet
			Waldinsektenfressende Vögel
			Wasservögel: Muschelfresser
			Wasservögel: Vegetarisch
			Watvögel in Riedgebieten nach Insekten suchend
			Watvögel Nahrung im Schlick suchend

Tabelle 3: Die Tabelle zeigt die von vier Experten gebildeten Kategorien zum Merkmal Nahrung bei 221 Vogelarten. Die Anzahl der gebildeten Kategorien reicht dabei von 5 bis 29 Kategorien. Die Kategorien, die die Annahme, werden auf unterschiedlichen Ebenen der Merkmalsdifferenzierung desselben hierarchisch organisierten Wissenssystems gebildet.

3.5. Ergebnisse

Aus den gewichteten individuellen Binärmatrizen werden neue gemittelte Koinzidenzmatrizen gebildet und einer NMDS unterzogen. Diese werden mittels Prokrustes-Transformation auf die ungewichteten Ausgangskarten gelegt. Die Resultate sind in Abbildung 25 und Abbildung 26 zu sehen.

Qualitativ ist festzustellen, dass für alle vier Karten die grundsätzliche Organisation bestehen bleibt: Bei den vollständigen Karten bleiben Dreiecksstruktur (Nahrung) und Simplex-Struktur (Habitat) bestehen; lediglich beim Habitat verschiebt es vier Vogelarten quer durch den Raum. Das erhöht hier den AverageLoss von 0.10 auf 0.16, jedoch handelt es sich auch hier noch um einen ausgesprochen kleinen Wert.

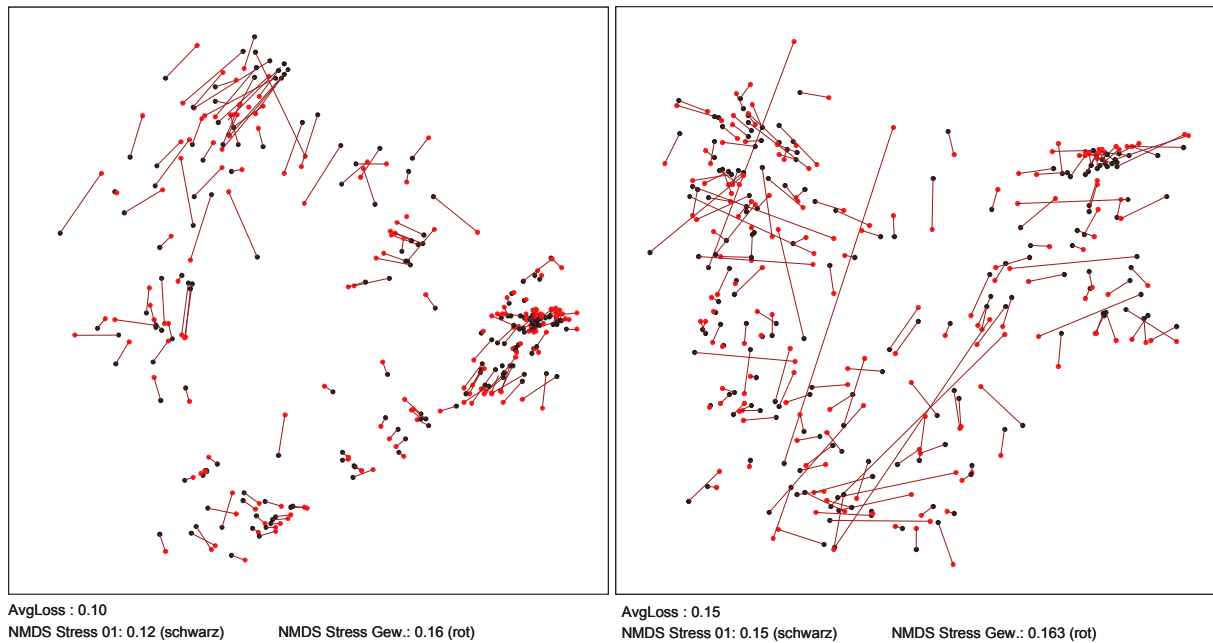


Abbildung 25: Prokrustes-Transformation der gewichteten mit den nicht gewichteten Nahrungs- und Habitatsmodellen. Objektziehung in beiden Merkmalsbereichen 187 Vogelarten.

Auch bei den Karten mit weniger Objekten bleiben die Grundstrukturen bestehen. In beiden Karten bestätigen sich aber gleichzeitig auch die Hypothesen hinsichtlich spezifischer Verschiebungen innerhalb der Hauptcluster (vgl. auch Abbildung 27): Der Sperling entfernt sich von den Finkenvögeln. Kleiber und Kohlmeise sondern sich von den Insektenfressern aber (und zwar in Richtung Körnerfresser), und in der 20iger Karte verlässt der Buntspecht dieses Cluster in die Gegenrichtung.

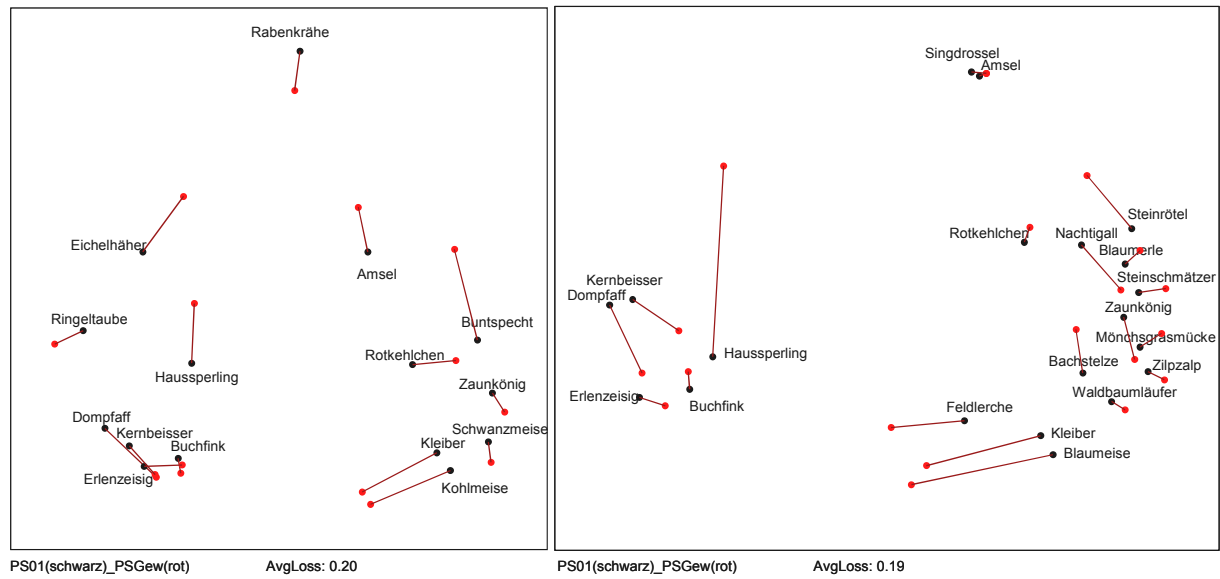


Abbildung 26: 15er und 20iger Objektziehungen aus dem Merkmalsbereich Nahrung, in ungewichteter und gewichteter Form miteinander prokrustet.

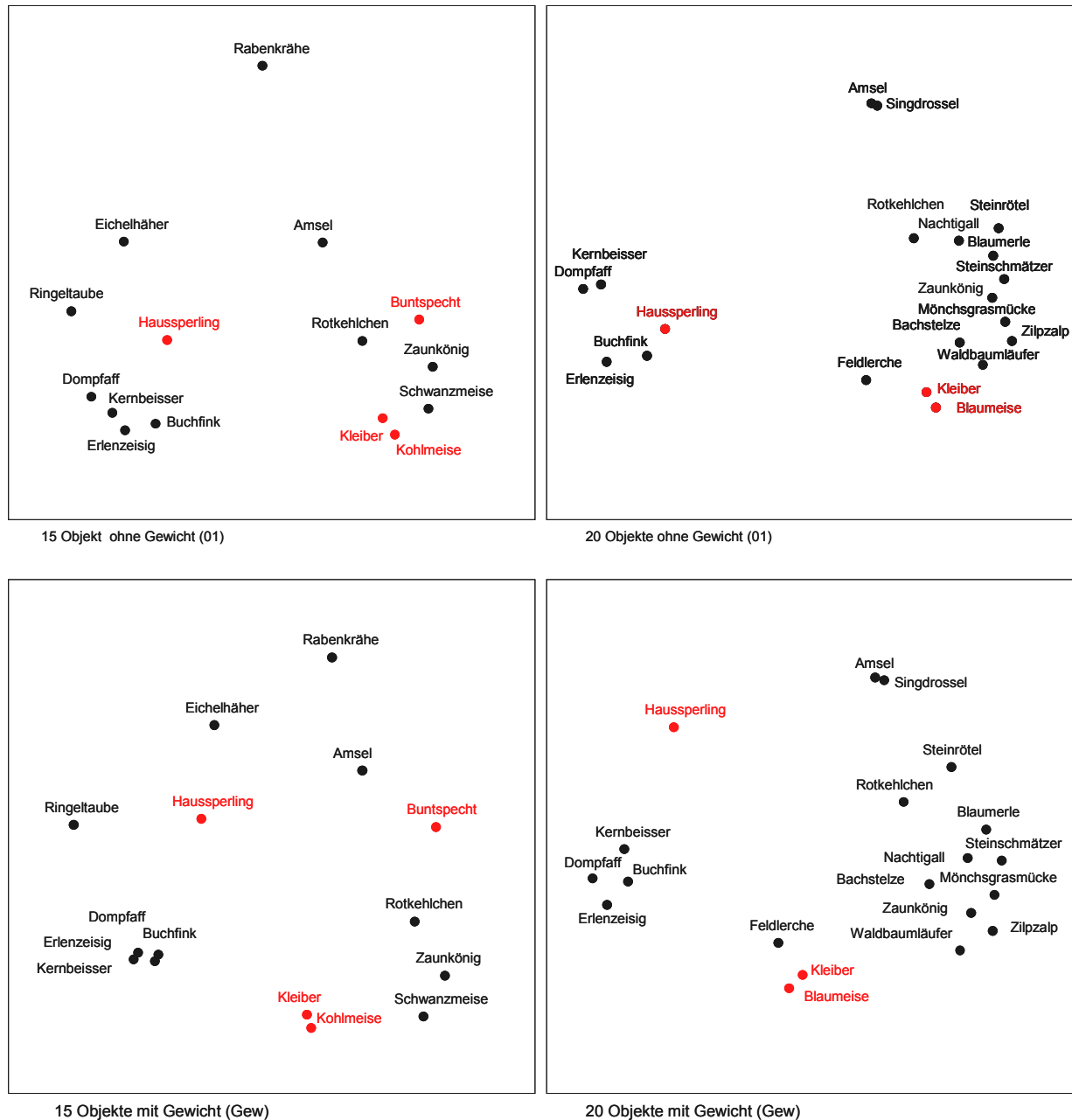


Abbildung 27: 15er und 20iger Karte vor und nach Gewichtung der einzelnen Sortierlösungen. Der Sperling entfernt sich von den Finkenvögeln. Kleiber und Kohlmeise sondern sich von den Insektenfressern aber (und zwar in Richtung Körnerfresser), und in der 20iger Karte verlässt der Buntspecht dieses Cluster in die Gegenrichtung.

3.6. Diskussion

Das Gewichten der individuellen Koinzidenzmatrizen anhand der Zahl der gebildeten Kategorien hat genau die Effekte gebracht, die bei ausgeprägter, aber nicht perfekter Kognitiver Parallelität prognostiziert waren: Die Grundstrukturen sind äusserst stabil geblieben, innerhalb der Cluster hat sich der Detailliertheitsgrad aber massiv erhöht.

Mit dem Konstatieren Kognitiver Parallelität lässt sich die Aussage machen, dass die Stichprobe der befragten Experten über eine Wissensbasis verfügt, die auf einem Baum beruht,

welcher auf unterschiedlichen Detailliertheitsebenen geschnitten werden kann. Man könnte es als Modell auch so formulieren, dass das Expertenwissen dieses Inhaltsgebietes „hierarchisch“ organisiert sei (wobei der Begriff „hierarchisch“ dann im Sinne einander inkludierender Mengen verstanden ist und nicht etwa im Sinne einer Wichtigkeitshierarchie voneinander unabhängiger Merkmale).

Die formale Logik der Kognitiven Parallelität selbst ist mit dem präsentierten Befund freilich nur zur Hälfte überprüft. Es müsste in einem weiteren Forschungsschritt anhand anderer Datensätze gezeigt werden, dass in Fällen fehlender Kognitiver Parallelität, in denen die Befragten im Parallelen Sortieren völlig unterschiedliche Merkmale heranziehen, die NMDS-Karten durch den Gewichtungsprozess massgeblich verändert werden. Geeignete Daten dazu könnten solche sein, bei denen fundamental unterschiedliche Sortierkriterien beim Parallelen Sortieren zu erwarten sind (z.B. Verwandtschaftsbegriffe, die sich u.a. nach Geschlecht, Linealität und Generationenfolge sortieren lassen – alles drei unkorrelierte Grunddistinktionen, die ja erst das Feld der Verwandtschaftsbegriffe konstituieren).

Als mit dieser Studie geklärt angesehen werden darf jedoch der positive Einfluss des Gewichtungsalgorithmus auf die Intracuster-Struktur. Hier hat sich in der Neuberechneten NMDS-Karte eine strukturell gut sichtbare und inhaltlich äusserst sinnvolle Ausdifferenzierung innerhalb der kritischen Cluster ergeben, und so kann dieses Verfahren bei künftigen Studien mit dem Parallelen Sortieren empfohlen werden. Zum einen ist es in der Lage, differenziertere (und damit schlichtweg bessere) Expertenmodell zu produzieren. Zum anderen liessen sich damit auch neue Fragestellungen überprüfen wie z.B. die, ob ein Experte (und zwar als Individuum, nicht als Abstraktion einer Stichprobe) sein Wissen im eingangs definierten Sinne hierarchisch strukturiert hat. Dazu müsse man ihn bitten, mehrere Parallele Sortierlösungen derselben Objektmenge abzugeben (wobei diese Sortierungen nicht identisch sein dürften). Der Einsatz von Gewichten beim Berechnen der gemittelten Koinzidenzmatrix würde dann zeigen, inwieweit die resultierende NMDS-Karte stabil bleibt oder kippt – und man hätte ein Mass für die Hierarchische Organisation der Wissensstruktur. Das vorgeschlagene Verfahren eröffnet also – das sollte dieser kleine Ausblick gezeigt haben – neue Perspektiven für die psychologische Expertenforschung.

3.7. Literatur

- Smith, B. (2002). Aristoteles, Kant und die Quantenphysik. In R. Hagenhuber (Ed.), *Philosophie und Wissenschaft* (S.79-97). Würzburg: Königshausen und Neumann.
- Borg, I. (1983). Scaling: A review of the German scaling literature of the last fifteen years. *The German Journal of Psychology*, 7, 63-79.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhard.
- Shepard, R.N., Romney, K.A. & Netherlove, S.B. (Eds.), (1972). *Multidimensional Scaling: Theory and applications in the Behavioural Science*. New York: Seminar Press.
- Sixtl, F. (1982). *Messmethoden der Psychologie*. Weinheim: Beltz.

4.

Kognitive Adäquatheit von Wissensstrukturkarten
—
eine explorative Studie

Zusammenfassung

Wissensstrukturkarten können im Hinblick auf ihre leichte Interpretierbarkeit unterschiedlich prägnant sein. Als "kognitiv adäquat" ist eine Strukturkarte zu bezeichnen, wenn sie dem bereichsspezifischen Wissen eines Betrachters angemessen ist und deswegen von ihm "gelesen" werden kann. Naturgemäss sind stark geclusterte Strukturen bereits mit kategorialem Wissen interpretierbar, während stärkere Gleichverteilung von Objekten im Raum ein dimensionales Wissen erfordert. Aus diesem Grund wird vorgeschlagen, statt einer direkten Berechnung von Wissensstrukturkarten aus Proximitätsdaten zunächst eine Clusteranalyse zu rechnen und die Distanzen aus dem sich ergebenden ultrametrischen Baum dann als NMDS-Karte zu skalieren. Dieses Verfahren wird am Inhaltsbereich Ornithologie exemplifiziert. Ein Experiment mit ornithologischen Laien zeigt, dass Wissensstrukturkarten grundsätzlich eine effiziente Navigationshilfe darstellen. Das Erhöhen der Kategorialität führt jedoch nur bis zu einem gewissen Grad zu einer zusätzlichen Nützlichkeit. Anhand zweier unterschiedlicher Strukturformen lässt sich einordnen, wann eine Erhöhung der Kategorialität zu erhöhter kognitiver Adäquatheit führt und wann nicht.

4.1. Einleitung: Die Prägnanz von Wissensstrukturkarten

4.1.1. Das didaktische Potential von Wissensstrukturkarten

Merkmalsgestütztes Objektwissen lässt sich über Nonmetrische Multidimensionale Skalierung (NMDS) in einem mehrdimensionalen Raum strukturell darstellen, indem sehr ähnliche Objekte nahe beieinander und unähnliche weit voneinander weg positioniert sind (zur Verfahrenslogik vgl. Borg & Groenen, 1997; Läge, Daub, Bosia, Jäger & Ryf, 2005).

Abbildung 28 gibt ein Beispiel für eine solche Wissensstrukturkarte: Für 20 in der Schweiz heimische Vogelarten wurden paarweise Proximitätswerte erhoben, indem 6 Experten die 190 Objektpaare auf einer 9stufigen Skala nacheinander auf ihre Ähnlichkeit hin einschätzten. Dabei waren die Experten in der Wahl ihrer Ähnlichkeitskriterien frei. Die Proximitätsmatrizen wurden über die Stichprobe gemittelt und durch NMDS (RobuScal-Algorithmus, Läge et al., 2005) zweidimensional skaliert.

Zum besseren Verständnis der Struktur wollen wir uns diese Karte etwas genauer ansehen. Auf der rechten Seite befindet sich ein Finkencluster. Finken sind typische Körnerfresser (Kernbeisser, Dompfaff, Buchfink und Erlenzeisig). Der Erlenzeisig ist wahrscheinlich wegen seines feinen Körperbaus etwas zu den Meisen hin verschoben. Auch der Haussperling lebt vornehmlich von Körnern und Sämereien. Wie die Finken hat er denn auch einen eher kräftigen Schnabel, mit dem sich die harten Schalen der Körner leicht knacken lassen. Auf der linken Seite der Karte finden sich die Drosseln (Steinschmätzer, dann die engen Verwandten Blaumerle und Steinrötel, die ebenfalls eng verwandten Amsel und Singdrossel, sowie die als Paar ebenfalls verwandtschaftlich zusammengehörenden Rotkehlchen und Nachtigall). Den Drosseln ist gemeinsam, dass sie vornehmlich von Bodeninsekten und anderen Kleintieren am Boden sowie von Früchten und Beeren leben. Der Unterschied in der Ernährung stellt eine wesentliche Bedingung für Körperbau und Lebensweise dar, weshalb die Experten die Finken den Drosseln gegenüber stellen.

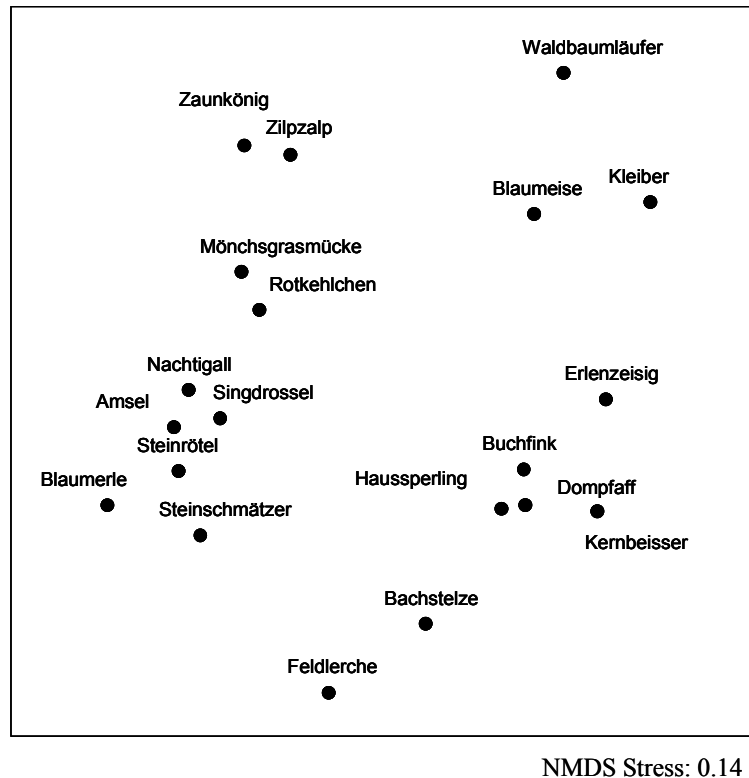


Abbildung 28: Expertenkarte (Globalähnlichkeiten) mit 20 heimischen Singvögeln

Im oberen Teil der Karte befinden sich vornehmlich Insektenfresser, die sich in eher geschlossenen Lebensräumen aufhalten (Hecken oder Wald). Dabei sollten Mönchsgrasmücke und Zilpzalp als sog. „Laubsänger“ eigentlich etwas näher zusammengehören, wobei erstere den kleinen Drosseln (Rotkehlchen und Nachtigall) in Art und Habitat sehr ähnlich ist und deswegen wahrscheinlich etwas weiter in diese Richtung platziert wird. Der Zaunkönig bildet den Abschluss dieser Busch- und Heckenbewohnenden Gruppe (am „südlichen“ Ende des Drosselclusters stehen mit Steinrötel, Blaumerle und Steinschmätzer eher die Vertreter offener Habitate).

Unten in der Karte befinden sich mit Feldlerche und Bachstelze zwei ebenfalls verwandte und hauptsächlich auf Insektennahrung spezialisierte Arten, welche offenes Gelände bevorzugen und deswegen gut hierher passen. Deren Antipode ist der Waldbaumläufer, welcher sein geschlossenes Habitat nicht verlässt. Unter ihm gruppieren sich Kleiber (der wie der Baumläufer vertikal entlang der Baumstämme läuft) und die Blaumeise. Beide gehören den Meisenartigen an (der Kleiber wird auch „Spechtmeise“ genannt). Obwohl sie eher feine, pinzettenförmige Schnäbel haben, werden sie von den Experten in die Nähe der Finken gerückt, weil die Meisen als Standvögel im Winter auf Körner und Sämereien ausweichen müssen und dann gemeinsam mit den Finken Lebensraum und Nahrung teilen.

Die auf Basis der Expertenbefragung entstandene Karte erweist sich also als ausgezeichnet interpretierbar. Trotz der enormen Vielfalt an möglichen Merkmalen, welche in die 6x190 Ähnlichkeitsurteile eingegangen sein dürften, reicht eine Verdichtung auf zwei „Hauptkomponenten“ (Nahrung/Schnabelform in der horizontalen und Habitat in der vertikalen Achse) aus, um die Semantik dieses Raumes in den wesentlichen Zügen zu beschreiben. Darüber hinaus ergeben sich spezifische Ähnlichkeiten im Körperbau, welche sich in besonderen Nachbarschaftsbeziehungen ausdrücken und summarisch als Clusterungen fassen lassen. Diese andeutungsweise zu findenden Clusterungen entsprechen weitgehend dem, was im zoologischen System als Familien innerhalb der Ordnung der Singvögel firmiert.

Wissensstrukturkarten machen also Beziehungen, Zusammenhänge und Abhängigkeiten von Sachverhalten leicht erkennbar. Aus pädagogischer Sicht sind sie deshalb äußerst interessant: Es ist anzunehmen, dass Wissensstrukturkarten bei der Vermittlung von Sachwissen ein hohes didaktisches Potential besitzen. Die Offenlegung globaler Zusammenhänge ermöglicht dem Lerner nämlich assimilatives Aufnehmen von Detailinformation in eine für ihn plausible Gesamtstruktur (vgl. dazu Kapitel 5, *Lernen mit Wissensstrukturkarten*).

4.1.2. Prägnanz: die Kognitive Adäquatheit von Wissensstrukturkarten

Wissensstrukturkarten können mehr oder weniger prägnant sein. Eine Wissensstruktur soll dabei genau dann als prägnant gelten, wenn sie in ihrer strukturellen Komplexität der Auffassungsgabe eines Menschen angepasst ist, wenn dieser sie also in ihren Grundzügen lesen und deuten kann. Diese Hinordnung von Prägnanz auf das individuell Intelligible könnte man auch als „Kognitive Adäquatheit“ bezeichnen. Und sie, das ergibt sich aus dieser Information, ist jeweils abhängig von allgemeiner Intelligenz und fachspezifischer Expertise. Für einen Experten (hier einen Ornithologen) wird also eine komplexe Kartenstruktur noch als prägnant erscheinen, während ein Laie bei derselben Karte möglicherweise bereits Interpretationsschwierigkeiten hat.

Erst mit ausreichender Zugänglichkeit hat eine Wissenskarte die Chance, für ihren Betrachter nützlich zu werden. Für einen Laien könnte sich etwa das oben angeführte Beispiel ohne nähere Anleitung als zu komplex erweisen, um es deuten zu können. Damit würde es seine mögliche Funktion als Orientierungs- und Lernhilfe einbüßen. Stimmen diese Überlegungen, so müsste für einen Laien eine Karte um so funktionaler sein, je einfacher – und das bedeutet: je kategorialer – sie strukturiert ist.

4.1.3. Vorschlag zur Vereinfachung von schwach klassifikatorischen Wissenssystemen

In Kapitel 2, *Die Modellierung von Klassifikatorischem Expertenwissen am Fallbeispiel der Ornithologie*, behandelte die Modellierung von klassifikatorischem Wissen für merkmalsgetütztes Objektwissen auf Expertenniveau. Exemplarisch wurden mittels der beiden Erhebungsverfahren 1. „Paralleles Sortieren“ und 2. „Ähnlichkeitsklassifikation mit Hilfe von BeschreibungsvARIABLEN“ Ähnlichkeitsdaten für die beiden ornithologischen Merkmalsgebiete „Habitat“ und „Nahrung“ erhoben (sie sind, wie das Einführungsbeispiel zeigt, für die Gesamtstruktur zentral).

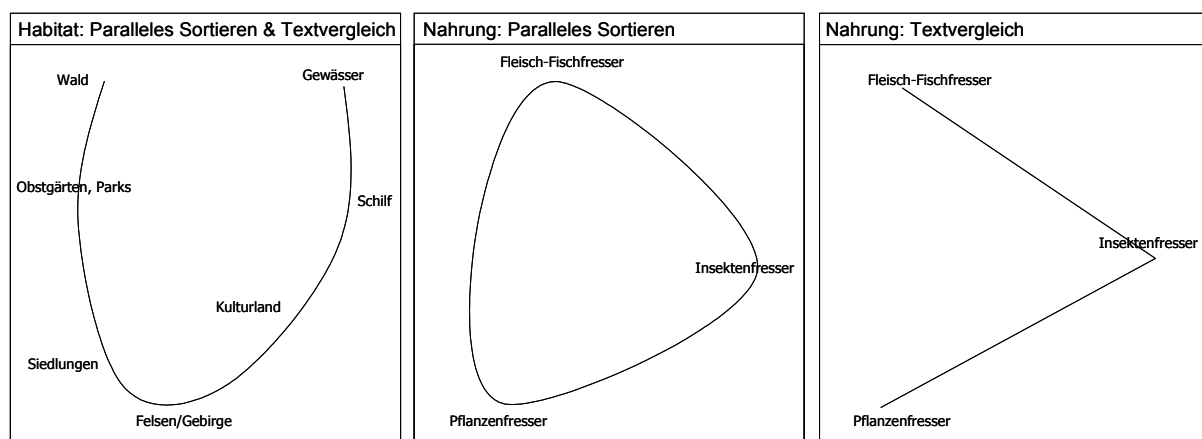


Abbildung 29: Vereinfachte Darstellungen der Expertenkarten aus Kapitel 2

Eine vereinfachte Darstellung der Wissensstrukturkarten findet sich in Abbildung 29. Beide Habitatskarten, sowohl jene aus dem „Paralleles Sortieren“ (Abbildung 10) als auch diejenige aus dem „Textvergleich“ (Abbildung 11) weisen eine gegen oben geöffnete Simplexstruktur auf (vgl. *Identifikation von Simplex Organisationen*, Gigerenzer, 1981, S.360). Die Simplexstruktur gilt allgemein als eine einfache Form kognitiver Organisiertheit. Die Bedeutungsträger werden entlang einer einzigen Bedeutungskomponente unterschieden.

Die Merkmalskarte „Nahrung“ aus dem Parallelen Sortieren (Abbildung 6) ist eine Annäherung an ein 3-Simplex (vgl. *Identifikation von nominalen Klassen*, Gigerenzer, 1981, S.364). Ein reguläres 3-Simplex beschreibt ein gleichseitiges Dreieck. Derartige kategoriale Lösungen sind in einer NMDS-Lösung dann zu erwarten, wenn die Ähnlichkeiten der Objekte innerhalb einer Äquivalenzklasse maximal und die Ähnlichkeit von Objekten verschiedener Klassen minimal ist und nur diese beiden Ähnlichkeitswerte auftreten. Die Merkmalskarte „Nahrung“ aus dem Parallelen Sortieren (Abbildung 7) ist definitiv eine Annäherung an eine solche Organisation. Allerdings sind die Übergänge zwischen den drei Eckpunkten („Pflanzenfresser“, „Fisch/Fleisch-Fresser“ und „Insektenfresser“) fließend, womit es sich nicht direkt um eine 3-Simplex handelt.

Die über Textvergleich erhobene Nahrungskarte zeigt eine große Streuung, mit leicht erhöhter Objektdichte im mittleren, rechten Teil der Karte (Insektenfresser). Wir finden zwar auch hier am oberen linken Randgebiet der Karte eher die Fisch-/Fleischfresser und im unteren linken Randgebiet eher die Pflanzenfresser. Objektcluster direkt sind jedoch keine erkennbar. Es stellt sich unter diesen Umständen die Frage, ob es sich für die Erhöhung der Verständlichkeit dieser Karten für fachliche Laien nicht lohnt, solch schwach-klassifikatorischen und, wie wir vermuten, kognitiv eher schwer zugänglichen Systeme, einer hierarchischen Clusteranalyse zu unterziehen und damit prägnanter zu gestalten.

Clusteranalyse. Es existieren unterschiedliche Verfahren zur Clusteranalyse. Die Festlegung auf eine geeignete Methode muss, wenn ein bestimmtes Ziel erreicht werden soll, von diesem her erfolgen. Und das Ziel ist in diesem Fall die Erhöhung der Prägnanz bei Beibehaltung sachlicher Korrektheit. Unter den Verfahrensklassen zu unterscheiden sind der klassische, eindeutige („scharfe“) Ansatz einer Clusteranalyse, bei dem angenommen wird, dass jedes Objekt in der betrachteten Menge von Objekten einer und nur einer Gruppe angehören kann, und der unscharfe Ansatz, bei dem eine mehrdeutige Mitgliedschaft der einzelnen Objekte in Bezug auf die einzelnen Gruppen zugelassen wird, z.B. Überlappende Gruppierung²⁶ oder Fuzzy clustering²⁷. Die Klasse der unscharfen Verfahren führt durch die erzeugten Ambiguitäten nicht zum Ziel, die Prägnanz zu erhöhen, und kann deswegen in diesem Fall vernachlässigt werden.

Bei den klassischen „scharfen“ Verfahren wird zwischen hierarchischen und partitionierenden Verfahren unterschieden. *Hierarchische Verfahren* fassen entweder ausgehend von der feinsten Gruppierung die einzelnen Objekte solange zusammen, bis alle in einer Gruppe sind (agglomerative Verfahren) oder beginnen mit einer Gruppe, die alle Objekte enthält, und spalten dann von dieser schrittweise Untergruppen ab, bis alle Objekte je eine Gruppe bilden (diversive Verfahren). Diversive Verfahren werden in der Praxis kaum verwendet. Gängige agglomerative hierarchische Verfahren sind: Single Linkage, Complete Linkage, Ward. Weitere

²⁶ **Überlappende Gruppierung:** Im Unterschied zu den scharfen Verfahren erfolgt bei der überlappenden Gruppierung nicht zwingende eine scharfe Aufteilung der Objekte auf Gruppen. Ein Objekt kann hier mehreren Gruppen gleichzeitig angehören. Dabei wird lediglich die Mitgliedschaft oder die Nicht-Mitgliedschaft festgestellt.

²⁷ **Fuzzy Clustering:** Ein Objekt gehört verschiedenen Gruppen mit einem bestimmten Anteil („membership“) an, wobei sich die Anteile zu 1 ergänzen. Gegenüber einer überlappenden Gruppierung werden also Aussagen zur Mitgliedschaft auf höherem Skalenniveau getroffen.

bekannte hierarchische Verfahren sind Zentroid²⁸, Average Linkage²⁹ und Flexible Methode. Hierarchisch-agglomerative Gruppierungsverfahren kommen im letzten Gruppierungsschritt immer zu einer Gruppenlösung. *Partitionierende Verfahren* beginnen mit einer Startlösung: die Objekte werden vorab in Gruppen aufgeteilt. Dann wird versucht diese Startlösung durch gruppieren der Objekte zu verbessern. Die Verfahren sind somit zweiteilig: Zum einen erfolgt die Auswahl der Startgruppen und zum anderen der Umgruppierungsprozess. Der am weitesten verbreitete Vertreter der partitionierenden Verfahren ist der K-Means-Algorithmus. Der offensichtlichste Unterschied zwischen hierarchischem und partitionierendem Clustering liegt darin, dass beim hierarchischen Clustering auch Subklassen entstehen (es existieren mehrere Ebenen), wogegen ein partitionierendes Clustering aus nur einer einzigen Ebene besteht (vgl. zu den Clusteranalyseverfahren: Rosenstiel & Heuser, 1998). Auf der Grundlage der Annahme potentiell hierarchischer Klassifikationssysteme bei der Wissensstrukturierung wählen wir ein hierarchisches Verfahren (im spez. Zentroidenclustering), welches sich zur Erhöhung der strukturellen Prägnanz ja auch auf einer bestimmten Hierarchieebene partitionierend interpretieren lässt.

Die NMDS besteht in einer monotonen Transformation der Unähnlichkeiten in euklidische Distanzen, die hierarchische Klassifikation aber in einer monotonen Transformation der Unähnlichkeiten in Distanzen, welche lediglich der ultrametrischen Ungleichung³⁰ zu genügen haben. Eine mittels hierarchischer Klassifikation identifizierte Cluster-Organisation ist deswegen auf die markantesten Unähnlichkeiten fokussiert und bildet diese, durch spätere Schritte unverrückbar, als die kleinsten Distanzen ab. Die NMDS hingegen berücksichtigt alle Proximitäten gleichmässig und ist somit sensibler für den Bereich mittlerer und kleiner Proximitäten. Eine hierarchische Clusteranalyse kann deshalb Informationen sichtbar machen, welche in der NMDS-Konfiguration allein nicht sichtbar wird (und umgekehrt; vgl. dazu auch Gigenzer, 1981, S. 371).

Das ultrametrische Modell. Um eine Vergleichbarkeit der erzielten Strukturen zu ermöglichen, leiten wir aus den Clusterbäumen wiederum Paarähnlichkeiten ab und stellen diese mittels NMDS zweidimensional dar. Die Paarähnlichkeiten bestimmen wir dabei nach dem ultrametrischen Modell. Das „ultrametrische“ Vorgehen (Hartigan, 1967; Jardin et al., 1967; Johnson, 1967) wurde neben dem Pfadmodell (Carroll & Chang, 1973, Carroll, 1976; Carroll & Pruzansky, 1975, 1980) von Marx & Läge (1995) unter anderem für die Berechnung der Distanzen beim „Hierarchischen Sortieren“ vorgeschlagen. Es baut auf der Vorstellung auf, dass die Distanz zwischen zwei Elementen in hierarchischen Bäumen als die „Höhe“ des ersten Knotens definiert wird, durch den die beiden Teilbäume, welche die betreffenden Elemente enthalten, miteinander verbunden sind. Beim ultrametrischen Modell sind alle Distanzen zwischen den Elementen zweier an einem Knoten auseinanderdividierter Mengen gleich groß. (Invers lassen sich die Distanzen als Ähnlichkeiten lesen.) Abbildung 30 zeigt den Clusterbaum der Clusteranalyse (Zentroidenclustering) über die Ähnlichkeitsmatrix, die als NMDS zu Abbildung 28 geführt hatte. Nach dem ultrametrischen Modell hat hier A zu B die Ähnlichkeit $p=9$, A und B haben zu C und D die Ähnlichkeit $p=6.7$. Alle Objekte im Cluster A bis L haben zu allen Objekten im Cluster M bis T die Ähnlichkeit $p=2.5$, u.s.w. (Hohe Ähnlichkeit wird, wie oben bereits dargelegt, auf diese Weise besonders fokussiert.) Die resultierende Proximitätenmatrix lässt sich mittels NMDS wiederum in eine Karte umsetzen.

²⁸ **Zentroid:** Bei diesem Verfahren wird jede Klasse durch ihren Schwerpunkt repräsentiert, und die Distanz zweier Gruppen wird durch die euklidische Distanz der beiden Gruppenschwerpunkte gemessen.

²⁹ **Average Linkage:** Bei diesem Verfahren wird die Distanz eines Objektes zu einer Gruppe als arithmetisches Mittel zu den Gruppenmitgliedern berechnet.

³⁰ Die ultrametrische Ungleichung besagt, dass die Distanz zweier Objekte nie grösser ist als deren Distanzen zu einem beliebigen dritten Objekt.

Abbildung 31 zeigt die mit RobuScal skalierten Paarähnlichkeiten aus dem ultrametrischen Modell, zweidimensional. Fünf Objektklassen werden hier klar unterschieden (in Abbildung 28 war eine solch klare Aufteilung der Objekte nicht gegeben). Die Wissensstrukturkarte zeigt damit eine in den Ähnlichkeitsdaten nicht real existierende klassifikatorische Organisiertheit und ist somit einfacher lesbar. Gleichzeitig damit besteht natürlich der Zwang einer eindeutigen Festlegung auf die Zugehörigkeit zu einem Cluster. Markant ist das für ein Cluster, welches sich als „Busch- und Heckenbewohner“ bezeichnen lassen könnte und das in der ursprünglichen Karte nicht sichtbar wurde: Die Nachtigall, in Abbildung 28 noch in der Nähe der übrigen Drosselvögel, weist hier eine enge Nachbarschaft zur Mönchsgrasmücke auf. Der Zaunkönig rückt in die Nähe des Rotkehlchens, bleibt aber weiterhin in der Nähe des Zilpzalps. Alle fünf zusammen bilden ein klar abgegrenztes Cluster. Durch die selektive Fokussierung hierarchischer Clusteranalysen auf die maximalen Proximitätswerte und durch die Abschottung der Clusterobjekte gegen Kovarianzen mit allen anderen Objekten werden die kategorienbildenden Merkmale durch das vorgeschlagene Verfahren akzentuiert, und genau das wirkt sich an dieser Stelle gut sichtbar aus: Der Habitus von Nachtigall und Rotkehlchen setzt sich durch, und die Tatsache, dass beide phylogenetisch zu den Drosselvögeln gehören, wird durch die Zwischenclusterung unterdrückt.

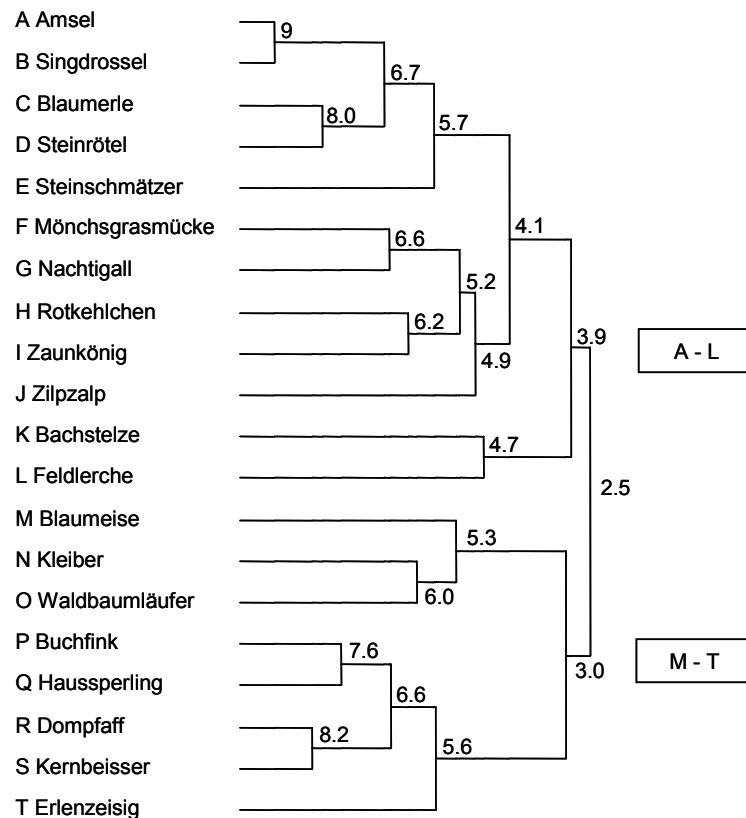


Abbildung 30: Clusterbaum der 20 heimischen Singvögel. Zentroidenclusteranalyse der Globalurteile.

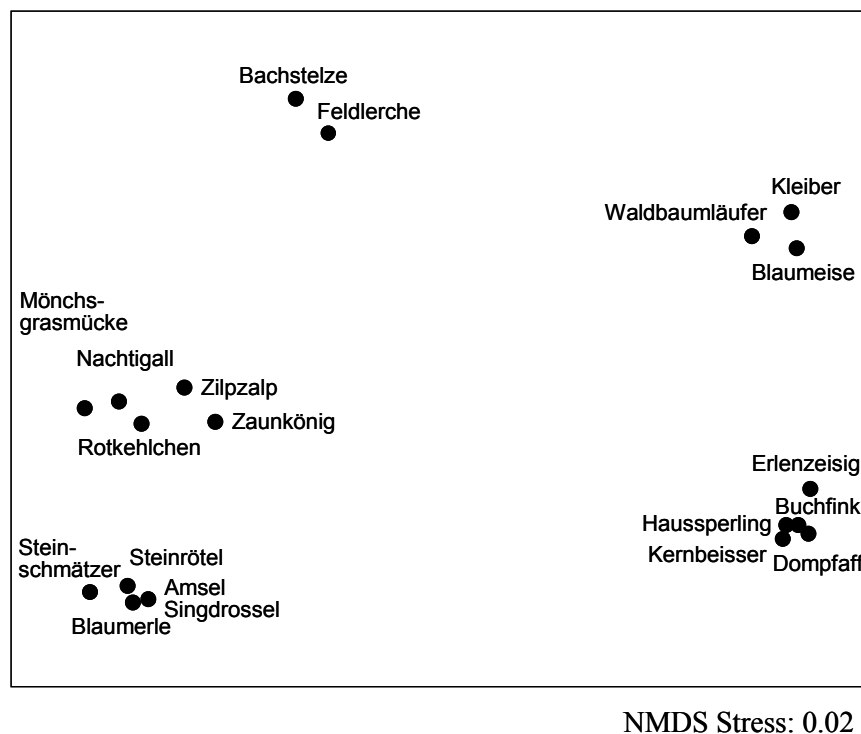


Abbildung 31: Expertenkarte (Globalurteile) mit 20 heimischen Singvögeln nach Clusteranalyse und Wiederaufbau der, nach dem ultrametrischen Modell gewonnenen, Paarähnlichkeiten.

4.2. Empirische Fragestellung

Zunächst einmal gilt es zu zeigen, dass Wissensstrukturkarten überhaupt in der Lage sind, einem Lerner als Orientierung zu dienen. Ist das nicht der Fall, so würden sich Fragen wie die nach der Steigerung der Verständlichkeit durch Erhöhung der Prägnanz als marginal erweisen. Aus diesem Grund sind im Experiment Wissensstrukturkarten einem (semantisch vernünftigen) Vergleichsmodell gegenüberzustellen.

In handelsgebräuchlichen zoologischen und botanischen Bestimmungsbüchern werden die Objekte dem Leser in fester Reihenfolge präsentiert (zumeist durch die phylogenetische Verwandtschaftsordnung definiert). Mit dieser Ordnung muss der Benutzer eines solchen Buches klarkommen. Also scheint das das geeignete Vergleichsmodell zu sein, mit dem sich die von uns vorgeschlagenen Wissensstrukturkarten messen können sollten. Nur wenn das der Fall ist, lohnt es sich, nach Modifikatoren für eine optimale Karte zu fragen.

Eine Wissensstrukturkarte sollte – so haben wir herausgearbeitet – den Lerner im Lernprozess dann am effektivsten unterstützen, wenn sie für ihn strukturell verstehbar ist. Der Lerner muss wissen, auf welcher semantischen Basis ein Objekt in der Karte seinen Platz findet. Eine Wissensstrukturkarte besitzt, so die Annahme, dann ein hohes didaktisches Potential, wenn Merkmalsdimensionen oder Äquivalenzklassen als solche erkennbar sind. Eindeutigkeit und Verständlichkeit sind die zentralen Eigenschaften, die eine prägnante Wissensstrukturkarte besitzen muss. Weiter oben haben wir dafür auch den Begriff „kognitive Adäquatheit“ verwendet.

Exemplarisch sollen die Wissenssysteme aus Kapitel 2 im Folgenden auf ihre kognitive Adäquatheit hin untersucht werden. Dabei interessieren wir uns unter anderem für die Frage, ob es gewinnbringend ist, schwach klassifikatorische Systeme einer hierarchischen Clusteranalyse zu unterziehen und sie damit für Laien prägnanter zu machen. Als didaktisches Hilfsmittel ist eine bis hin zur Karikatur vereinfachte Struktur eventuell sinnvoller als die objektiv-adäquate Verteilung der Objekte im Raum.

Kapitel 2 liefert zu den beiden Wissensbereichen Habitat und Nahrung je zwei unterschiedliche Strukturvarianten. Zum einen wurden Wissenskarten auf der Basis von Daten aus dem Parallelen Sortieren von Experten errechnet (und haben wegen der diskreten Datenform eine Tendenz zu Kategorien), zum anderen stammen sie aus einer automatisierten Analyse von Textüberlappungen, was eher eine breitere Verteilung der Proximitätswerte in der Datenmatrix zur Folge hatte. Im Hinblick auf ihre Prägnanz und damit auf ihre Nützlichkeit für den Lerner könnten sich diese beiden Karten systematisch voneinander unterscheiden. (Tun sie es nicht, so wäre dies – als „Nebenprodukt“ des im Folgenden berichteten Experimentes – ein Indiz für die Substituierbarkeit aufwändiger Expertenurteile durch automatischen Textvergleich, zumindest wenn es um den Zweck der Orientierungshilfe beim Lernen geht.)

Die in Kapitel 2 erhobene Nahrungsstrukturkarte aus dem Parallelen Sortieren tendiert in Richtung eines 3-Simplex, die über eben dieselbe Methode erhobene Habitatskarte bildet ein offenes Simplex (vgl. Abbildung 29). Da ein 3-Simplex eher eine kategoriale Struktur in sich trägt als ein als Skala interpretierbares offenes Simplex, würde es nicht verwundern, wenn beim 3-Simplex der Vorteil einer durch das Verfahren noch weiter kategorisierten Lösung eher im Experiment sichtbar würde als die Kategorisierung eines offenen Simplex.

4.3. Versuchsdurchführung

Die Prägnanz der Wissensstrukturkarten messen wir über die Qualität einer Karte als Navigationshilfe.

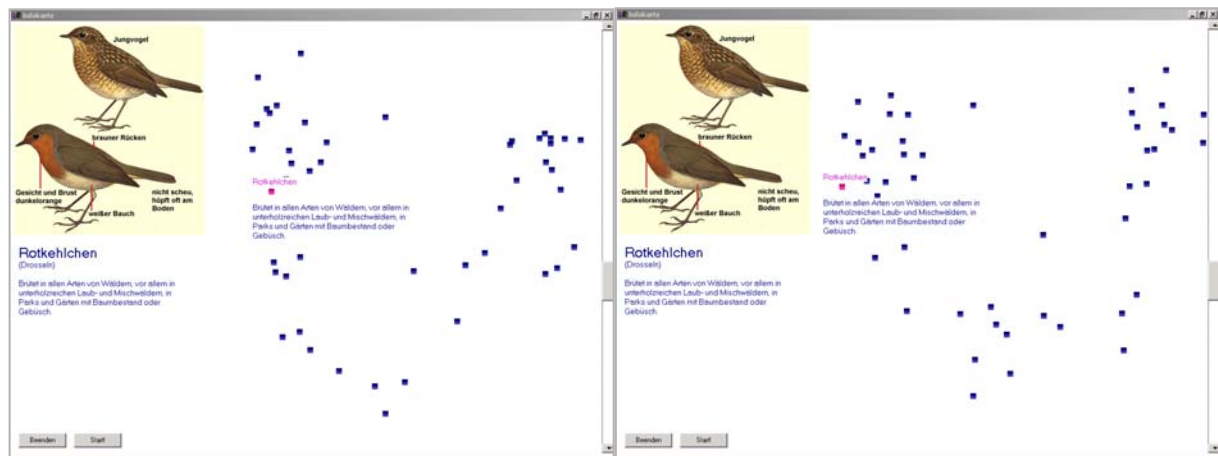
In einer ersten Phase des Experiments ließen wir die Versuchspersonen 10 Minuten (Material A) oder 15 Minuten (Material B) lang eine Wissensstrukturkarte frei explorieren, damit sie sich mit der Karte vertraut machen konnten. In der technischen Umsetzung am Computer ließ sich dafür ein Punkt in der Karte per „Mouse-over“ aktivieren. Abbildung 32 und Abbildung 35 zeigen die acht verwendeten Wissensstrukturkarten für das Gebiet der Ornithologie in ihrer technischen Umsetzung für diese Explorationsphase. Bei Aktivierung eines Punktes erschienen Artnamen, Merkmalsbeschreibung, Abbildung und die Familienzugehörigkeit des Vogels. Ansonsten waren die Karten stumm.

Nach der Explorationsphase wurde getestet, wie gut sich die einzelnen Merkmalstexte in der Karte wieder auffinden lassen. Der Versuchsperson wurde dazu ein Merkmalstext vorgelegt (respektive unterhalb der Punkteverteilung eingeblendet). Wiederum über Aktivierung der Punkte mit „Mouse-over“ sollte sie nun den zum eingeblendeten Merkmalstext adäquaten Merkmalstext ausfindig machen. Die Suchzeiten wurden dabei als kritische Variable registriert. Je schneller die Merkmalstexte gefunden werden, desto höher die Nützlichkeit der Karte.

Versuchspersonen. Insgesamt nahmen 64 Versuchspersonen an der Studie teil. Es handelte sich dabei um StudentInnen der Universität Zürich im Alter zwischen 22-35 Jahren ohne ornithologische Vorkenntnisse.

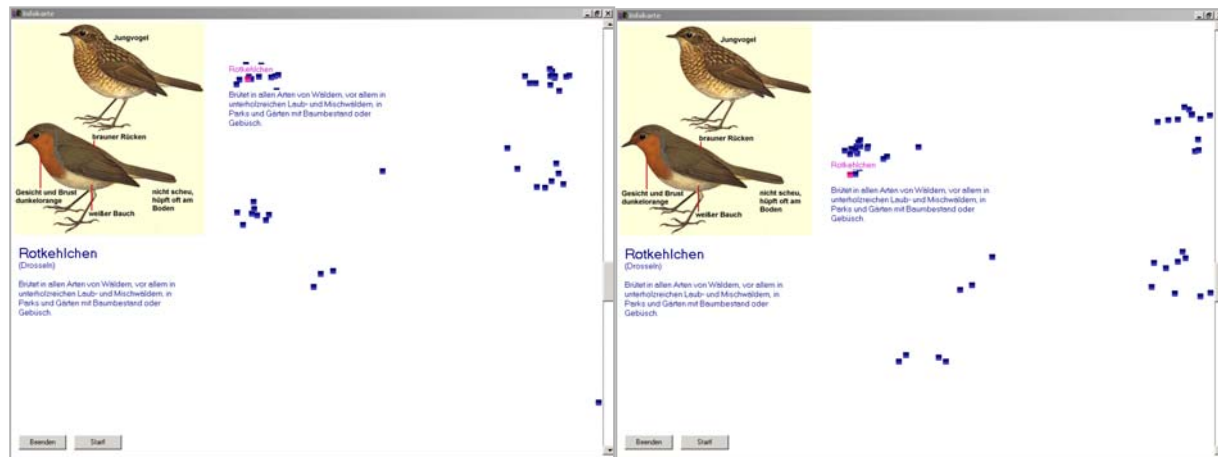
4.4. Material A: Habitatskarten

Vier Habitatsstrukturkarten wurden untersucht (Abbildung 32). Die beiden Wissensstrukturkarten PS/mtx („für Paralleles Sortieren / Matrix“) und TV/mtx (für „Textvergleich / Matrix“) sind Auszüge aus dem Datenmaterial der beiden in Kapitel 2 mittels „Parallelem Sortieren“ (PS) und „Textvergleich“ (TV) erhobenen Wissensstrukturen. Die Ähnlichkeitsmatrizen wurden dabei für dieses Experiment auf 48 Vogelarten reduziert und mittels NMDS neu skaliert. Die Auswahl der 48 Vogelarten geschah so, dass die ursprüngliche Simplexstruktur erhalten blieb.



PS/mtx: Simplexstruktur

TV/mtx: Simplexstruktur



PS/CL: kategoriale Struktur (hoch klassifikatorisch)

TV/CL kategoriale Struktur (hoch klassifikatorisch)

Abbildung 32: Bildschirmpräsentation der Wissensstrukturkarten im Experiment. Mit „Mouse-Over“ über einen Punkt in der Karte erscheinen links außen: Bild, Artname und Familienzugehörigkeit sowie Merkmalstext. Direkt beim Positionspunkt kann man zusätzlich Artname und Merkmalstext lesen. Der aktivierte Positionspunkt leuchtet rot auf.

PS/CL (für Paralleles Sortieren / Clusteranalyse“) und TV/CL (für „Textvergleich / Clusteranalyse“), sind die klassifikatorisch vereinfachten Strukturversionen dieser beiden Wissenssys-

teme. Die beiden Ähnlichkeitsmatrizen von PS/mtx und TV/mtx, wurden dabei einer Clusterranalyse (CL) unterzogen (Zentroidenclustering), die Paarähnlichkeiten nach dem ultrametrischen Modell bestimmt und mittels NMDS neu skaliert. Beide Karten sind im Vergleich zu den „unbehandelten“ Karten hochgradig klassifikatorisch.

Die Stichprobengröße pro Versuchsbedingung betrug $n=4$ Versuchspersonen. Über ein Abtasten der Positionspunkte musste die Versuchsperson nach einer 10-minütigen Explorationsphase 15 der insgesamt 48 Habitatstexte wieder auffinden. Die 15 Habitatstexte waren bei allen Versuchspersonen dieselben und wurden in immer derselben Rangfolge präsentiert. Quelle für die verwendeten Habitatstexte sind „Die Vögel Mitteleuropas“ von Singer (2000).

Buch-Bedingung. Als grundsätzliches Vergleichsmodell für alle Karten sollte die Struktur handelsüblicher Bestimmungsbücher dienen. Abbildung 33 zeigt die für das Computereperiment entwickelte Realisierung. Die Reihenfolge der Vogelarten entstammt hier dem Bestimmungsbuch von Singer (2000). Die Reihenfolge der Vogelarten entspricht der stammesgeschichtlichen Verwandtschaftlichkeit der Vögel. Das Buch bietet dem Benutzer verschiedene

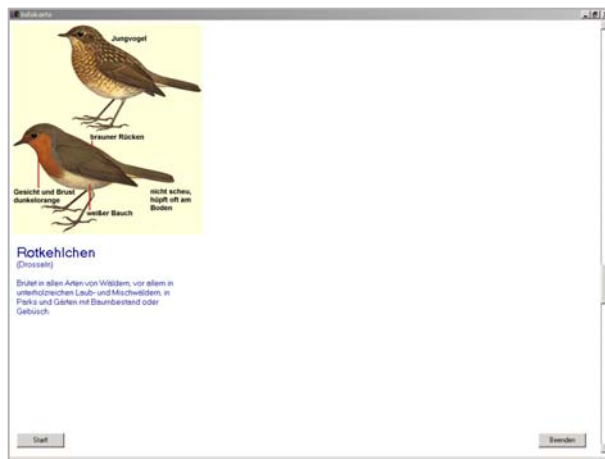


Abbildung 33: Buchbedingung

Explorationsmöglichkeiten an. Einfaches Vor- und Zurückblättern (in der Computerversion realisiert mit Vor- und Zurückbuttons sowie Oben-Unten-Pfeile auf dem Scrollbalken) oder auf irgendeiner Seite aufschlagen (Klick irgendwo auf dem Scrollbalken). Auch für die Buchbedingung wurde die Navigationsgüte erhoben. 12 Versuchspersonen wurden mittels dieser Versuchsbedingung getestet. Die Explorationsphase betrug – genau wie in allen Kartenbedingungen – 10 Minuten, und auch hier mussten 15 Merkmalstexte wiedergefunden werden. Diese Suche war hier natürlich über den Scrollbalken vorzunehmen.

Wir erwarten, dass das Wiederauffinden von Habitatstexten in der Buchversion viel Zeit in Anspruch nimmt. In der Wissensstrukturkarte lassen sich Äquivalenzklassen direkt ansteuern. Es sei denn, die über Wissensstrukturkarten transportierten Ähnlichkeitsinformationen widersprechen dem Merkmalsverständnis unserer Versuchspersonen, sind im Sinne obiger Definition also nicht prägnant.

4.4.1. Ergebnisse

Pro Versuchsbedingung wurden die Ausreißer (Suchzeiten außerhalb des Konfidenzintervalls von 95%) eliminiert.

Auswertung 1: Im Mittel brauchten die 16 Versuchspersonen mit Kartenbedingung 16.6 Sekunden, um einen Merkmalstext wiederzufinden, die 12 Versuchspersonen mit der Buchbedingung brauchten dazu im Schnitt 28.1 Sekunden. Die Kartennavigierer waren signifikant schneller ($t(28)=-6.56$; $p<0.001$, einseitig). Dies entspricht unserer Vermutung und zeigt, dass die Karten generell einen Vorteil für die Orientierung besitzen.

Auswertung 2 (Abbildung 34): Die Vergleiche zwischen den vier Kartenbedingungen beurteilen wir – aus Gründen der statistischen Überprüfbarkeit – für die 15 Merkmalstexte. Da die einzelnen Vogelarten sehr unterschiedlich schwierig zu bearbeiten sein könnten, ist es ange-

bracht, dabei t-Tests für abhängige Stichproben zu verwenden. Dazu werden alle Messwerte zu einer Vogelart in jeder Stichprobe gemittelt (pro Bedingung Wissensstrukturkarte je 4 Zeitmessungen pro Vogelart, in der Buchbedingung 12 Zeitmessungen).

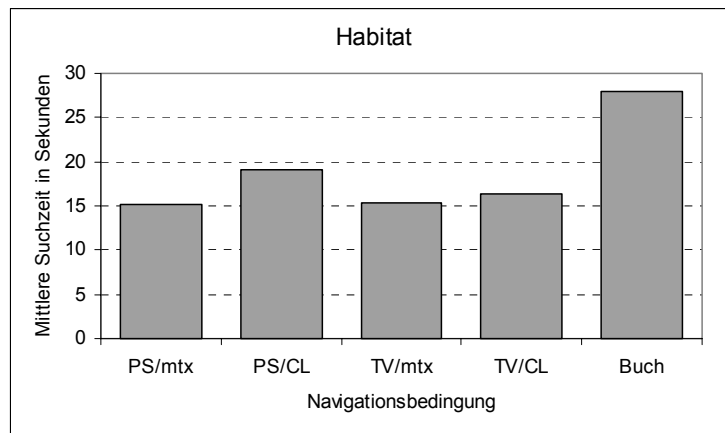


Abbildung 34: Mittlere Bearbeitungszeit für die 15 Merkmalstexte in unterschiedlichen Navigationsbedingungen. In den beiden Simplexstrukturen, PS/mtx und TV/mtx, lässt es sich am einfachsten navigieren.

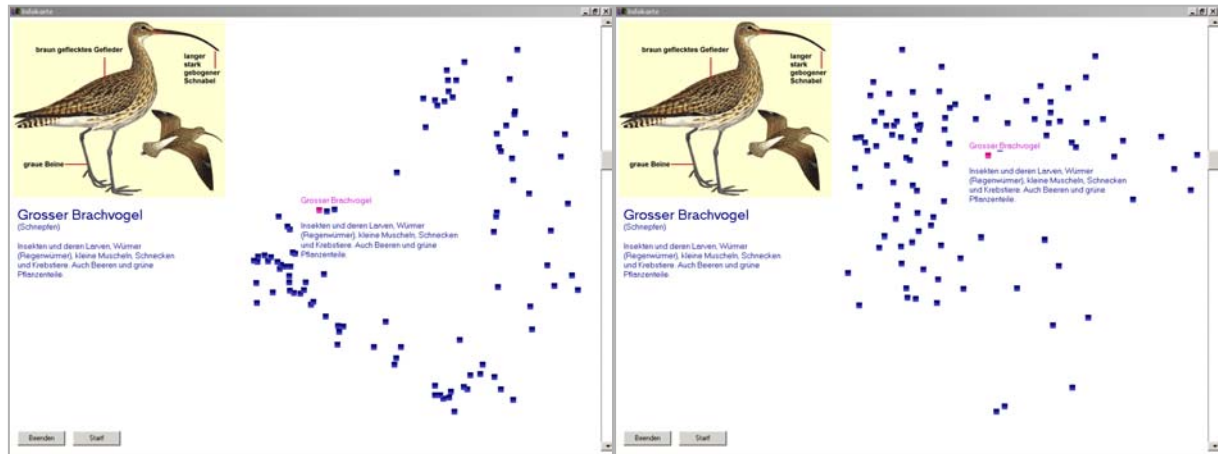
Entgegen unserer Vermutung, eine kategoriale Vereinfachung von schwach klassifikatorischen Karten durch Clusteranalyse fördere die Navigationsleistung, werden die Merkmalstexte in den beiden Simplexstrukturen PS/mtx und TV/mtx schneller wiederaufgefunden als in PS/CL und TV/CL. Wir haben ja bereits einleitend diskutiert, dass Simplexstrukturen als einfache kognitive Organisationen gehandelt werden (vgl. *Identifikation von Simplex Organisationen*: Gigerenzer, 1981, S.360). Scheinbar ist die Strukturprägnanz tatsächlich bereits so gut, dass eine Strukturvereinfachung hier das Verständnis für die Merkmalszusammenhänge eher senkt denn fördert. Signifikant wird dieses Ergebnis bei den Karten aus dem Parallelen Sortieren ($t(30)=-3.01$; $p<0.01$, einseitig), nu in der Tendenz signifikant aber bei den Karten aus dem Textvergleich ($t(30)=-1.41$; $p<0.1$, einseitig).

Wenn wir die Messdaten von PS/mtx und TV/mtx auf der einen und PS/CL und TV/CL auf der anderen Seite für die einzelnen Mermalstexte zusammenlegen (das sind dann pro Merkmalstext nicht mehr nur 4 sondern 8 Messungen), so lässt sich dieses Ergebnis gesamthaft bestätigen ($t(30)=-3.27$; $p<0.01$, einseitig). Die Orientierung in den geclusterten Karten war demnach tatsächlich schwieriger als in den beiden Simplexstrukturen.

Zwischen den beiden originalen Simplexstrukturen (aus Parallelem Sortieren und durch automatisierten Textvergleich) ist kein Prägnanzunterschied feststellbar. Beide besitzen sie dieselbe Navigationsgüte.

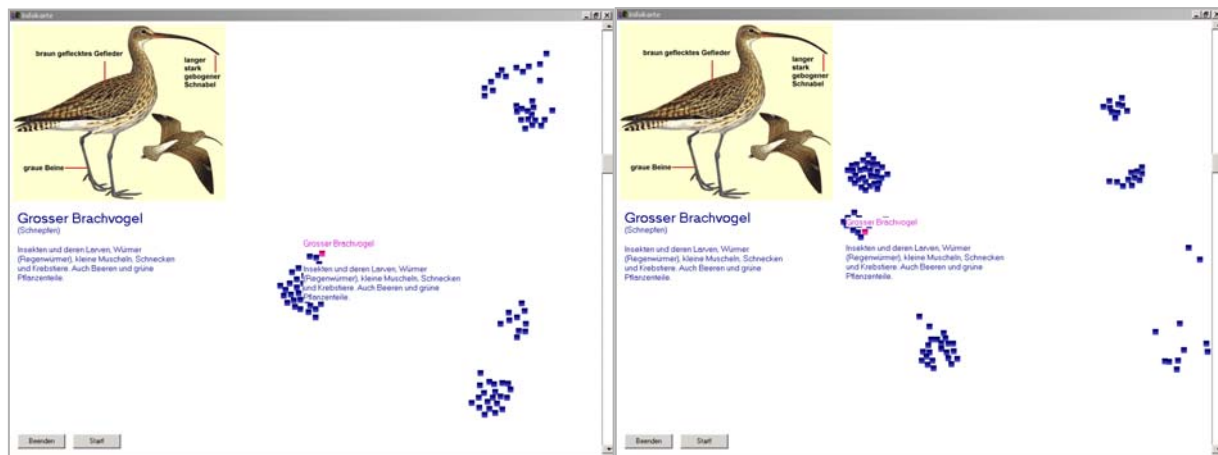
4.5. Material B: Nahrungskarten

Vier Nahrungsstrukturkarten werden hier untersucht Abbildung 35. Die Karten enthalten je 100 ausgewählte Vogelarten der Schweizer Vogelwelt. Material B unterscheidet sich von Material A sowohl durch den Wissensbereich als auch durch die Größe der Objektziehung.



PS/mtx: Annäherung an irreguläre 3-Simplex

TV/mtx: nicht-klassifikatorische Struktur



PS/CL: irreguläre 3-Simplex (hoch klassifikatorisch)

TV/CL: kategoriale Struktur (hoch klassifikatorisch)

Abbildung 35: Bildschirmpräsentation der Wissensstrukturkarten im Experiment. Mit „Mouse-over“ über einen Punkt in der Karte erscheinen links Außen: Bild, Artname und Familienzugehörigkeit und Merkmalstext. Direkt beim Positionspunkt: Artname und Merkmalstext (direkt beim Positionspunkt). Der aktivierte Positionspunkt leuchtet rot auf.

PS/mtx (für „Paralleles Sortieren / Matrix“) und TV/mtx (für „Textvergleich / Matrix“) sind Auszüge aus dem Datenmaterial der beiden 187iger Nahrungskarten aus Kapitel 2. Die auf 100 reduzierten Ähnlichkeitsmatrizen wurden mittels NMDS zweidimensional skaliert. PS/mtx zeigt eine an eine irreguläre 3-Simplex angenäherte Organisation, in TV/mtx streuen die jedoch Objekte hoch, so dass die durch Textvergleich gefundene Karte von der Organisa-

tion her weit von einer klassifikatorischen Lösung entfernt ist. Für die oben genannte Frage des Vergleichs der beiden Datentypen ist diese Bedingung also von besonderem Interesse.

PS/CL (für „ParallelesSortieren / Clusteranalyse“ und TV/CL (für „Textvergleich / Clusteranalyse“) sind die klassifikatorisch vereinfachten Wissensstrukturkarten von PS/mtx und TV/mtx. Die beiden Ähnlichkeitsmatrizen wurden mittels Zentroidenclustering (CL) analysiert, die Paarähnlichkeiten nach dem ultrametrischen Modell neu definiert und anschließend mit NMDS skaliert. Beide Karten sind hochgradig klassifikatorisch. PS/CL zeigt eine reguläre 3-Simplex-Organisation.

Durchführung. Im Unterschied zur Habitatsbedingung mussten die Versuchspersonen hier 14 Merkmalstexte wiederauffinden. Wegen der höheren Objektzahl in der Karte wurde die Explorationsphase von 10 auf 15 Minuten erhöht. Pro Kartenbedingung nahmen $n=6$ Versuchspersonen am Experiment teil, bei der Buchbedingung $n=12$.

4.5.1. Ergebnisse

Pro Versuchsbedingung wurden die Ausreißer (Suchzeiten außerhalb des Konfidenzintervalls von 95%) eliminiert.

Auswertung 1: Im Mittel brauchten die 24 Versuchspersonen mit Kartenbedingung 28.5 Sekunden, um einen Merkmalstext wiederzufinden, die 12 Versuchspersonen mit der Buchbedingung brauchten dazu im Schnitt 42.8 Sekunden. Die Suche in der Karte gelingt somit signifikant schneller als das Blättern im Buch, was das Ergebnis aus der Untersuchung mit Material A bestätigt ($t(36)=-3.43$; $p<0.01$, einseitig).

Auswertung 2 (Abbildung 36): Die Vergleiche zwischen den vier Kartenbedingungen beurteilen wir, wiederum über die Merkmalstexte ($m=15$). Es gilt wieder die Annahme, dass die Vogelarten nicht einfach Ziehungen aus demselben Topf sind, sondern womöglich sehr unterschiedlich schwierig zu bearbeitende Objekte. Deshalb rechnen wir mit t-Tests für abhängige Stichproben. Dazu werden wieder alle Messwerte zu einer Vogelart in jeder Sichtprobe gemittelt. Pro Bedingung Wissensstrukturkarte, sind das Pro Merkmalstext je 6 Zeitmessungen ($n=6$ Versuchspersonen).

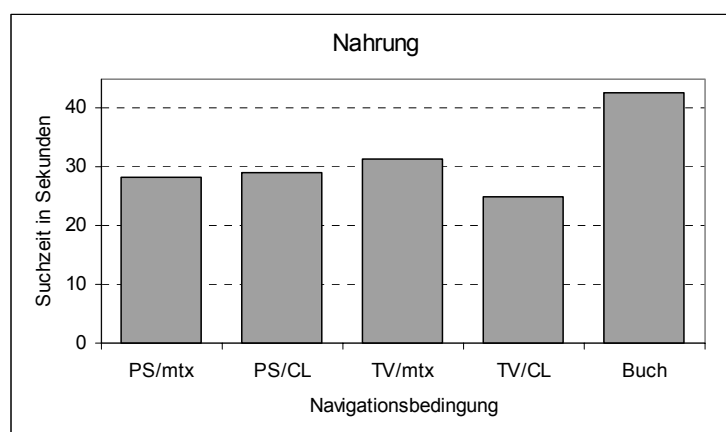


Abbildung 36: Mittlere Bearbeitungszeit für die 14 Merkmalstexte in unterschiedlichen Navigationsbedingungen. Auffällig und ebenso signifikant ist der Navigationsvorteil von TV/CL gegenüber von TV/mtx.

Die Vermutung, eine Vereinfachung von schwach klassifikatorischen Systemen fördere die Navigationsleistung, wird hier für TV/mtx bestätigt. Die Orientierung in TV/CL fällt den Versuchspersonen sehr viel leichter als in TV/mtx ($t(28)=1.49$; $p<0.05$, einseitig). PS/CL scheint

im Bezug auf die gestellte Aufgabe jedoch nicht prägnanter als PS/mtx, nur eben klassifikatorischer. Es kommt hier keine Leistungssteigerung zu Tage. Ein Leistungsabfall, wie bei Material A, findet allerdings auch nicht statt. Bezüglich der gestellten Aufgabe geben sich beide Karten also in etwa gleich prägnant ($t(28)=-0.22$; $p=0.41$, einseitig). Auch zwischen PS/mtx und TV/mtx lässt sich statistisch kein signifikanter Unterschied beweisen, obwohl die Richtung unseren Erwartungen entspricht ($t(28)=-1.33$; $p=0.10$, einseitig).

4.6. Diskussion

Zunächst einmal machen die Ergebnisse den enormen Vorteil der Wissensstrukturkarten gegenüber der festen linearen Reihenfolge des Bestimmungs“buches“ deutlich. Da die durch phylogenetische Verwandtschaft bestimmte Reihenfolge nur unvollkommen mit den Merkmalen Habitat und Nahrung korreliert (vgl. Kapitel 2.5, Abbildung 21), wird das Buch auf der Suche nach den Merkmalstexten von der ersten bis zur letzten Vogelart wohl einfach durchgeblättert. Für die Vermittlung bestimmter Merkmalsbereiche ist ein Abkoppeln des jeweiligen Systems von der phylogenetischen Verwandtschaftlichkeit deshalb durchaus sinnvoll. Die isolierte Darstellung von Teilbereichen, gibt dem Lerner die Möglichkeit, Systemzusammenhänge zu erkennen, die ihm über die gewohnte Art der Darstellung für lange Zeit verborgen blieben und die aber den Experten schlussendlich ausmachen. Mit den beiden Wissensstrukturkarten haben wir zur Realisierung eines solchen Vorgehens ein erwiesenermaßen vorteilhaftes Werkzeug vorgeschlagen.

Die bezüglich ihrer Navigationsgüte untersuchten Habitats- und Nahrungskarten sind in der Art ihrer Organisation sehr unterschiedlich. Die beiden Primärstrukturen „Habitat“ zeigen beide eine Simplexorganisation. Beide sind gleich prägnant und die Vereinfachung dieser beiden Wissenssysteme hatte einen deutlichen Leistungsabfall in beiden Systemen zur Folge.

Die Primärstruktur aus dem Parallelen Sortieren für das Merkmal Nahrung zeigt eine Dreiecksorganisation (Annäherung an einen irregulären 3-Simplex). Dagegen ist die, über Merkmalsbeschreibungen berechnete Struktur ganz und gar unklassifikatorisch. Schwachklassifikatorische Strukturen, so die Annahme, sind wenig prägnant und eignen sie sich also wohl eher schlecht zur Vermittlung von Objektmerkmalen. Das Navigieren in einer schwachklassifikatorischen Struktur sollte deshalb besondere Schwierigkeiten mit sich bringen (vgl. zur Verdeutlichung dieser Annahme Kapitel 2, Abbildung 8b). Ganz so hilflos wie erwartet verhielten sich die Versuchspersonen im Umgang mit der nicht-klassifikatorischen Nahrungskarte denn doch nicht, auch wenn sie einige Sekunden länger brauchten als bei den drei anderen Karten. Die Vögel wurden in der Karte aus dem Parallelen Sortieren zwar tatsächlich im Mittel 4.9 Sekunden schneller gefunden, statistisch ließ sich dieser Unterschied aber leider nicht bestätigen. Eine Vereinfachung der Karte aus dem Parallelen Sortieren, bringt, zumindest bezüglich der gestellten Aufgabe, weder Gewinn noch Verschlechterung. Eine deutliche Leistungssteigerung finden wir hingegen bei der über Clusteranalyse vereinfachten nicht-klassifikatorischen Nahrungskarte.

Die Resultate dieser Untersuchung zeigen nur in diesem letzten Fall einen Vorteil der durch Clusteranalyse erzielten strukturellen Vereinfachung der Karte. Die Generalhypothese „so einfach wie möglich“ muss hier schlussendlich abgelehnt werden, für den Fall einer strukturell wenig geordneten Karte könnte sich jedoch tatsächlich ein Vorteil ergeben. Doch diese Distinktion aus einem einzigen Experiment ableiten zu wollen, wäre vermutlich etwas zu gewagt. Hier muss weitere Forschung den Weg weisen.

Leistungsunterschiede zwischen den beiden Verfahren „Paralleles Sortieren“ und „Merkmalsvergleich“ tauchten nicht systematisch auf. Auch hier ist, was ein abschliessendes Urteil anbe-

langt, noch Vorsicht geboten, doch deuten die Ergebnisse der vorgelegten Studie darauf hin, dass zumindest für den Zweck der strukturellen Orientierung die aufwändigen Expertenurteile durch automatisierten Textvergleich ohne nennenswerten Qualitätsverlust substituierbar sind. Die Möglichkeit, dokumentiertes Expertenwissen in semantisch sinnvolle und praktisch brauchbare Wissensstrukturkarten umzusetzen, bietet vor allem die Chance, Objektbeziehungen zwischen Mengen beliebiger Grössenordnung zu bearbeiten.

4.7. Literatur

- Aichele, D. & Golte-Bechtle, M. (1997). *Was blüht denn da?*. Franckh-Kosmos: Stuttgart.
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling – Theory and applications*. New York: Springer.
- Carroll, J.D. & Chang, J.J. (1973). A method for fitting a class of hierarchical tree structure models to dissimilarities data, and its application to some “body parts” data of Miller’s. *Proc. 81st Ann. Conv. Am. Psychol. Assoc.*, 8, 1097-1098.
- Carroll, J.D. & Pruzansky, S. (1975). *Fitting of hierarchical tree structure (HTS) models, mixtures of HTS models, and hybrid models, via mathematical programming and alternating least squares. Presented at U.S.-Japan Semin. Multidimensional Scaling*, Univ. Calif. San Diego: La Jolla.
- Carroll, J.D. & Pruzansky, S. (1980). Discrete and hybrid scaling models. In E.D. Lantermann & H. Feger (Eds.), *Proceedings of Aachen Symposia on Decision Making and Multidimensional Scaling*. Berlin: Springer.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhard.
- Hartigan, J.A. (1967). Representation of similarity matrices by trees. *J.Am.Stat.soc.*, 62, 1140-1158.
- Jardine, C.J., Jardine, N. & Sibson, R. (1967). The structure and construction of taxonomic hierarchies. *Math.Biosci.*, 1, 173-179.
- Johnson, S.C. (1976). Hierarchical clustering schemes. *Psychometrika*, 32, 241-254.
- Läge, D., Daub, S., Bosia, L., Jäger, C. & Ryf, S. (2005). *Die Behandlung ausreisserbehafteter Datensätze in der Nonmetrischen Multidimensionalen Skalierung – Relevanz, Problemanalyse und Lösungsvorschlag*. AKZ-Forschungsbericht Nr. 21. Zürich: Angewandte Kognitionspsychologie.
- Marx, W. & Läge, D. (1995). *Der ideologische Ring*. Göttingen: Hogrefe.
- Rosenstiel, W. & Heuser, U. (1998). *Internetsuche und neuronale Netze: Stand der Technik*. Universität Tübingen, WSI-98-10, ISSN 0946-3852
- Singer, D. (2000). *Die Vögel Mitteleuropas*. Franckh-Kosmos: Stuttgart.

5.

Lernen mit Wissensstrukturkarten

Zusammenfassung

Relationale Wissensstrukturkarten sollten sich gut eignen, um als Strukturhilfen den Aufbau von Merkmalswissen zu unterstützen. Diese Hypothese wird am Beispiel ornithologischen Wissens experimentell überprüft. Das dazu entwickelte Multimedia-Lerndesign präsentiert zusätzlich zu dem zu lernenden Merkmalswissen (in diesem Fall die Nahrung von Vögeln), die relationale Einordnung des zu lernenden Objekts (der Vogelarten) in einer Wissensstrukturkarte der Schweizer Vogelwelt. Ein Längsschnitt-Experiment dokumentiert den im Zeitverlauf zunehmenden Lernvorteil aufgebauten Strukturwissens gegenüber herkömmlichem Merkmalslernen.

5.1. Einleitung

In der Wissenspsychologie geht man davon aus, dass Menschen ihr Wissen bezüglich verschiedenartigster Phänomene in Form von mentalen Modellen (Wissenskonstruktionen) organisieren. Seel (2000, Glossar) beschreibt die mentalen Modelle als kognitive Konstruktionen, mittels derer eine Person ihre Erfahrung oder ihr Denken derart strukturiert, dass sie eine systematische Repräsentation ihres Wissens erreicht, um subjektive Plausibilität zu erzeugen oder spezifische Vorgänge der Objekt- und Ereigniswelt in der Vorstellung zu simulieren. Der Mensch unterscheidet und differenziert Gegenstände, Ideen, Emotionen, Einstellungen, Konstrukte u.s.w. Der Mensch klassifiziert. Dabei haben durchaus verschiedene Ordnungssysteme nebeneinander ihre Existenzberechtigung.

Verantwortlich für das Entstehen solcher Ordnungssysteme (Wissensmodelle) sind Lernprozesse. Lernen, so Piaget (1976), ist ein Wechselspiel von assimilativen und akkomodativen Prozessen. Assimilation bezeichnet das Phänomen, dass das Individuum neue Elemente in sein Wissen integriert, ohne dabei sein bestehendes Wissen strukturell in Frage stellen zu müssen. Die Wissensstrukturen können dabei freilich in ihrem Umfang erweitert oder strukturell verfeinert werden. Akkomodation bezeichnet im Gegensatz zur Assimilation die Veränderung bestehender Wissensstrukturen im Sinne einer mehr oder weniger starken Umstrukturierung. Ausgelöst wird sie durch neue Informationen, deren Integration in das bestehende Wissenssystem zu unüberbrückbaren logischen Widersprüchen führen würde.

Eine wichtige Implikation dieser heute allgemein akzeptierten Erkenntnis ist, dass neues Wissen – und bestehe es nur aus einfachen Fakten – gleichzeitig auch immer strukturbildend ist (strukturweiternd bei der Assimilation und strukturverändernd bei der Akkomodation). Damit gibt es auch kein Lehren ohne das Vermitteln von Struktur, denn Wissen an sich ist Struktur. Im Bezug auf bestimmte Handlungskonzepte oder angestrebte Ziele kann Wissen aber unangemessen oder schlecht strukturiert sein oder sich als lückenhaft erweisen. Es kann auch vorkommen, dass Wissen ohne Ziel und Zweck vermittelt wird. Wenn der Lerner nicht weiss, warum er etwas lernt, fällt es ihm natürlich schwer, die oft nur „implizit“ sichtbaren Wissensstrukturen eines Tutors auch erkennen zu können.³¹

³¹ Nach Meinung der Autorin, gibt es leider viele Lehrkörper, welche sich über die Struktur ihres Wissens selbst nicht so ganz im Klaren sind. Diese, im Bezug auf Handlungskonzepte, fehlerhaft strukturierte Präsentation von Information kann das Lernen sogar oft noch erschweren. In diesem Falle ist eine so weit wie möglich unstrukturierte Präsentation des Datenmaterials wohl eher angebracht. Der Lerner soll sich seine Struktur den eigenen Bedürfnissen angepasst, selbst zusammenbasteln. Didaktischer Vorschlag für die Pilzkunde: Pilzkarten an Stelle von Pilzbüchern.

Der Nutzen eines vernetzten Wissens besteht ohne jede Frage in der Leichtigkeit, mit der einzelne Wissens Elemente im Gedächtnis aufgerufen werden können. Gleichzeitig ermöglichen assimilative Prozesse aber auch ein schnelleres und tieferes Enkodieren neuer Fakten. Diese Vorzüge sind evident, und sie kontrastieren mit dem enormen kognitiven Aufwand von Akkomodationsprozessen. Aus pädagogischer Sicht stellt sich deswegen die Frage, wie Wissensmodelle von außen, d.h. durch bestimmte pädagogische Massnahmen, gezielt in eine langfristig nützliche Richtung beeinflusst werden können. Norman (1983) beispielsweise sprach sich explizit dafür aus, den Lernenden durch die Bereitstellung von *konzeptuellen Modellen* bei der Konstruktion und Entwicklung geeigneter mentaler Modelle zu unterstützen.

Unter einem konzeptuellen Modell versteht Norman ein nach didaktischen Gesichtspunkten konstruiertes Modell, das in einer physikalischen Nachbildung, aber auch einfach in einer grafischen Darstellung des Originals bestehen kann. „Kognitive Karten“, die das Merkmalswissen einer Person über eine Anzahl von Objekten darstellen (Läge, 2001; Marx & Hejj, 1989; Marx & Läge, 1995), erfüllen diese Definition. Als dimensionale Gebilde zeigen sie die Relationen, die eine Person zwischen diesen Objekten sieht, in Form von kleineren oder grösseren Distanzen auf und lassen als Gesamtbild damit Rückschlüsse auf die strukturierenden (klassifizierenden) Merkmale zu. Die Kognitiven Karten erlauben es, das Wissen über ein spezifisches Objekt durch deren Position in der Gesamtstruktur nachzuvollziehen. In Kapitel 2 haben wir solche Strukturkarten am Beispiel des Wissensgebietes Ornithologie auf Expertenniveau erhoben.

Im Sinne eines Kognitiven Modells sollte eine derartige Struktur im Extremfall auch dann in der Lage sein, eine schnellere und tiefere Enkodierung (und damit ein besseres Lernen) neuer Wissens Elemente zu unterstützen, wenn sie selbst nicht bereits Bestand respektive Ausdruck des Wissens der lernenden Person ist.

Diese Hypothese wird im Folgenden (unseres Wissens nach zum ersten Mal) für Modelle in Form Kognitiver Karten experimentell überprüft. Für die Klarheit der Fragestellung scheint es uns wichtig zu betonen, dass damit nicht der Vorteil assimilativen Lernens an sich in den Blick genommen wird (denn die lernende Person verfügt ja noch gar nicht über ein elaboriertes System), sondern vielmehr die Orientierungskraft eines adäquaten geometrischen Modells, in welchem Ähnlichkeiten zwischen Objekten durch räumliche Nähe wiedergegeben ist.

Gelernt werden sollen diejenigen detaillierten Merkmalsinformationen, deren Summe die Position eines Objektes in einer Struktur bestimmt. Dazu werden wir auf die zweidimensionalen Expertenkarten aus Kapitel 2 zurückgreifen. Eignen sie sich zur effizienteren Vermittlung von Sachwissen? Durch die Offenlegung globaler Zusammenhänge hoffen wir dem Lerner assimilatives Lernen auf Merkmalsebene zu ermöglichen, wo noch gar keine adäquate Wissensstruktur ist. Strukturhilfe als Mittel zur Integration von bruchstückartigem Faktenwissen gewissermassen. Dabei soll neben einem kurzfristigen Effizienzgewinn beim Lernen auch die längerfristige Behaltensleistung mit ins Auge gefasst werden.

5.2. Methode

Aufbau und Durchführung des Experiments. In einer Längsschnittuntersuchung hatten die Versuchspersonen die Aufgabe, in drei Blöcken (Lernblock A, Lernblock B und Lernblock C, in Abbildung 37 als ovale Kreise dargestellt) zu jeweils 12 Vogelarten einen Text über deren Nahrung zu lernen.

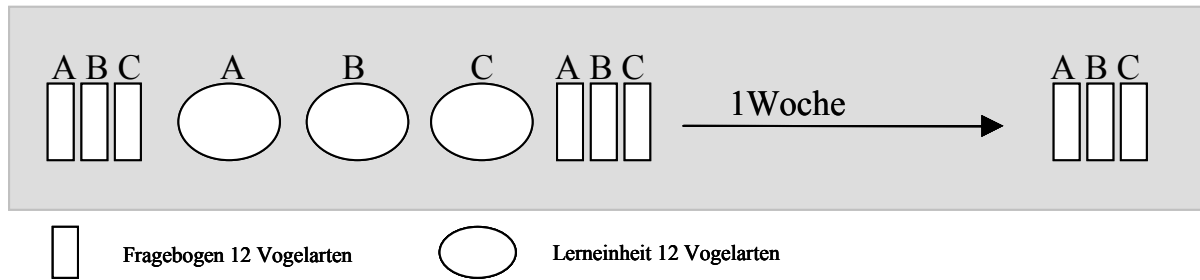


Abbildung 37: Schematische Darstellung des Ablaufs

Vor dem ersten Lernen mussten alle Versuchspersonen über die 3 x 12 zu lernenden Vogelarten einen Merkmalsfragebogen ausfüllen (Fragebogen A, B und C; in Abbildung 37 mit Rechtecken gekennzeichnet).

Jede Versuchsperson wurde zufällig einer von drei Lernbedingungen zugeteilt: Merkmalslerner (M), Strukturlerner (S) und Strukturpluslerner (SP). Danach wurde gelernt. Für jeden 12er-Lernblock (A, B, C) standen jeweils 15 Minuten Lernzeit zur Verfügung, in der sie sich zu jeder Vogelart deren Nahrung einprägen sollte. Die Nahrungstexte stammen aus dem Vogelbestimmungsbuch „Die Vögel Mitteleuropas“ vom Kosmos Verlag (Singer, 2000). Die jeweils 12 Vögel wurden am Computerbildschirm einzeln und bei allen Versuchspersonen und Versuchsgruppen gleich in immer derselben Reihenfolge präsentiert. Durch Anklicken des Buttons „Nächster Vogel“ wurde ein nächster Vogel mit dem dazugehörigen Nahrungstext aufgerufen. Die Präsentationszeit für jede einzelne Lernkarte konnte die Versuchsperson somit selbst bestimmen. Auch die Anzahl Lerndurchgänge war frei wählbar. Die Lerner konnten so ihr Lerntempo und die Lernstrategie nach eigenem Gusto frei bestimmen, sofern sie in der vorgegebenen Gesamtzeit von 15 Minuten blieben.

Den drei Lernblöcken folgte eine 10minütige Pause. Dabei wurden die Versuchspersonen bei einer Tasse Kaffee in belanglose Gespräche verwickelt. Nach der Pause mussten die drei Fragebögen (Block A, Block B, Block C) erneut ausgefüllt werden. Diese Wissenserhebung wurde eine Woche später dann nochmals wiederholt.

Drei Lernbedingungen (Abbildungen 38a, b und c.). Allen drei Lerngruppen gleich war die Präsentation von Bild, Artnamen und Merkmalstext. Den beiden Lernergruppen S (Strukturlerner) und SP (Strukturpluslerner) wurde zusätzlich eine zweidimensionale Wissensstruktur mit insgesamt 162 Punkten präsentiert. Jeder Punkt steht für eine Vogelart, und die Beziehungen zueinander sind durch die Nahrung dieser Arten bestimmt (Expertenmodell aus Kapitel 2). In dieser stummen Karte waren auch die 36 zu lernenden Vogelarten durch einen Positionspunkt vertreten. Bei den S-Lernern war immer gerade dann, wenn eine bestimmte Vogelart gelernt wurde, der in der Karte entsprechende Positionspunkt rot gekennzeichnet und mit dem Artnamen beschriftet (Abbildung 38b). Beim SP-Lerner wurden immer alle 12 Vogelarten des betreffenden Durchgangs mit Namen angezeigt. Zusätzlich leuchteten Positionspunkt und Namen des gerade zu lernenden Vogels rot auf (Abbildung 38c). Die Gruppe der Merkmalslerner (M) sah keine derartige Strukturkarte (Abbildung 38a).

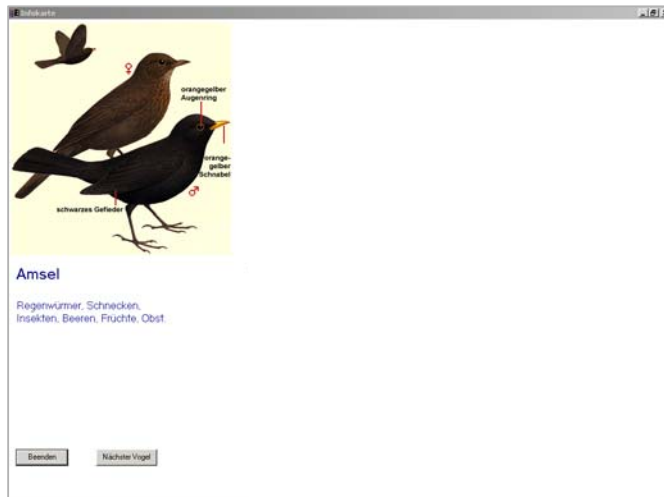


Abbildung 38a: Merkmal (M)

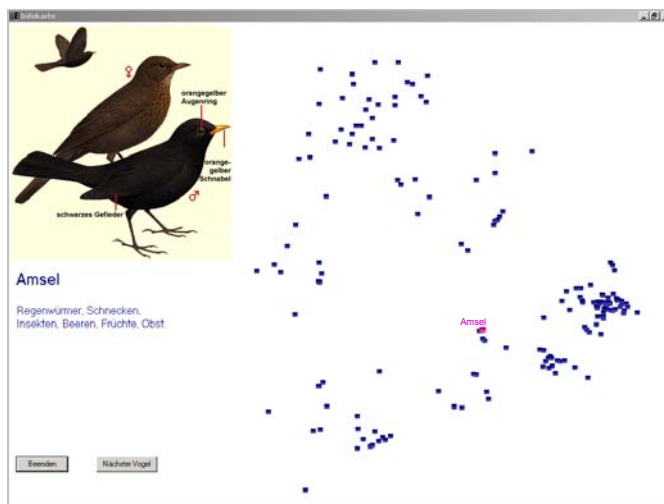


Abbildung 38b: Struktur (S)

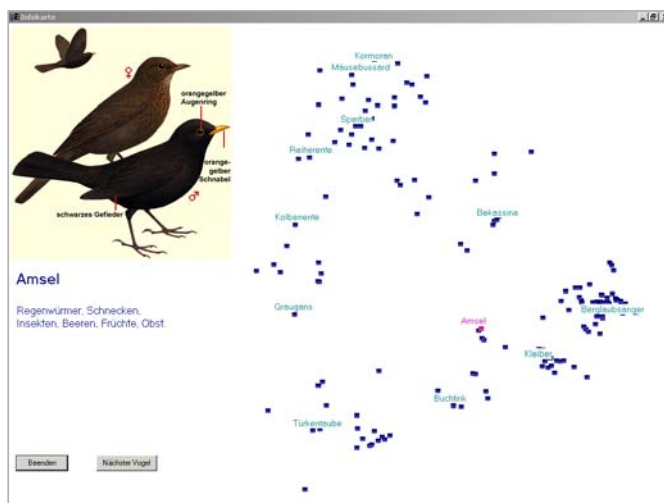


Abbildung 38c: StrukturPlus (SP)

Versuchspersonen. An der Untersuchung nahmen 3x12 Versuchspersonen ohne spezifische ornithologische Vorkenntnisse teil. Insgesamt beteiligten sich 27 (8, 9, 10) Männer und 9 (4, 3, 2) Frauen, hauptsächlich Studentinnen und Studenten der Universität Zürich, im Alter zwischen 20 und 35 Jahren.

Wissensstrukturkarte Nahrung. Die präsentierte Wissenskarte ist eine Teilstruktur der aus Expertendaten gewonnenen Nahrungskarte (vgl. Kapitel 2, Abbildung 6).

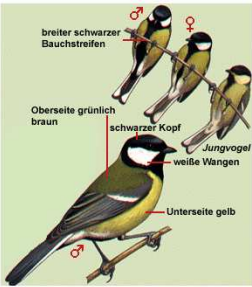
Fragebogen. Die drei Fragebögen waren einfache Nahrungslisten. Die Versuchsperson kreuzte pro Vogelart jeweils diejenigen Nahrungsmittel an, von der sie annahm, dass der Vogel sie frisst (Abbildung 39). Die Versuchsperson konnte dabei so viele Kreuze machen wie sie wollte, wurde aber bei Falschantworten mit Punkteabzügen bestraft. Bei einem Vielköstler konnte eine Versuchsperson mehr, bei einem Spezialisten weniger Punkte (dafür um so mehr Minuspunkte) erzielen. Die erzielte Punktzahl wurde der Versuchsperson immer direkt nach jeder einzelnen Vogelart gemeldet. Minimum- und Maximumpunkte in den drei Fragebögen insgesamt: Fragebogen A: Min. -211, Max. 102, Fragebogen B: Min. -208, Max. 102, Fragebogen C: Min. -199, Max. 132. Die Abfragerihenfolge der Vogelarten war in den Fragebögen zufällig.

In allen drei Objekt- und Merkmalspräsentationen waren gleich: Abbildung der Art, Artname, Nahrungstext & Reihenfolge der Präsentation.

Fragebogen

Versuchsperson: Heiri W.

Kohlmeise



OK Abbrechen

Insekten

☐ INSEKTEN/INSEKTENLARVEN (Blattläuse, Käfer, Grillen, Schnaken, Hummeln ..., Raupen, Maden, ...)

evtl. spezifizieren

☐ HOLZBEWOHNENDE INSEKTEN (Holzfressende Käfer und Larven ...)

☐ FLUGINSEKTEN (Mücken, Fliegen, Käfer im Flug ...)

☐ WASSERINSEKTEN (Wasserkäfer, Libellenlarven, Mückenlarven ...)

☐ AMEISEN

☐ RAUPEN

andere Kleintiere

☐ SPINNEN

☐ WÜRMER (z.B. Regenwürmer)

☐ SCHNECKEN

☐ MUSCHELN

☐ KLEINKREBBE

☐ FRÖSCHE

☐ KAULQUAPPEN

☐ REPTILIEN (Eidechsen, Blindschleichen ...)

Pflanzlich

☐ SAMEN (Grassamen, Getreidesamen, Baumsamen, Nüsse ...)

☐ GRÜNE PFLANZENTEILE (Gräser, Knospen, Triebe, Blüten, Sprosslinge ...)

☐ WASSERPFLANZEN

☐ BEEREN, FRÜCHTE

Fleisch und Fisch

☐ FISCH

☐ TOTE FISCH

☐ VÖGEL

☐ SÄUGETIERE (Mäuse, junge Hasen, Hasen ...)

☐ AAS (tote Säugetiere oder Vögel)

Abfall

☐ ESSENSRESTE, BROT ...

Abbildung 39: Nahrungsfragebogen

5.3. Erwartete Befunde

Der Unterschied zwischen der Strukturbedingung und der Strukturplus-Bedingung ist der, dass bei der SP-Bedingung die Ähnlichkeitsverhältnisse zwischen den 12 in einem Block zu lernenden Vogelarten permanent sichtbar sind. Der Lerner muss sich die Positionen der einzelnen Vogelarten nicht einprägen (wie in der S-Bedingung der Fall). Zudem ist eine Revisualisierung der bereits gelernten Strukturpositionen bei Hinzukommen eines neuen Objektes nicht notwendig, um dieses mit dem Gelernten zu verknüpfen. Damit sollte sich der kognitive Aufwand des Lerners deutlich verringern.

Hypothese 1: Die SP-Lerner nehmen das Sachwissen über die Nahrungstexte am effizientesten auf und erzielen deshalb am meisten Punkte im Fragebogen. Gegenüber den reinen Merkmalslernern weisen auch S-Lerner dank der Strukturhilfe einen Vorteil auf.

Hypothese 2: Der in Hypothese 1 erwartete Effekt tritt erst nach längerer Beschäftigung mit der Karte ein, ist also nach Lerneinheit A noch nicht oder nur rudimentär vorhanden, entsteht nach Lerneinheit B (oder wird deutlicher) und ist nach Lerneinheit C noch einmal grösser als nach B.

Der Grund für diese Vermutung liegt darin, dass sich die zunächst nur aus unbeschrifteten Punkten bestehende Karte erst mit zunehmendem Wissen, d.h. der Kenntnis der Position einiger Arten und deren Nahrung, für den Lernenden entschlüsselt werden kann. Integrative Kerninformation wie z.B. die, dass sich die Nahrungskarte in ihrer Grobstruktur über die drei Kategorien Fisch/Fleisch-Fresser, Pflanzenfresser und Insektenfresser interpretieren lässt, wird dem Lerner nicht Freihand geliefert, sondern der Lerner muss sich dieses Strukturverständnis erst selbst erarbeiten. Erst mit zunehmender Beschäftigung mit dem Lernmaterial sollte deshalb der Lernvorteil der SP-Lerner gegenüber den S-Lernern und der S-Lerner gegenüber den M-Lernern zunehmen.

Hypothese 3: Findet sich ein Lernvorteil der S- und SP-Lerner gegenüber den M-Lernern, so ist dieser zeitstabil: auch nach einer Woche Pause sollte er noch erkennbar sein. Der Grund für diese Annahme liegt darin, dass es eben die Vernetzung ist, die den Lernvorteil für das Merkmalswissen bewirkt.

5.4. Resultate

Vor dem Lernen unterschieden sich die drei Versuchspersonengruppen in ihrem Vorwissen nicht. Die Fragebögen wurden mit Mittel um die 25% +/- 3% richtig beantwortet. Das Basiswissen der Versuchspersonengruppen war also in allen drei Stichproben gleich niedrig.

Alle drei Hypothesen lassen sich bei rein visueller Betrachtung von Abbildung 40 bestätigen. Beginnen wir mit **Hypothese 2**, die besagt, dass der Lernvorteil der S- und SP-Lerner gegenüber den M-Lernern erst nach einer bestimmten Beschäftigungszeit mit der Struktur eintritt. Wir sehen im *Fragebogen A*, dass die M-Lerner gegenüber den S-Lernern sogar deutlich mehr Punkte erzielen, und zwar sowohl zum Zeitpunkt 2 ($t(24) = 2.07$; $p < 0.05$, zweiseitig) als auch zum Zeitpunkt 3 ($t(24) = 2.03$; $p < 0.1$, zweiseitig). Dieser Unterschied verschwindet im Vergleich zur SP-Gruppe weitgehend wieder.

Wir nehmen an, dass das in den Hypothesen nicht erwartete Einknicken der Strukturlerner in Durchgang A damit zusammenhängt, dass sie aufgrund der einzelnen Punkte in der zunächst ja völlig stummen Karte viel Mühe darauf verwenden, die Positionen einzelner Vögel zu lernen. Und das führt dann dazu, dass sie in den ersten 15 Minuten signifikant weniger lernen als die Merkmalslerner, die diese „Ablenkung“ nicht haben. Für die Strukturplus-Lerner ergibt sich diese Schwierigkeit ja nicht, denn sie sehen zumindest 12 bedeutungshaltige Punkte laufend in der Karte und brauchen sich deren Positionen deswegen nicht so intensiv einzuprägen. Zwar hat auch bei ihnen die Karte in der ersten Phase (A-Phase) leichte „Kosten“, dank der von vornherein besseren Integration des Wissens ist die Vergessensrate bei den SP-Lerner gegenüber den M-Lernern jedoch kleiner als bei den S-Lernern (vgl. Abbildung 40, Block A, Zeitpunkt 2 und 3). Der Wissensstand in Block A ist nach einer Woche dann gegenüber den M-Lernern wieder ausgeglichen.

Wie in **Hypothese 2** vermutet, zeigt sich aber dann der erwartete Haupteffekt: Die Bestätigung von **Hypothese 1**, dass die SP-Lerner auf Grund der Strukturhilfe einen grösseren Wissenszuwachs haben, wird erst im Lernblock B und später auch in C bestätigt. Statistisch absichern lässt sich diese Gerichtetheit der Effekte im Fragebogen B leider nicht. In Fragebogen C lässt sich der Lernvorteil der SP- gegenüber den M-Lernern immerhin noch auf dem Signifikanzniveau von 10% als statistische Tendenz bestätigen, und zwar sowohl zum Zeitpunkt 2 ($t(24) = -1.39$; $p < 0.1$, einseitig) als auch eine Woche später ($t(24) = -1.42$; $p < 0.1$, einseitig). Dieser Effekt ist also auch noch nach einer Woche vorhanden und ist somit eine Bestätigung von **Hypothese 3**. Auch wenn in der Grafik ein Lernvorteil der S-Lerner gegenüber den M-Lernern in Fragebogen B und C erkennbar wird, so lässt sich dieser Effekt wegen der enormen Streuungen zwischen den individuellen Fragebögen nicht statistisch absichern. **Hypothese 1** bestätigt sich damit also nur teilweise, und es bleibt unklar, ob die S-Lerner gegenüber den M-Lernern tatsächlich einen Lernvorteil hatten. Auch **Hypothese 3** gilt für die S-Lerner nicht. Der Vorteil der S-Lerner gegenüber den M-Lernern ist nämlich auch nicht zeitstabil. Die Strukturhilfe war, wenn überhaupt, nur eine kurzfristig wirksame Lernhilfe. Im Fragebogen C lässt sich der Lernvorteil der SP-Lernern gegenüber den S-Lernern immerhin noch auf dem Signifikanzniveau von 10% als eine statistische Tendenz bestätigen ($t(24) = 1.51$; $p < 0.1$, einseitig).

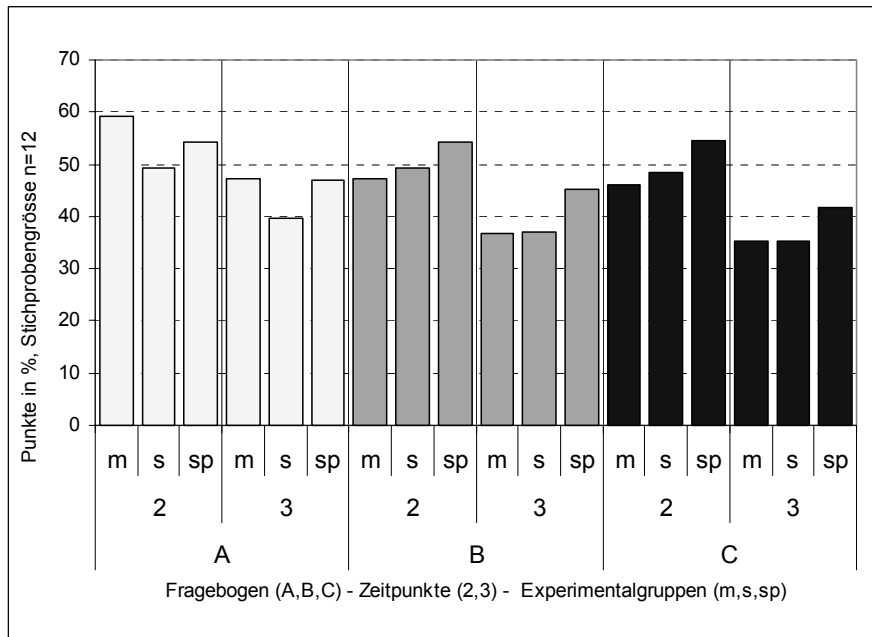


Abbildung 40: Die Grafik zeigt die im Mittel erreichte Punktezahl in % der möglichen Punkte, aufgeteilt nach den drei Fragebögen A, B und C, und darin wiederum in die Zeitpunkte 2 (direkt nach dem Lernen) und 3 (Nacherhebung eine Woche später) unterteilt. Innerhalb dieser Struktur sind die drei Versuchsgruppen mit je 12 Vpn positioniert: S-Lerner (s), SP-Lerner (sp) und M-Lerner (m).

Für das Experiment wurde mit der Stichprobengröße von jeweils 12 Versuchspersonen der minimale Datenumfang erhoben, um überhaupt t-Test-Statistik betreiben zu können. Die Streuungen zwischen den einzelnen Individuen erwiesen sich jedoch als beträchtlich. Um die gefundenen statistischen Tendenzen ein wenig stabiler überprüfen zu können, machen wir uns das Resultat zunutze, dass die Zeitpunkte B und C in den Mittelwerten einander verblüffend gleichen. Allein das spricht tatsächlich für die Nicht-Zufälligkeit der gefundenen Effekte. Da die Messungen mit etwas gutem Willen ja als unabhängig behandelt werden können (auch wenn dann in den Verteilungen natürlich von jeder Person nun zwei Messwerte stecken), werfen wir die Messzeitpunkte B und C zusammen, um die beiden **Hypothesen 1 und 3** auf einer etwas stabileren Datenbasis ($n = 24$) noch einmal statistisch überprüfen zu können.

Bestätigen lassen sich nun die folgenden Effekte: Die SP-Lerner erzielen in den Fragebögen B und C direkt nach dem Lernen deutlich mehr Punkte als die M-Lerner ($t(48)=-0.6$; $p<0.05$, einseitig) und auch mehr als die S-Lerner ($t(48)=1.32$; $p<0.1$, einseitig). Dasselbe Ergebnis findet sich auch eine Woche später: Die SP-Lerner machen deutlich mehr Punkte als die M-Lerner ($t(48)=-1.9$; $p<0.05$, einseitig) und als die S-Lerner ($t(48)=1.7$; $p<0.05$, einseitig). Jedoch lässt sich die Hypothese, dass auch die S-Lerner besser lernen als die M-Lerner, nach wie vor nicht durch einen signifikanten Unterschied statistisch untermauern.

5.5. Diskussion

Trotz enorm grosser Streuung im Fragebogen liess sich zeigen, dass es sich lohnt, Kognitive Karten (Wissenskarten) dem Lerner als Strukturhilfe und Mittel zur Integration von Faktenwissen zur Verfügung zu stellen. Dabei ist zu vermuten, dass die assimilativen Effekte um so grösser sind, je mehr der Lerner bereits weiss, je mehr er also die klassifikatorischen Möglichkeiten eines solchen Kognitiven Modells ausnutzen kann. Die Merkmalslerner hatten vor allem in der ersten Lernsitzung (A-Block) gegenüber den andern beiden Versuchsgruppen den

Vorteil, dass sie sich neben dem Merkmalslernen nicht noch zusätzlich mit der Geartetheit der Struktur auseinander zusetzen hatten. Die M-Lerner konnten sich hier voll und ganz auf die einzelnen Vogelarten und ihre Nahrungsmerkmale konzentrieren. Die beiden anderen Lernergruppen waren derweil zusätzlich damit beschäftigt, die vorliegende Wissenskarte inhaltlich mit Sinn aufzufüllen. Je mehr die Strukturlerner (S- und SP-Lerner) den Objektraum mit Bedeutung zu versehen in der Lage waren, desto effizienter konnten neue Vogelarten und ihre Nahrungsmerkmale wohl integriert werden. Kommt eine neue Vogelart dazu, so reicht aufgrund des bestehenden Strukturwissens nun schon ein Blick auf seine Position im Raum, um zu wissen, was dieser Vogel in etwa frisst; das Objekt mitsamt seinen Merkmalsausprägungen wird aufgrund der bestehenden Wissensstruktur assimiliert. Das Feintuning, die Spezifizierung im Sinne von Abgrenzung von anderen Vogelarten, kann nun immer noch über eine genauere Überprüfung des Merkmalstextes stattfinden. Eine erste Klassifizierung des Vogels über die Wissenskarte ist äusserst zeitsparend, ermöglicht also effizientes Lernen.

Die inhaltliche Interpretation der Karte fiel den SP-Lernern bestimmt sehr viel leichter als den S-Lernern. Die äusserst informationsarme Präsentationsform der „stummen“ Wissensstruktur bei den S-Lernern war bewusst gewählt worden, um die eigentliche Power von Kognitiven Strukturen im Unterschied von S und SP zu überprüfen. Ein Unterschied der S-Gruppe zu den Merkmalslernern war allerdings doch erwartet worden, jedoch fiel dieser direkt nach dem Lernen statistisch nicht signifikant aus und verschwand nach einer Woche vollständig. Das spricht dafür, dass die stumme Karte mit nur dem eingetragenen Zielobjekt in der knappen verfügbaren Zeit nicht so gut gelernt worden ist, dass sie eine echte Lern- und Behaltenshilfe gewesen wäre. Struktur benötigt zum Verständnis eben auch Semantik, und die setzt im Fall Kognitiver Karten voraus, dass ausreichend viele Objektpunkte in ihrer Position interpretierbar sein müssen. Das hier präsentierte Experiment ist in seinen Ergebnissen noch nicht das letzte Wort (dazu müssten langfristige und semantisch reichhaltigere Lernexperimente durchgeführt werden), aber es stellt einen vielversprechenden Anfang dar. Und der deutet sehr stark darauf hin, dass sich Kognitive Karten in die Reihe der lernfördernden Kognitiven Modelle einreihen lassen. Es würde sich also lohnen, den hier präsentierten Prototyp mit einem noch umfangreicheren Sachwissen auszustatten und dann unter den Bedingungen realer Ausbildung zu evaluieren.

5.6. Literatur

- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophische Fakultät der Universität Zürich.
- Läge, D., Daub, S., Bosia, L., Jäger, C. & Ryf, S. (2005). *Die Behandlung ausreisserbehafteter Datensätze in der Nonmetrischen Multidimensionalen Skalierung – Relevanz, Problemanalyse und Lösungsvorschlag*. AKZ-Forschungsbericht Nr. 21. Zürich: Angewandte Kognitionspsychologie.
- Marx, W. & Hejj, A. (1989). *Subjektive Strukturen*. Göttingen: Hogrefe.
- Marx, W. & Läge, D. (1995). *Der ideologische Ring*. Göttingen: Hogrefe.
- Norman, D.A. (1983). Some observations on mental models. In Gentner, D., Stevens, A.L. (Eds.), *Mental models*. Erlbaum: Hillsdale, NJ, 7-14.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Seel, N.M. (2000). *Psychologie des Lernens*. Ernst Reinhardt: München.
- Singer, D. (2000). *Die Vögel Mitteleuropas*. Stuttgart: Franckh-Kosmos.

6.

DIVE IN

—

**adaptiver Unterricht mit hierarchisch
ineinandergreifenden Wissensstrukturkarten**

Zusammenfassung

Im Bereich des Information Retrieval realisiert die Informationstechnologie die Abbildung facettenreicher Objektbereiche zunehmend in hierarchisch verknüpften Strukturkarten: Man erhält eine grosse Zahl von Wissensobjekten in ihren relationalen Bezügen (z.B. Textübereinstimmungen) in einer Karte abgebildet. Mit der Auswahl eines Kartenbereiches werden dann detailliertere Karten aufgebaut, die eine genauere Orientierung ermöglichen. DIVE IN überträgt die grundsätzliche Logik dieses Verfahrens auf Wissensstrukturkarten und ermöglicht somit ein kartenunterstütztes multimediales Merkmalslernen. Es ermöglicht dem Lerner die Navigation auf unterschiedlichen detaillierten Merkmalsebenen, je nach aktuellem Wissensstand über einzelne Objekte. Der Vorteil besteht darin, dass Detailwissen und Überblickswissen auf diese Weise parallel zueinander aufgebaut werden können. Exemplifiziert wird DIVE IN am Beispiel ornithologischen Wissens.

6.1. Hierarchisch gegliederte Wissensstrukturkarten

In den vorausgehenden Kapiteln wurden Wissensstrukturkarten eingesetzt, um Expertenwissen abzubilden und um Lerner auf der Basis dieser relationalen Information zu unterrichten. Die in diesem Kontext präsentierten Expertenkarten fielen enorm variabel in ihrem Umfang aus und erreichten bis an die 187 Objekte. Aus Gründen der Übersichtlichkeit für eine betrachtende Person kann eine geometrische Darstellung nur in einem niedrigdimensionalen Raum erfolgen, und so muss die aus der Karte ablesbare Merkmalsinformation notwendigerweise auf einige wenige Basisinformationen begrenzt bleiben. Im modellierten Expertenwissen selbst stehen jedoch weitaus detailliertere Informationen zur Verfügung, welche jedoch nur dann in einem niedrigdimensionalen Raum zum Tragen kommen, wenn die beherrschenden Basisinformationen für die skalierte Objektmenge relativ konstant ausfallen. Da die detaillierteren Informationen bei einer umfassenden Karte innerhalb der Hauptcluster also verloren gehen, müsste die Auswahl einer geeigneten Objektmenge dadurch geschehen, dass Hauptcluster der Gesamtkarte gewissermaßen „aufgeblasen“ werden. Auf diese Weise entstehen zusätzlich Wissenskarten höherer Detailliertheit: Die Überblickskarten differenzieren die einzelnen Regionen anhand der wichtigsten Unterscheidungsmerkmale, die für ein gesamtes Objektfeld gelten, während detaillierte Karten auf genau diese Information weitgehend verzichten, dafür aber die innerhalb eines Teiles des Feldes geltenden Differenzierungskriterien herausarbeiten.

Für die Strukturierung von Wissen sind in den letzten Jahren verschiedenste Verfahren entstanden, welche den geäusserten Gedanken Rechnung tragen. Gerade im Zusammenhang mit einer wachsenden Informationsmenge via elektronische Medien erhofft man sich von der Visualisierung der Datenflut in einem geeigneten Format eine maßgebliche Verbesserung der Benutzerfreundlichkeit im Dokumenten- oder generell im Informationsretrieval. Ein eindrückliches Beispiel solcher Informationsvisualisierung kann mit DocMiner³² gegeben wer-

³² DocMiner ist ein Kooperationsprojekt der folgenden Institute: Fraunhofer Institut für Medienkommunikation, Abteilung für künstliche Intelligenz Universität Dortmund & Fraunhoferinstitut für Arbeitswirtschaft und Organisation. Kurze Projektbeschreibung bei Becks & Host (2000a), www.doculine.com/news/2000/0700/VisuellRetrieval.htm. Weitere vergleichbare Ansätze zum Dokumenten Retrieval, welche auf die Methapher von Dokumentenlandkarten zurückgreifen, sind in Becks (2001, S. 51) ausführlich beschrieben.

den. DocMiner arbeitet mit der Metapher von „Dokumentenlandkarten“. Technisch integriert DocMiner dabei „statistisch-linguistische“ Verfahren³³, leistungsfähige Dimensionsreduktionsverfahren und eine Visualisierungstechnologie auf der Basis von neuronalen Netzen. DocMiner ermöglicht es dem Nutzer, eine Dokumentendatenbank in seinem ganzen Umfang zu explorieren, ohne dabei den Überblick zu verlieren. Die einzelnen Dokumente werden dem Nutzer als Punkte in einer zweidimensionalen Karte präsentiert. Dokumente ähnlichen semantischen Inhalts liegen nahe beieinander, unterschiedliche Dokumente weit auseinander. Die Struktur repräsentiert ein Klassifikationssystem semantischer Ähnlichkeiten. Beim normalen Retrieval besteht das Problem, dass der Nutzer auch bei einer erfolgreichen Anfrage immer nur einen winzigen Ausschnitt präsentiert bekommt und nie weiß, inwieweit die Gesamtmenge an Dokumenten seinem Bedürfnis entspricht. Anstatt dem Benutzer nur auf Anfrage Information in kleinen Häppchen zu geben, wird bei DocMiner als Ausgangspunkt deshalb die graphische Darstellung inhaltlicher Zusammenhänge direkt präsentiert. Er kann nun die Daten zunächst als Ganzes aus verschiedenen Blickwinkeln betrachten und dann durch Selektion von Teilbereichen auf bestimmte Ausschnitte der Daten fokussieren. Auf diese Weise ist eine explorative Analyse der Daten möglich.

6.2. Adaptiver Unterricht mit Intelligenten Tutoriellen Systemen

Verfahren dieser Art sollten sich nicht nur zum Finden von Information eignen, sondern auch dem Lernen eines Wissensgebietes förderlich sein. Dazu bietet es sich an, die hierarchische Ordnung der Karten adaptiv, d.h. bezogen auf das jeweilige Wissen des Lernenden, zu nutzen. Insbesondere könnte ein Lernverfahren nicht nur dem Lernenden Feedback über seine Leistungen in einer gerade erledigten Aufgabe geben, sondern es könnte gleichzeitig auch den weiteren Verlauf so steuern, dass sich der Lernende bei jedem Objekt stets auf der seinem Wissen angemessenen Detailliertheitsebene befindet.

Damit ist die Kernidee von adaptiven Lehrmaschinen formuliert, so wie diese im lerntheoretisch fokussierten Bereich der Psychologie des 20. Jahrhunderts konzeptionell und teilweise auch real entwickelt wurden. Leutner (1992a) stellt in seinem Buch „Adaptive Lehrsysteme“ einen historischen Abriss über die Entwicklung von Unterrichtsformen dar, welche, angefangen mit ersten Ansätzen in den zwanziger Jahren des letzten Jahrhunderts mit Maschinen zur „Intelligenzprüfung und Informationsvermittlung“ (vgl. Pressey, 1926) und deren Weiterentwicklung durch Skinner (1958) und Crowder (1959) in den sechziger Jahren in den ersten computerisierten Systemen mündeten (CAI: computer assisted instruction). Die Feedback-Systeme funktionierten in dem durch Skinner geprägten Ansatz ähnlich dem oben skizzierten Vorschlag, die Auswahl von Schwierigkeit der folgenden Lernitems dem gerade eben Erreichten individuell anzupassen.

Mit dem Aufkommen der Künstlichen Intelligenz wurden an die CAIs neue Anforderungen gestellt. Die nun als Intelligente Tutorielle Systeme (ITS)³⁴ bezeichneten Programme, sollen

³³ Zur autonomen, maschinellen Erstellung von Wissenskarten aufgrund von semantischen Beziehung zwischen Text-Dokumenten kann hier auf die Methoden des Dokumenten-Clusterings der semantischen Netzwerke von Lin et. al. (1991), Kohonen et al. (2000), Becks et. al. (2000b), Becks (2001) verwiesen werden.

³⁴ Auf definitorische Unterscheidungen zwischen den unzähligen und zum Teil gleichbedeutenden Bezeichnungen für (virtuelle) Lehrsysteme wie bspw. ITS (Intelligente Tutorielle Systeme), ATS (Adaptive Tutorielle Systeme) CBT (Computer Based Training) usw. wird hier nicht weiter eingegangen. In der vorliegenden Arbeit wird der etwas allgemeinere und geläufigere Terminus ITS verwendet und meint allgemein, dass sich das Lehrsystem an die individuellen Bedürfnisse des Lernenden anpassen kann. Im Speziellen bedeutet dies für diese Arbeit, dass das System auf individuelle Wissensdefizite entsprechend didaktisch und inhaltlich reagiert.

sich dem Instruktionsverhalten einer menschlichen Lehrkraft angleichen (vgl. Goldstein, 1982), sie sollen also über Lehrstoffexpertise verfügen. Erst der Vergleich zwischen dem IST-Zustand des Lerners und dem SOLL-Zustand des Expertenmodells führt zu „intelligentem Verhalten“ des Systems und bildet die Basis für einen ökonomischen und individuell angepassten Lernvorgang (vgl. zu den Komponenten eines ITS: Van Merriënboer, Van Den Berg & Maaswinkel, 1989).

Normativ betrachtet muss ein derartiger „adaptiver Unterricht“ zwei Zielen genügen: Zum einen der Fokussierung auf die Gestaltung didaktisch optimal aufbereiteter multimedialer Unterrichtsinhalte, zum anderen der überaus wichtigen, verlässlichen und effizienten Statusdiagnose als Grundlage für die Erstellung der kommenden Lektion. Eine lernprozessbegleitende Diagnose des Wissens führt jedoch bei diesen zwei Zielen zu einem Dilemma („Testlängendilemma“), auf das bereits Leutner (1992b, 1993) hingewiesen hat: Auf der einen Seite soll eine möglichst umfassende und genaue Diagnose erstellt werden, um bei der Zusammenstellung der folgenden Lektionen so adaptiv wie möglich sein zu können. Auf der anderen Seite will man aber nicht zuviel Zeit und kognitive Ressourcen der Lernenden für die Diagnose beanspruchen, weil beides von der verbleibenden Lernzeit abgeht. Die Adaptivität im Inhalt gewinnt ihre Relevanz dadurch, dass zum einen Mehrfachlernen weitgehend vermieden und zum anderen, dass falsches Wissen (welches über einen selbstgesteuerten Lernprozess nur schwerlich detektiert wird, vgl. Charney, Reder & Kusbit, 1990 und Strohschneider, 1990) ausgemerzt werden kann.

Klein (2000) resümiert in diesem Zusammenhang die Schwierigkeiten einer solchen Adaptivität. Er führt aus, dass die Adaptierbarkeit (d.h. die Anpassbarkeit des Systems durch den Nutzer, wie z.B. die Wahl unterschiedlicher Bildschirmarrangements oder das An- und Ausschalten von Hilfefunktionen) technisch wie konzeptionell wesentlich leichter zu lösen ist als die inhaltliche Adaptivität des Lernmaterials (vgl. auch Leutner, 1992a, 1995). In der Tat beschränken sich die heutigen ITS-Systemen weitgehend auf die Vermittlung regelbasierten Wissens, Wissens also, welches sich mit Wenn-dann-Regeln charakterisieren lässt (als Umsetzungs-Beispiele für Mathematik- und Physikwissen s. Albert & Lukas, 1999 und Siemer & Angelides, 1998; für einen erweiterten Überblick über bestehende ITS-Systeme vgl. Daub, 2001) und vielfach durch einen kleinen Satz an Entscheidungs- und Transformationsregeln abgedeckt werden kann.

Das Effizienzproblem in der Wissensüberprüfung wird jedoch besonders deutlich bei der Diagnostik merkmalsbezogenen Sachwissens. Geht es um die unzähligen, für die Qualität von Expertenentscheidungen aber zwingend notwendigen Fakten, müsste für eine zuverlässige Diagnose der gesamte Wissensstand innerhalb eines Fachgebietes mit Hunderten von Fragen abgetestet werden. Nur so wäre man in der Lage, Lücken und Falschwissen einer Person zu detektieren, was in der Praxis ein nicht durchführbares Unterfangen darstellt. Im Folgenden wird mit DIVE IN ein Verfahren vorgestellt, welches feedbackbasiert zu den Skinner'schen Wurzeln adaptiver Lernsysteme zurückgeht, diese jedoch auf hierarchische Wissensstrukturkarten anwendet.

6.3. DIVE IN – adaptiv und hierarchisch

Zur Realisierung einer „feedbackgestützten Lernmaschine“ findet die Wissensdiagnostik in DIVE IN nicht in einem von den Lerneinheiten abgekoppelten Modul statt. Das Diagnostiktool ist vielmehr auch gleich Wissensvermittlung. Auf diese Art und Weise lässt sich das „Testlängendilemma“ umgehen, denn wo Diagnostik = Lernen, ist der Zeitverlust für die Diagnose auf Kosten des Lernens gleich Null. Die Wissensdiagnostik geschieht in DIVE IN über ein einfaches Protokollieren oder Beobachten individueller Explorationstätigkeit beim Lö-

sen von Aufgaben. Es nutzt dabei das Zusammenspiel von Hypothesengenerieren und Hypothesenüberprüfen, welches für exploratives Vorgehen typisch ist.

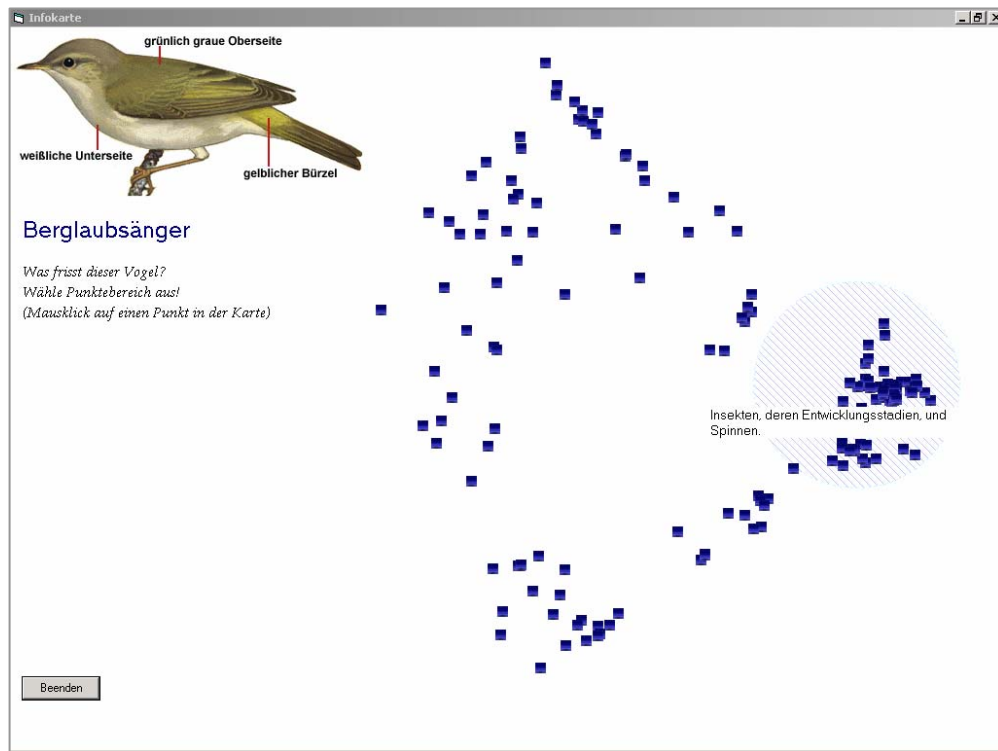


Abbildung 41: Beispiel **Level 1** für die adaptive Wissensvermittlung von Nahrungsmerkmalen.

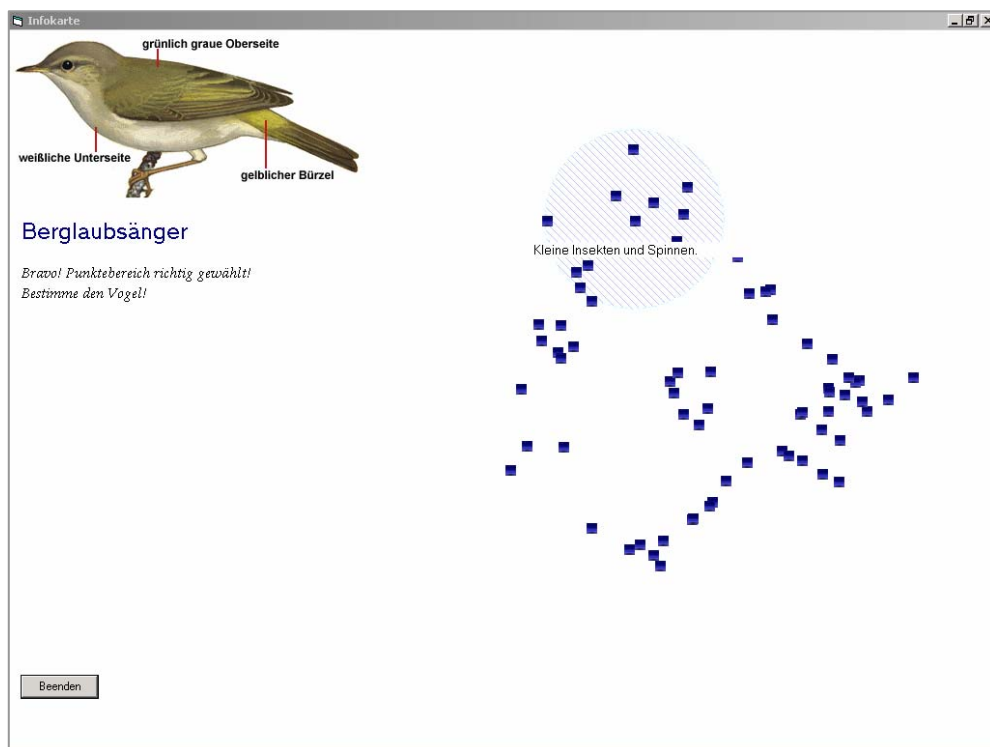


Abbildung 42: Beispiel **Level 2**, DiveInAuswahl von „Insektenfressern“. **Level 3** ist hier die direkte Bestimmung des Positionspunktes im blauen Kreis.

Abbildung 41 zeigt die Bildschirmoberfläche der technischen Realisierung von DIVE IN. Der Lerner hier etwas über die Nahrung einzelner Vogelarten lernen. Dazu wird ihm links auf dem Bildschirm eine Vogelart vorgelegt. Anhand der Abbildung des Vogels, dem Artnamen und der Familienzugehörigkeit soll der Lerner seine eigenen Nahrungshypothesen aufstellen und über die Wissensstrukturkarte überprüfen. Jeder einzelne Punkt in der Karte repräsentiert dabei eine Vogelart. Vögel mit ähnlicher Ernährungsweise liegen nahe beieinander, Vögel mit unterschiedlichen Ernährungsgewohnheiten sind weit voneinander entfernt. Mit „Mouse-over“ über einen Punkt in der Karte lassen sich die Fressgewohnheiten der einzelnen Vogelarten absキャンen. Dabei bleiben in dieser Version von DIVE IN die anvisierten Vogelarten inkognito. Kognitionspsychologisch gesehen sind hier zwei Prozesse für den Lernerfolg verantwortlich: Hypothesengenerierung und anschließende Überprüfung. Der Lerner sucht sich aktiv einen Bereich in der Karte aus, von dem er denkt, dass der vorgegebene Vogel zu den ihm angezeigten Merkmalsbeschreibungen passt. In dieser Version von DIVE IN geschieht dies mit dem Setzen eines blauen Kreises.

In Abbildung 41 ist der Berglaubsänger das zu bestimmende Objekt. Die hier präsentierte Wissensstrukturkarte zeigt Level 1 des Wissenssystems. Im gezeigten Beispiel wählt der Lerner einen Bereich der Karte aus, in der sich die Vögel vor allem von Insekten und Spinnen, zum Teil auch von Beeren oder Würmern ernähren. Zur Physiognomie des Vogels, feingliedrig und feiner Schnabel, scheint das Merkmal Insekten sehr gut zu passen. DIVE IN bestätigt seine Vermutung. Der ausgewählte Bereich wird in DIVE IN nun als detailliertere Karte aufgezeichnet (Abbildung 42). Die Objektsuche wird nun auf diesem Level 2 fortgesetzt. Ist auch das zweite Level geschafft (Auswahl eines Bereiches mit blauem Kreis), so folgt nun die direkte Auswahl eines Punktes. Wird auch diese Aufgabe richtig gelöst, gilt der Vogel bezüglich des Merkmals Nahrung als gelernt.

Kann die Nahrung einer Vogelart nicht durch Einkreisen (Level 1 und 2) oder direktes Anklicken bestimmt werden, so leuchtet der korrekte Punkt in der Karte rot auf, und eine Beschreibung der Nahrung des Vogels wird links eingeblendet. Diese Information kann sich der Lerner nun anschauen. Er weiß dabei, dass die Vogelart im Laufe des Lernprozesses wieder auftauchen wird.

Durch die Beschäftigung mit einzelnen Arten hat der Lerner die Möglichkeit, ein Strukturwissen über die Nahrungszusammenhänge in der Vogelwelt aufzubauen, indem er nach und nach die Karte auf ihren verschiedenen Levels strukturell zu interpretieren lernt.

Die Vogelarten lassen sich übrigens alle auch auf Level 1 durch Anklicken direkt bestimmen. Ein unmittelbarer Treffer spricht für gutes Wissen über diese Art, so dass die weitere Beschäftigung damit vom System gleich abgebrochen wird. So wird der Prozess zeiteffizient gehalten, indem er sich nur auf die noch nicht gewussten Objekte konzentriert.

Der Lerndurchgang beginnt mit einer Zusammenstellung von 12 Objekten. In einem zweiten Durchgang werden die Vogelarten in der Reihenfolge, die Bestgewussten zuerst, die Schlechtesten am Schluss durchexerziert. Einmal perfekt bestimmte Objekte fallen ab dem nächsten Durchgang heraus und werden durch neue Vogelarten ersetzt. Auf diese Weise beherrscht der Lerner zum Schluss alle 162 Vogelarten, welche sich derzeit in diesem System befinden, und hat sie jeweils nur so lange bearbeitet, bis er sie perfekt lokalisieren konnte. (Wenn ein Lerner allerdings eine Vogelart einmal auf Level 1 falsch bestimmt hat, muss er ab dem nächsten Durchgang den Treffer im Level 2 landen. Schliesslich wurde ihm die korrekte Position des Vogels ja bereits nach dem ersten Bestimmungsversuch offen gelegt und könnte einfach ohne Sachverstand wiedererinnert werden. Auf Level 2 darf der Vogel durchaus wiederholt direkt zu bestimmen versucht werden, denn alle Karten auf dem Level 2 sehen immer wieder anders aus und das "gewusst" kann demnach nicht auf blosses Wiederauffinden von Positionspunk-

ten zurückgeführt werden. Bei einem Treffer auf Level 2 gilt eine Vogelart deswegen in jedem Fall als gewusst und wird nicht weiter wiederholt.)

Die Wissensstrukturkarte basiert auf den in Kapitel 2 erhobenen Expertendaten. Dabei ist die 162er-Karte eine Neuskalierung der reduzierten Ähnlichkeitsmatrix mittels NMDS-Algorithmus RobuScal. Einige der ursprünglich 187 Vogelarten waren aus Gründen offensichtlicher Falschplatzierung durch die Experten eliminiert worden. Die Datenreduktion auf 162 Vogelarten ist allerdings auch eine Anpassung an das für die Wissensvermittlung eingesetzte Vogelbestimmungsbuch „Die Vögel Mitteleuropas“ (Singer, 2000), das bezüglich einzelner Merkmale und Vogelarten Lücken aufweist.

DIVE IN steht also für die Idee, den Lerner von einem Überblickswissen anfangen auf immer differenziertere Ebenen der Merkmalsunterscheidung eintauchen zu lassen. Dazu werden, ähnlich wie in DocMiner, einzelne Teilbereiche der Karte zur genaueren Betrachtung herangeholt. Das Heranholen von Teilstrukturen ist in DIVE IN jedoch keine Strukturvergrößerung im geometrischen Sinn. Wie einleitend erläutert gehen wir davon aus, dass detailliertere Wissenssysteme sich zwar zu Oberklassen vereinen, dass die Objektrelationen innerhalb einer Oberklasse im Vergleich zum Gesamtsystem jedoch nach eigenen Merkmalen strukturiert sind. Aus diesem Grund wird für Level 2 jeweils die Teilmenge im Umkreis um den angeklickten Vogel neu durch NMDS skaliert, indem auf die Teilmatrix der Proximitäten zwischen genau diesen Objekten zurückgegriffen wird. Dieses Verfahren der Neuskalierung garantiert, dass die Detailkarten wirklich auf denjenigen Merkmalen beruhen, die innerhalb genau dieses Bereiches der Karte differenzieren.

6.4. Fixe Klassen bestimmen

Die Auswahl von Objektbereichen für die Neuskalierung von Level 2 ist, wie gerade dargelegt, in der vorgestellten Version von DIVE IN flexibel; sie richtet sich nach den Objekten, die sich in dem auf Level 1 vom Lerner gesetzten Kreis befinden.

Eine Alternative würde darin bestehen, auf Level 1 die Objektklassen für Level 2 vorzubestimmen. Abbildung 43 zeigt ein solches Vorgehen am Beispiel einer Nahrungskarte mit der Objektziehung von 187 Vogelarten. Auf Level 1 (oben) lassen sich die drei Bereiche der NMDS-Karte mit Fisch-Fleischfresser (blau), Pflanzenfresser (grün) und Insektenfresser (orange) umschreiben. Klickt eine Person nun zur Identifikation der gesuchten Art auf einen der Punkte, so wird kein Radius um diesen Punkt für Level 2 berücksichtigt, sondern die entsprechende a-priori-Kategorie, der dieser Punkt semantisch zugeordnet ist.

Alle Objekte des entsprechenden Bereiches - im Beispiel von Abbildung 43 ist das der blaue mit den Fisch-Fleischfressern - werden nun als neu berechnete NMDS-Karte auf Level 2 präsentiert. Diese Karte besteht in diesem Beispiel wiederum aus drei semantisch gut voneinander abgrenzbaren Clustern (untere Karte). Eine inhaltliche Umschreibung der drei Klassen würde lauten: Fischfresser, Fleischfresser, würmerfressende Watvögel. Innerhalb jeder Klasse kann sich ein Lerner, bis hierher vorgedrungen, dann auf die Suche nach dem korrekten Punkt machen. Erreicht er zumindest das richtige Cluster, so hat er einen Teilerfolg erzielt und kann hier fortfahren, wenn 12 Objekte später dieselbe Vogelart wieder vom System aufgerufen wird.

Der Vorteil des Verfahrens Fixer Klassen ist dieser: Die Versuchspersonen arbeiten (in diesem Beispiel) nur noch mit 4 verschiedenen Strukturkarten, was das Lernen wahrscheinlich erleichtert. Außerdem lässt sich die Adaptivität des Systems erhöhen. Während der ornithologische Laie auf der obersten Ebene des Systems anfängt, könnte man fortgeschrittene Lerner nach einer bestimmten Anzahl auf erster Ebene richtig gelöster Aufgaben auch direkt auf unteren Levels explorieren lassen (und nur bei hoher Fehlerquote wieder auf die oberen Levels zurückwechseln).

Die fixe Zuteilung aller Objektpunkte zu a-priori-Kategorien kann in einer hoch klassifikatorischen Wissensstrukturkarte von Auge erfolgen. Bei Karten mit fließenden Übergängen zwischen den Clustern (wie der in Abbildung 43 oben) ist hingegen eine Clusteranalyse der Ähnlichkeitswerte zur Bestimmung semantisch beschreibbarer Cluster zu empfehlen.

Wo und wie der Schnitt bei einer Clusteranalyse gemacht werden soll, liegt im Ermessen des Modellbauers. Eventuell wäre der Cut auf einer anderen Ebene hinsichtlich der Wissensbasis oder hinsichtlich der Erlernbarkeit adäquater.

Nun kann es vorkommen, dass die Objektcluster aus der hierarchischen Clusteranalyse nicht perfekt der räumlichen Anordnung der Objekte in der NMDS-Abbildung entsprechen. Eine fixe Zuweisung der Unterkarten auf die Oberkarte führt dann dazu, dass diese sich räumlich an den Rändern überlappen. Das erschwert dann dem Lerner natürlich die Orientierung. Derartige Karten sollten deswegen vermieden werden. Wir schlagen in einem solchen Fall vor, die NMDS-Karte aufgrund der ultrametrischen Distanzen im resultierenden hierarchischen Clusterbaum zu berechnen (vgl. Kapitel 4). Auf diese Weise entsteht in jedem Fall eine kategoriale Struktur mit scharfen Rändern. Und da ein hierarchischer Clusterbaum auf jeder beliebigen Ebene ohne Veränderung der darunterliegenden Struktur geschnitten werden kann, erübrigt sich bei einem solchen Verfahren sogar das Neuskalieren der Karten auf Level 2, denn sie sind strukturidentisch mit dem Cluster in der Hauptkarte.

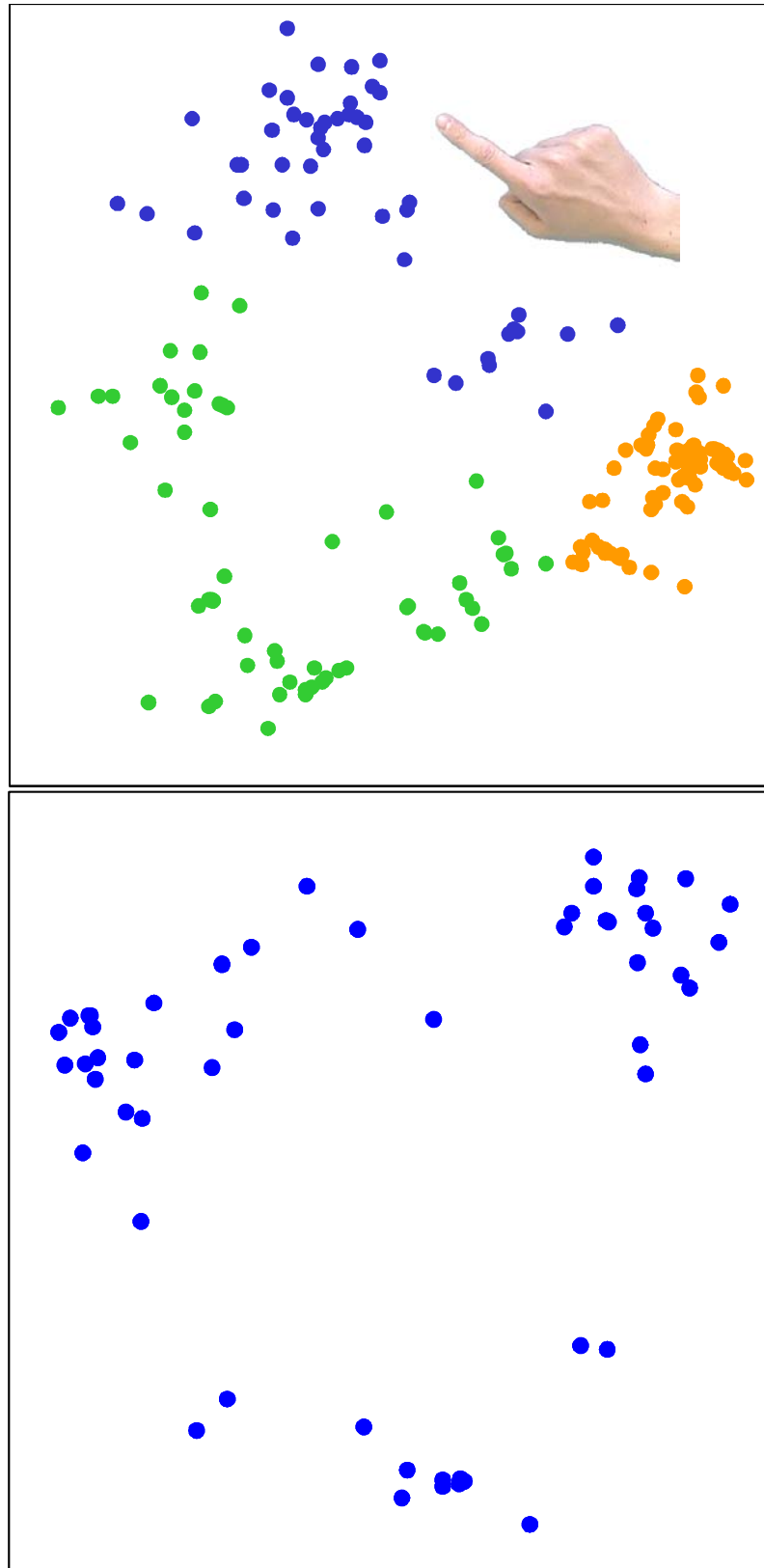


Abbildung 43: Vorgegebene Klassen auf Level 1 bestimmt mit Clusteranalyse (AvgVerfahren). Der Blaue Bereich bildet auf dem zweiten Level klare Objektklassen: Es sind dies hier um ihnen einen Namen zu geben die Fischfresser, die Fleischfresser und als drittes einige Würmer- und Insektenlarvenfressende Watvögel. Solch klare Klassen sind allerdings nur über Daten aus dem Parallelen Sortieren möglich, da Experten in dieser Erhebungsmethode Klassen bilden. In der Wirklichkeit sind solche Klassengrenzen nicht vorhanden.

6.5. Diskussion

Mit DIVE IN wird ein Verfahren vorgeschlagen, welches die hierarchisierten Realisierungen im kartenbezogenen Information Retrievals (z.B. DocMiner) mit dem Anspruch adaptiver Lernsysteme verknüpft. Die als Beispiel präsentierte Nahrungskarte sollte – was die Auswahl des Materials angeht – jedoch eher in ihrem Charakter als Demonstration einer Verfahrensidee gewürdigt werden denn als perfektes Lernmaterial. Sowohl auf der Darstellungs- als auch auf der Durchführungsebene gäbe es mehrere Möglichkeiten zur Verbesserung. Vor allem aber die aus Vogelbüchern ausgelesenen Merkmalstexte erwiesen sich im Nachhinein nicht als besonders geeignet, denn zum Teil weisen hier mehrere Vogelarten dieselben Texte auf. Solche Vogelarten wären durch einen einzigen Punkt in der Karte deshalb besser vertreten als jetzt. Dass sie gemäss der Expertendaten jedoch etwas unterschiedliche Koordinaten bekommen, muss als ein Hinweis darauf gesehen werden, dass die Texte nicht fein genug differenzieren. Doch geht es, wie gesagt, hier nicht um eine Kritik am vorliegenden Lernmaterial, sondern um die Diskussion der Verfahrensidee an sich.

Und diese Verfahrensidee beurteilen wir – vorbehaltlich einer Austestung hinsichtlich ihrer Effizienz als Lernmittel – positiv. Sie verbindet didaktische Forderungen nach aktivem, explorativem Lernen mit knallhartem Feedback hinsichtlich des gezeigten Verhaltens (und zwar durch entsprechende Auswahl des später noch einmal zu wiederholenden Stoffes). Zudem eröffnet sie den Weg zum einfachen Erlernen einer bereits von vornherein adäquaten Wissensstruktur, die im Laufe des Lernprozesses assimilatives (und damit leichteres) Aufnehmen neuer Fakten ermöglicht (konkret weiß der Lerner nach einer gewissen Zeit durch die reine Position in der Karte bereits viel über die Nahrung).

Diese lernpsychologischen Argumente lassen es angeraten erscheinen, den hier konzeptionell erarbeiteten und in einem Prototyp programmiertechnisch realisierten Ansatz nun in einem zweiten Schritt einer empirischen Überprüfung zu unterziehen. Dazu schlagen wir vor, das DIVE IN einer Beschäftigung mit einem Vogelbuch gegenüberzustellen und die Lernleistung im Merkmalswissen nach mehreren definierten Zeitpunkten (z.B. 1 Stunde, 2 Stunden, 3 Stunden etc.) Beschäftigung mit dem Material zu vergleichen. Unsere Prognose wäre diejenige, dass Adaptivität, aufgebautes adäquates Strukturwissen und vielleicht sogar auch der motivationale Vorteil medial gesteuerter Explorationstätigkeit dazu führen werden, dass der Vorteil an Lerneffizienz umso größer wird, je länger das Lernen mit DIVE IN andauert.

6.6. Literatur

- Alber, D. & Lukas, J. (Eds.), (1999). *Knowledge spaces. Theories, empirical research and applications*. Mahwah: Erlbaum.
- Becks, A. & Host, M. (2000a). *Visuell gestütztes Retrieval mit Dokumentenlandkarten*. www.doculine.com/news/2000/0700/VisuellRetrieval.htm. (zuletzt besucht Juni 2004)
- Becks, A., Sklorz, S., Jarke, M.: A (2000b). *Modular Approach for Exploring the Semantic Structure of Technical Document Collections*, ACM Proc. Of AVI 2000, Palermo.
- Becks, A. (2001). *Visual Knowledge Management with Adaptable Maps*. Dissertationsschrift an der Technischen Hochschule Aachen.
- Charney, D., Reder, L. & Kusbit, G.W. (1990). Goal setting and procedure selection in acquiring computer skills. A comparison of tutorials, problem solving and learner exploration. *Cognition and Instruction*, 7, 323-342.
- Crowder, N.A. (1959). Automating tutoring by means of intrinsic programming. In E. Ganter (Eds.), *Automatic teaching. The state of the art* (pp. 109-116). New York: John Wiley and Sons.

- Daub, S. (2001). *Similarity Simulation – Wie man den Code des Globalurteils knackt*. Dissertationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Goldstein, I.P. (1982). The genetic graph: A representation for the evolution of procedural knowledge. In D. Sleeman & J.S. Brown (Eds.), *Intelligent tutorial systems* (pp. 51-78). New York: Academic Press.
- Klein, B. (2000). *Didaktisches Design hypermedialer Lernumgebungen. Die adaptive Lernumgebung „incops“ zur Einführung in die Kognitionspsychologie*. Marburg: Tectum.
- Kohonen, T., Kaski, S. et al. (2000). Self Organization of a Massive Document Collection. *IEEE Transactions on Neural Networks*, Vol. 11, No. 3.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophische Fakultät der Universität Zürich.
- Leutner, D. (1992a). *Adaptive Lehrsysteme. Instruktionspsychologische Grundlagen und experimentelle Analysen*. Weinheim: Psychologie Verlags Union.
- Leutner, D. (1992b). *Das Testlängendilemma in der lernprozess-begleitenden Wissensdiagnostik*. *Zeitschrift für Pädagogische Psychologie*, 24, 232-248.
- Leutner, D. (1993). Das gleitende Testfenster als Lösung des Testlängendilemmas: Eine Robustheitsstudie. *Zeitschrift für Psychologie*, 7, 33-45.
- Leutner, D. (1995). Adaptivität und Adaptierbarkeit multimedialer Lehr- und Informationssysteme. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (S. 140-149). Weinheim: Psychologie Verlags Union.
- Lin, X., Soergel, D., Marchionini, G., (1991): A self-organizing semantic map for information retrieval. In Proc. Of 14th ACM/SIGIR conf. *Research and Development in Information Retrieval*.
- Pressey, S.L. (1926). A simple apparatus which gives tests and scores and teaches. In (reprint) A. A. Lumsdaine & R. Glaser (Eds.), (1969). *Teaching machines and programmed learning* (pp. 35-41). Washington D.C.: National Education Association of the United States.
- Siemer, J. & Angelides, M.-C. (1998). Towards an intelligent tutoring system architecture that supports remedial tutoring. *Artificial Intelligence Review*, 12, 469-511.
- Singer, D. (2000). *Die Vögel Mitteleuropas*. Stuttgart: Franckh-Kosmos.
- Skinner, B.F. (1958). Teaching machines. In (reprint) A. A. Lumsdaine & R. Glaser (Eds.), (1960). *Teaching machines and programmed learning* (pp. 137-158). Washington D.C.: National Education Association of the United States.
- Strohschneider, S. (1990). *Wissenserwerb und Handlungsregulation*. Wiesbaden: Deutscher Universitätsverlag.
- Van Merriëboer, J.J.G., Van Den Berg, K.G. & Maaswinkel, D.M. (1989). Some experiences with two intelligent tutoring systems for teaching computer programming: PROUST and LISP-tutor. In J. M. Pieters (Ed.), *Intelligent tutorial systems and instruction* (pp. 7-44). Enschede: OTG Onderwijsleerprocessen.

7.

**Die Veränderung von Wissenskarten beim
Merkmalslernen**

Zusammenfassung

Merkmalswissen über eine Anzahl von Objekte lässt sich relational in Form einer Wissenskarte abbilden. Mit fortschreitendem Lernen sollten sich solche individuellen Karten systematisch verbessern. Dazu wird folgendes Messverfahren vorgeschlagen: Eine Matrix paarweiser Ähnlichkeitsurteile zwischen einem Satz von Objekten wird mittels Nonmetrischer Multidimensionaler Skalierung in eine zweidimensionale Wissenskarte überführt. Diese Karte wird durch Prokrustes-Transformation auf eine Zielstruktur gelegt. Die numerische Abweichung der beiden Karten (der AvgLoss) dient dann als kritisches Maß für die Güte des Wissens.

Zwei Lernexperimente zeigen den angenommenen Zusammenhang zwischen Merkmalswissen und Güte der Wissenskarte auf. Experiment 1 überprüft das vorgeschlagene Verfahren mittels eines komplexen Wissensgebietes (der Ornithologie) und belegt die systematische Verbesserung der Karten. Experiment 2 dokumentiert dann an einfacherem Lernmaterial (Länder) die Konvergenz der Wissenskarten untereinander und gibt einen ersten Hinweis darauf, wie ähnlich individuelle Wissenskarten einer Zielstruktur werden können.

Schlüsselwörter: Wissenspsychologie, Sachwissen, Merkmalswissen, Wissenskarte, Kognitive Struktur, Kognitive Karte, Lernexperiment, Ähnlichkeitsurteil, Nonmetrische Multidimensionale Skalierung, Prokrustes-Transformation

7.1. Einleitung

Elemente unseres Wissens existieren nicht einfach unverbunden nebeneinander, sondern sie bilden zusammen ein Ganzes. Diese alte Erkenntnis der Piaget'schen Tradition in der Pädagogischen Psychologie (Aebli 1976; Ausubel, 2000; Inhelder, 1974; Inhelder, Sinclair & Bovet, 1974; Montada, 1995; Piaget, 1973, 1976, 2001; Piaget & Inhelder, De Caprona & Cornu-Wells, 1987; Steiner, 1996) hat die Wissenspsychologie seit ihrer Wiederbegründung als eigenständiger Forschungszweig (vgl. Mandl & Spada, 1988) in ihren neueren Entwicklungen aufgenommen. Die modernen Modelle der Wissensmodellierung zeichnen sich durch eine Art System aus, d.h. die Elemente konstituieren unwillkürlich einen Satz von Relationen. Z.B. unter Rückgriff auf Konzeptionen eines Semantischen Netzwerkes (wie das von Quillian, 1968) wurde diese relationale Facette von Sachwissen aufgegriffen und mit Mapping-Techniken abgebildet. Damit tragen diese Modelle dem Umstand Rechnung, dass sich Wissensstrukturen durch Differenziertheit, Vollständigkeit und Komplexität auszeichnen, also weit über eine einfache Anhäufung unverbundener Fakten hinausgehen (vgl. Mandl & Fischer 2000). Dabei bedienen sich die Mapping-Techniken vorwiegend graphischen Verfahren in der Datenerhebung. Exemplarisch zu nennen sind die zahlreichen Arbeiten mit der Heidelberger Strukturlegetechnik (basierend auf Scheele & Groeben, 1984) und die verwandten Verfahren der graphentheoretischen Darstellung von Wissensstrukturen (Bonato, 1990), die Netzwerk-Elaborierungs-Technik (Eckert, 1998, 1999, 2000), das Concept Mapping Software Tool (COMASOTO) zur Diagnose strukturellen Wissens (Weber & Schumann, 2000), die Wissensvermittlung, tutoriell unterstützte Wissensanwendung und Wissensdiagnose mit Begriffsnetzen nach Plötzner, Beller & Härder (2000) oder die Knowledge-Tracking-Methode von Janetzko & Strube (2000).

Mandl & Fischer (2000) resümieren, dass nach anfänglichen Problemen bei der Wissensmodellierung (hoher Idiosynkrasie-Anteil, fragliche Validität und mangelnde Re-Test-Reliabilität) inzwischen theoretische Grundlagen und auch konkret eingesetzte Diagnoseverfahren und Benutzerinterfaces vorliegen. Sie verweisen in ihrem historischen Abriss auch auf

die psychometrisch-strukturelle Perspektive, die „... mit verschiedenen Formen der Ähnlichkeitsskalierungen und ihrer graphischen Darstellung (z.B. Multidimensionale Skalierung, Clusteranalysen) ... [den Versuch unternehmen] ... psychische Phänomene und deren Veränderung mathematisch zu quantifizieren.“ (S. 4). Über die vorhandenen graphischen Mapping-Techniken hinaus sprechen die Autoren auch explizit die Möglichkeit an, auf der Basis nicht-graphischer Daten (z.B. multiple Paarvergleiche) über komplexe Algorithmen auf Kognitive Wissensstrukturen zu schließen.

Im Folgenden wird ein Weg zur Diagnose merkmalsbezogenen Sachwissens vorgestellt, der genau einer solchen Grundidee folgt: Über numerische relationale Urteile soll für einen Objektbereich festgestellt werden, wie eine individuelle Person dieses Feld für sich selbst strukturiert hat. Ein geeignetes, weil verhältnismäßig expertiseunabhängiges Messinstrument dafür sind paarweise Ähnlichkeitsurteile (vgl. Klauer, 1989). Bereits mit wenigen Informationen sind wir Menschen nämlich in der Lage, ein ungefähres Urteil über die Ähnlichkeit zweier Dinge oder Personen abzugeben. Dabei vertrauen wir offenbar auf unser „gutes Auge“, mit dem wir wichtige von unwichtigen Merkmalen unterscheiden und die verbleibenden wichtigen Merkmale in eine wohlproportionierte Relation zueinander bringen. Laien sind genauso in der Lage wie Experten, die Ähnlichkeit zwischen zwei Dingen abzuschätzen – auch wenn das Urteil natürlich beim Experten qualitativ besser, d.h. sachadäquater im Vergleich zu einer Norm ausfallen sollte. Läge (2004) konnte zeigen, dass sich Ähnlichkeitsurteile mit zunehmendem Expertisegewinn systematisch verändern.

Nach unserem Dafürhalten teilen Ähnlichkeitsurteile mit den bestehenden Mapping-Techniken gegenüber einer reinen Abfrageliste von Sachfragen den Vorteil, dass sie relationale Gefüge, Integration und Elaboration (Gewichtungen, kategoriale Strukturierung usw.) der Wissensstruktur mit messen. Zusätzlich sind sie unabhängig von Frageformaten und relationalen Vorschriften. Allerdings liegt die resultierende relationale Information bei Ähnlichkeitsurteilen genauso wenig unmittelbar auf der Hand wie das zu ermittelnde Merkmalswissen über die einzelnen Objekte (welches mit Mapping-Techniken nur umständlich zu erheben ist). Es bedarf vielmehr komplexer Auswertealgorithmen, um relationales Wissen und Merkmalswissen aufzuzeigen.

Durch Ähnlichkeitsurteile ermittelte Proximitäten zwischen Paaren von Objekten lassen sich mittels Nonmetrischer Multidimensionaler Skalierung (NMDS) als Distanzwerte verrechnen und zur Grundlage von „Wissenskarten“ machen (die Logik dieses Verfahrens ist anwendungsorientiert bei Borg & Groenen, 1997 und mathematisch grundlegender bei Mathar, 1997 beschrieben). Auf diese Weise gewinnt man die relationale Information über das aktuelle Sachwissen. Solche individuellen „Landkarten“ können durch Prokrustes-Transformation mit einem Expertenmodell verglichen werden. Bereits ohne großen mathematischen Aufwand springt bei einem solchen Vergleich ins Auge, welche Objekte ein Lerner gut kennt (d.h. korrekt verortet) und welche er falsch platziert (als Beispiel vgl. Abbildung 44).

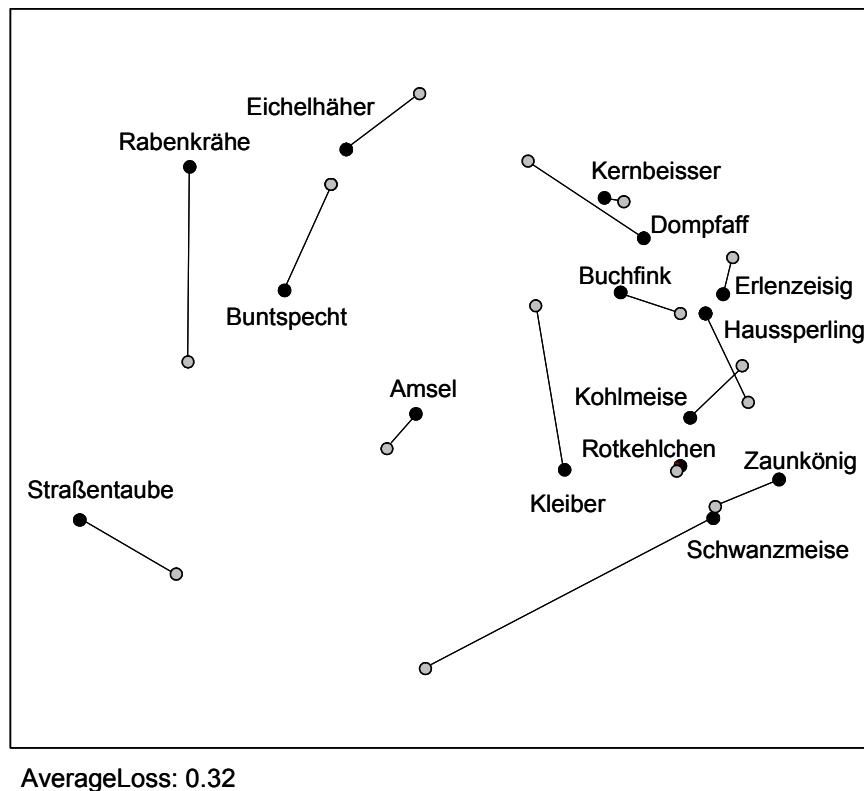


Abbildung 44: Prokrustes-Transformation eines Expertenmodells mit einer Lernerkarte. Die schwarzen Punkte stehen für die Positionen der Vogelarten im Expertenmodell, die grauen für diejenigen in der Lernerkarte (vgl. Experiment 1).

Als Gesamtabweichungswert einer Prokrustes-Transformation hat Läge (2001) den Average-Loss (kurz: AvgLoss) eingeführt. Dieses Pauschalmaß könnte also als Schätzung für die Güte des aktuellen Wissensstandes einer Person herangezogen werden. Im Laufe des Expertisegewinns sollte sich dieser Abweichungswert von einer Wissensnorm systematisch verringern.

Ein solcher Vorschlag, um die Güte individuellen Wissens in einer vernetzten Form zu berechnen, ist unseres Wissens bislang noch nicht gemacht worden. Zwar gibt es eine Reihe von Anwendungen Nonmetrischer Multidimensionaler Skalierung in der Psychologie (etwa Gigerenzer, 1981; Marx & Hejj, 1989; Marx & Läge, 1995; Edelman, 1998; Lee, Byatt & Rhodes, 2000), und auch Prokrustes-Transformationen wurden vereinzelt bereits durchgeführt (vgl. Gower & Dijksterhuis, 2004), aber geschah dieses doch durchgängig für gemittelte Proximitätsmatrizen befragter Personenstichproben. Technisch betrachtet hatte das seinen guten Grund, denn die bislang bestehenden Verfahren der Multidimensionalen Skalierung waren sehr anfällig für Ausreißerwerte in einer Proximitätsmatrix. Solche Ausreißer (in Form extremer Inkonsistenzen hinsichtlich des Gesamt der Messwerte) passieren immer einmal wieder bei der Abgabe von Urteilen, sind jedoch a priori nicht zu detektieren. Jedoch können sie die resultierenden NMDS-Karten in der Gesamtstruktur massiv beeinflussen, indem alle Punkte in ihrer Struktur verändert werden, nur um einen einzelnen zu großen Einpassungsfehler abzufangen. Bei gemittelten Stichprobenwerten verschwinden solche unsystematischen Ausreißer, weswegen sie dann für die Skalierung kein Problem mehr darstellen. Erst mit der Entwicklung eines robusten Verfahrens der NMDS, welches Ausreißer im Prozess der schrittweisen Genese der geometrischen Konfiguration detektiert und eliminiert, lassen sich nun aber auch Wissenskarten auf individuellem Niveau verlässlich nachzeichnen (RobuScal-Algorithmus; Läge et al., in review). Analoges gilt für die Prokrustes-Transformation: Auch hier kann ein einzelnes schlecht verortetes Objekt das Übereinanderlegen der Strukturen massiv beeinflussen. Deswegen ist auch hier eine robuste Version angebracht (Läge, 2001).

Dank dieser robusten Verfahren ist es nun möglich, die individuellen Ähnlichkeitsurteile in Wissenskarten umzusetzen. Dabei ist es interessant, die Veränderung dieser Wissenskarten im Verlauf der Beschäftigung mit den Merkmalen von Objekten zu beobachten. Die Zunahme oder (im Falle inadäquaten Vorwissens) die Veränderung von Merkmalswissen sollte ja eine Auswirkung auf die relationalen Ähnlichkeitsurteile haben, die eine Person abgibt. Und die Matrix dieser veränderten Ähnlichkeitsurteile wiederum sollte eine Restrukturierung der Karte als ganzer oder zumindest einzelner Objekte innerhalb einer als ganzer stabil bleibenden Karte bewirken.

Dabei setzt diese Kette eine gewisse Integration des durch die Beschäftigung neu erworbenen Wissens voraus. Merkmale, die ohne jedes Verständnis einfach nur auswendig gelernt werden, mögen in reproduktiven Testfragen (wie z.B. den meisten Multiple-Choice-Aufgaben) abrufbar sein. In die Veränderung eines Ähnlichkeitsurteils sollten sie aber nur dann eingehen, wenn die tatsächliche Bedeutung für dieses Objekt und für das Gesamt des Wissensfeldes auch verstanden wurde.

Im Folgenden werden zwei systematische Lernstudien vorgestellt, die sich der Frage der Veränderung von Wissenskarten beim Merkmalslernen widmen. Experiment 1 bedient sich dazu des Wissensgebietes der Ornithologie, weil hier hinreichend klar definierte Objekte (Vogelarten) mit komplexen definierenden und charakterisierenden Merkmalssätzen vorliegen. Experiment 2 schließt im Aufbau an Experiment 1 an, versucht aber den Gedanken der Konvergenz der individuellen Wissenskarten noch stärker herauszuarbeiten. Dazu wird dem Lernexperiment der Inhaltsbereich der Geographie zugrunde gelegt, und zwar sind Länder anhand von drei klar definierten Merkmalen (Bruttosozialprodukt, Einwohnerzahl und Größe) auf ihre Ähnlichkeit hin einzuschätzen. Mit diesem – im Vergleich zu Experiment 1 leicht lern- und integrierbaren – Material kann der maximale Grad an Konvergenz individueller Wissenskarten ausgelotet werden. Am Ende dieser beiden Experimente soll dann ein Einblick in Ausmaß und Richtung der Veränderung einer Wissenskarte beim Lernen stehen.

7.2. Experiment 1

7.2.1. Fragestellung und Hypothesen

Experiment 1 misst den Fortschritt des Lernens in individuellen Wissenskarten, indem es den jeweiligen Kenntnisstand in einem konvergenten Wissensgebiet (der Ornithologie) mit einem Expertenmodell vergleicht. Dabei wird als erster Hauptbefund eine gerichtete und zudem konvergente Veränderung der individuellen Wissenskarten untereinander erwartet. Zweitens sollte diese Veränderung in eine systematische Annäherung der individuellen Wissenskarten an ein Expertenmodell resultieren (Im Idealfall sollte ein Lerner so dicht an die Normstruktur herankommen wie ein individueller Experte, was aber in einem Experiment dieser Art schwerlich zu erreichen sein wird).

7.2.2. Methode und Durchführung

Aufbau des Experiments. In einer Längsschnittuntersuchung hatten die Versuchspersonen die Aufgabe, 15 Vogelarten anhand vier charakteristischer Merkmalsbereiche zu lernen. In drei Sitzungen, die jeweils im Wochentakt stattfanden, erhielten sie dazu Informationen zu diesen 15 Objekten. Mittels paarweiser Ähnlichkeitsurteile wurde dann die Auswirkung der Lerneinheiten auf die individuellen Wissenskarten der Versuchspersonen überprüft.

Die Versuchspersonen wurden vor dem Experiment zufällig in drei Gruppen aufgeteilt: zwei Experimentalgruppen mit unterschiedlichem Treatment (Exkursionslerner und Kärtchenlerner) und eine Kontrollgruppe³⁵. Die Kärtchenlerner (KL) erhielten in allen drei Lernphasen Kärtchen der 15 Vogelarten mit merkmalsstrukturierten Informationen. Die Exkursionslerner (EL) nahmen im ersten Lerndurchgang mit einem Ornithologieexperten an einer Exkursion teil, an der sie alle 15 Arten beobachten konnten. Der Exkursionsleiter konzentrierte sich in seinen Erläuterungen vornehmlich auf die Merkmalsinformationen, die auch die Kärtchenlerner zur Verfügung hatten. Die beiden folgenden Lernphasen unterschieden sich nicht von denen der Kärtchenlerner. Die Kontrollgruppe durchlief nur die Wissensdiagnose, nahm jedoch nicht an den Lerneinheiten teil. Der Aufbau des Experiments ist in Abbildung 45 skizziert.

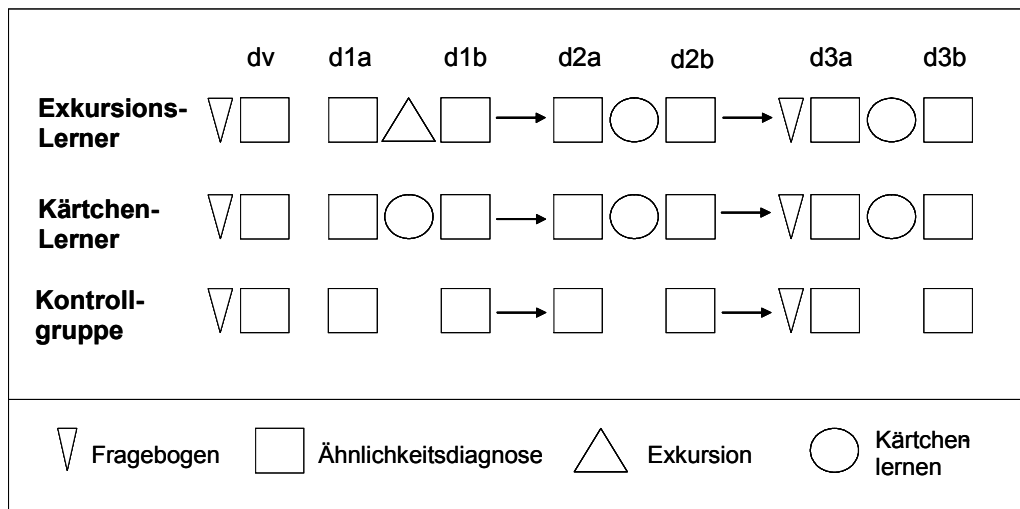


Abbildung 45: Schematische Darstellung des Ablaufs von Experiment 1

Vor der ersten und vor der dritten Lernphase durchliefen alle Versuchspersonen einen merkmalsorientierten Sachwissenstest, indem ein Multiple-Choice-Fragebogen zu allen 15 Vogelarten die vier Merkmalsbereiche aus den Lerneinheiten testete (Der zweite Test fand mit Absicht nicht unmittelbar nach einer Lerneinheit statt, um Kurzzeiteffekte zu vermeiden. Vor der dritten Lerneinheit lag eine Woche lernfreie Zeit, so dass nur der langfristige Erfolg aus den beiden ersten Durchgängen gemessen wurde.). Vor und nach jeder Lerneinheit wurde der aktuelle Wissensstand aller Versuchspersonen mittels paarweiser Ähnlichkeitsurteile auf der Strukturebene erhoben. Vor der ersten Lerneinheit gaben die Versuchspersonen den Satz von 105 Ähnlichkeitsurteilen zweimal ab, damit sie beim ersten Mal (dv) etwas Übung mit diesem Verfahren bekamen (Der zweite der beiden Datensätze (d1a) wird dann bei der Auswertung als Messung des Vorwissens angesehen). Schließlich wollten wir die Veränderung der Karten eindeutig auf das Lern-Treatment zurückführen und nicht auf einen Gewöhnungseffekt bei den Ähnlichkeitsurteilen. Die Kontrollgruppe durchlief dieselbe Wissensdiagnostik.

Auswahl der Vogelarten und Bestimmung der Zielstruktur. Die Auswahl der 15 Vogelarten, so wie sie sich im Expertenmodell (Abbildung 46) präsentiert, hatte einigen Kriterien zu folgen. (1) Für die Exkursion mussten alle Arten im Raum Zürich im selben Habitat leben und

³⁵ Ursprünglich sollten mit diesen beiden Lernbedingungen mögliche Unterschiede zwischen einem entdeckenden Lernen (Exkursion) und einem rein instruktionalen Lernen (Kärtchen) aufgedeckt werden. Da sich jedoch, wie die Ergebnisse zeigen werden, die beiden Gruppen nach zufälliger Zuordnung der Versuchspersonen in den Startbedingungen unterschieden, kann auf diese Teilfragestellung in diesem Bericht nicht weiter eingegangen werden.

dort leicht zu finden sein, und zwar auch im Winter, als dieses Experiment durchgeführt wurde. (2) Einzelne Arten sollten nicht dermaßen von allen anderen Vögeln verschieden sein, dass auswertungstechnisch betrachtet die Gefahr degenerierter NMDS-Lösungen bestanden hätte. Auch sollten die Arten nicht nur wenigen, scharf getrennten Verwandtschaftsgruppen angehören, um übermäßig kategoriale Lösungen zu vermeiden. Das war technisch wichtig, weil sich sonst ein möglicher Lernfortschritt nicht numerisch fair durch Prokrustes-Transformation von NMDS-Karten messen lassen würde. (3) Ca. sechs Arten sollten Laien (zumindest oberflächlich) bekannt sein, damit eine gewisse Grundstruktur vorhanden war. Mindestens sechs Arten sollten aber den Versuchspersonen völlig unbekannt sein, um ausreichenden Spielraum für das Lernen zu besitzen. Die ersten beiden Restriktionen führten zu einer eingegrenzten Liste möglicher Arten. Diese wurden in einem Vorversuch 30 Studierenden der Universität Zürich in einem Fragebogen vorgelegt, um Bekanntheitsgrad und Vorwissen einschätzen zu können. Daraus erfolgte dann die Auswahl der 15 am geeignetsten erscheinenden Arten.

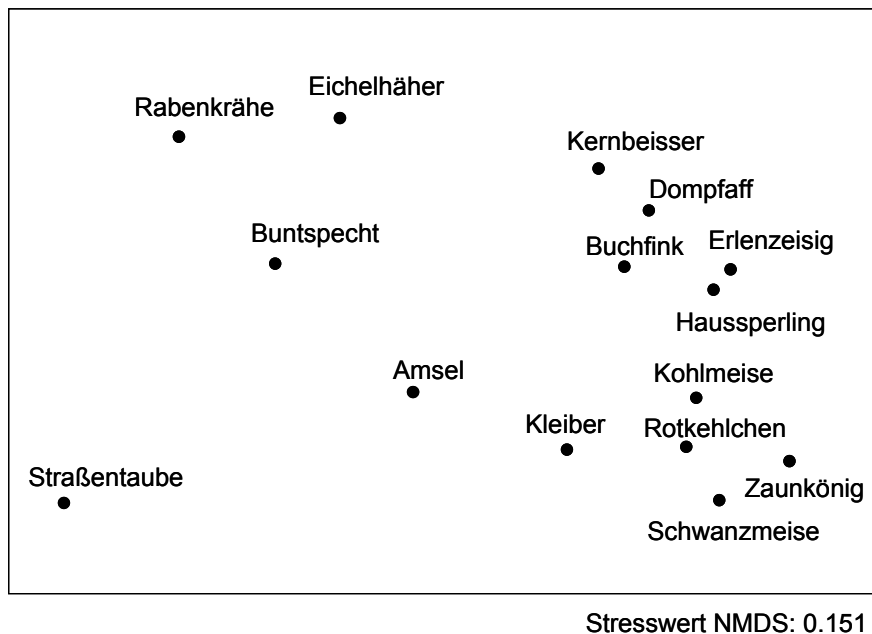


Abbildung 46: NMDS-Karte, die in Experiment 1 die Zielstruktur des Lernens ergibt.

Auf Strukturebene benötigen wir eine Wissensnorm, mit der jede Lernerkarte zu jedem Zeitpunkt verglichen werden kann. Da diese Karte für die numerische Bestimmung des Lernerfolgs sehr wichtig ist, wurde große Mühe darauf verwendet, sie so akkurat wie möglich zu erheben. Deshalb wurden die 15 ausgewählten Arten acht erfahrenen Ornithologie-Experten vorgelegt, die jeweils die paarweisen Ähnlichkeiten einzuschätzen hatten. Bei vier von diesen Experten konvergierten die resultierenden NMDS-Karten in sehr hohem Maße, so dass ihre gemittelten Urteile mit einem Gewicht von 50% zur Grundlage für die Normkarte herangezogen wurden. Um sicherzustellen, dass das Normmodell auch wirklich die zu lernenden Merkmalsbereiche abdeckte, wurden die individuellen Ähnlichkeitsurteile nach dem letzten Lernen für jede der beiden Treatment-Gruppen gemittelt. Jede dieser zwei Ähnlichkeitsmatrizen floss mit einem Gewicht von 25% in die durch NMDS zu bildende Normkarte ein (Abbildung 46).

Lernmaterialien. Die Exkursion fand in zwei Sechsergruppen statt und dauerte ca. eineinhalb Stunden. Die Lerner konnten alle 15 Arten unter Anleitung und begleitenden Kommentaren eines erfahrenen Exkursionsleiters per Fernglas beobachten. Zusätzlich zu seinen Auskünften

hatten die Exkursionsteilnehmer die Lernkärtchen in der Hand und konnten sich bei Bedarf daran orientieren. Diese Lernkarten setzten sich je Vogelart aus drei Teilen zusammen: aus a) einer farbigen Abbildung der Vogelart, aus b) einem kurzen, beschreibenden Text, der neben Besonderheiten und leicht eingänglichen Kuriositäten über die Art eine Merkmalsbeschreibung zu Lebensraum, Nahrung und Nistgewohnheiten enthielt (so wie man sie in Standard-Vogelbüchern findet), sowie aus c) einer Merkmalsliste, die alle wichtigen Merkmale aus dem beschreibenden Text tabellarisch aufführte.

In den individuell durchgeführten Lerndurchgängen erhielten die Versuchspersonen jeweils dieselben Lernmaterialien wieder vorgelegt. Sie waren in der Organisation ihres Lernens weitgehend frei, hatten sich aber in den beiden ersten Sitzungen jeweils 30 Minuten und in der dritten Sitzung 20 Minuten damit zu beschäftigen. Als Kriterien wurde ihnen an die Hand gegeben, dass sie versuchen sollten, aufgrund der Abbildungen die Arten zu benennen und dass sie in der verbleibenden Zeit alle Kärtchen mindestens einmal durchgelesen haben sollten. Sie wussten, dass sie Ähnlichkeitsurteile über diese Arten abzugeben hätten und konnten bereits beim Lernen die Kärtchen strukturiert auf einem Tisch ausbreiten.

Versuchspersonen. An der Hauptuntersuchung nahmen 31 Studierende der Universität Zürich im Alter von 22–27 Jahren teil. Sie besaßen keine spezifischen ornithologischen Vorkenntnisse. In den beiden Experimentalgruppen befanden sich jeweils 4 Frauen und 8 Männer, die Kontrollgruppe bestand aus 5 Frauen und 4 Männern. Die Gruppenzuteilung erfolgte vor dem Experiment und zufällig.

7.2.3. Resultate

Merkmalswissen wird gelernt. Vor der ersten und vor der dritten Lerneinheit durchliefen alle Versuchspersonen einen allgemeinen Wissenstest. Zu allen 15 Vogelarten wurde dabei das Merkmalswissen über vier komplexe, mehrere Einzelmerkmale enthaltende Fragen in Form eines Multiple-Choice-Tests erhoben. Maximal konnten in dem Test 60 Punkte erzielt werden. Die Gruppe der Exkursionslerner verbesserte sich von Test 1 zu Test 2 von durchschnittlich 34 auf 46.75 Punkte, die Kärtchenlerner von 30 auf 43.75 Punkte. Beide Verbesserungen sind im t-Test auf dem 0.1% -Niveau signifikant (EL: $t(11) = -10.15$, $p < 0.001$, zweiseitig; KL: $t(11) = -6.67$, $p < 0.001$, zweiseitig). Die Versuchspersonen der Kontrollgruppe hingegen verbesserten sich nicht signifikant (von 28 auf 30.77 Punkte; $t(8) = -2.03$, $p = 0.07$, zweiseitig). Im Test 2 sind die beiden Lerngruppen deutlich besser als die Kontrollgruppe (EL: $t(19) = -5.04$; $p < 0.001$; zweiseitig; KL: $t(19) = -3.92$; $p < 0.001$, zweiseitig), unterscheiden sich jedoch nicht signifikant untereinander ($t(22) = 1.28$; $p = 0.21$; zweiseitig). Die erzielten Punktzahlen zeigen deutlich an, dass die beiden Experimentalgruppen in den beiden ersten Lerndurchgängen tatsächlich gelernt hatten, nicht aber die Kontrollpersonen.

Die reine Ratequote in diesem Test lag bei 15 Punkten (60 Fragen mit vier Antwortmöglichkeiten, von denen jeweils eine richtig war). Der Wissenstest dokumentiert also, dass die Versuchspersonen bereits über ein Eingangswissen verfügten. Das war ja auch so geplant, denn es waren bewusst sechs Vogelarten ausgewählt worden, die im Vorversuch bereits bei Laien einen gewissen Bekanntheitsgrad aufwiesen. Gleichzeitig war – wie ebenfalls angezielt – noch nach oben hin genügend Spielraum für eine Verbesserung des Wissens. Die beiden Treatmentgruppen erreichten bis zu Beginn der dritten Lerneinheit etwa die Hälfte des möglichen Wissenszuwachses, so dass in der Längsschnittuntersuchung kein Deckeneffekt auftrat. Mit dieser Überprüfung sind die notwendigen Voraussetzungen hinsichtlich des Merkmalswissens bestätigt: Anders als die Kontrollpersonen verbessern die Lerner ihr Wissen. Damit können nun die Hauptfragestellungen beantwortet werden.

Konvergente Veränderung der Wissenskarten. Gemäß der ersten Hypothese ist zu überprüfen, ob die Wissenskarten der Lerner im Laufe des Prozesses systematisch konvergieren. Wir ver-

gleichen dazu die Wissensstruktur unmittelbar vor dem ersten Lernen (Zeitpunkt d1a) und unmittelbar nach dem dritten Lernen (d3b) miteinander. Dazu werden aus den Ähnlichkeitsurteilen jeder Person zu jedem der beiden ausgewählten Zeitpunkte zweidimensionale NMDS-Karten gerechnet. Diese Karten werden jeweils paarweise mit Prokrustes-Transformation übereinander gelegt. Die resultierende Matrix der AvgLosses lässt sich mit NMDS in einen Raum legen, so dass die Ähnlichkeitsstruktur zwischen den individuellen Wissenskarten sichtbar wird. Letzteres geschieht unter Einschluss der Zielstruktur (um die Richtung einer systematischen Veränderung gleich qualitativ einordnen zu können).

Abbildung 47 zeigt die resultierende NMDS-Karte für die Gruppe der Exkursionslerner. Die Struktur lässt deutlich eine gerichtete Verschiebung erkennen: Die Punkte für die individuellen Wissenskarten zum Zeitpunkt d1b liegen unten links (Ausnahme: EL4-d1a), die zum Zeitpunkt d3b oben rechts (Die Zielstruktur ist in letzterer Hälfte der Lösung enthalten, so dass diese auch die systematisch bessere darstellt.). Neben der Gerichtetheit konvergieren die individuellen Wissenskarten der Exkursionslerner auch mit der Zeit stärker (durchschnittlicher AvgLoss zum Zeitpunkt d1a = 0.70 und zum Zeitpunkt d3b = 0.62; $t(22) = 3.20$, $p < 0.001$, zweiseitig).

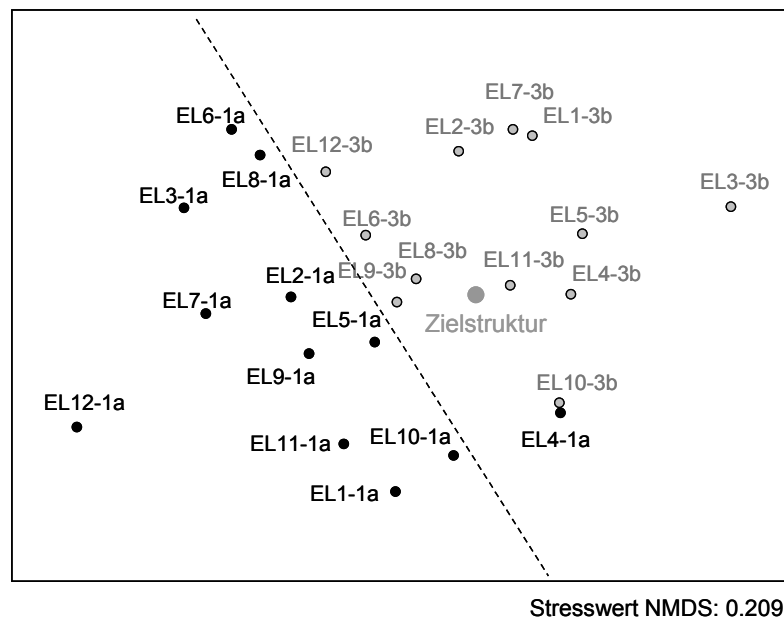
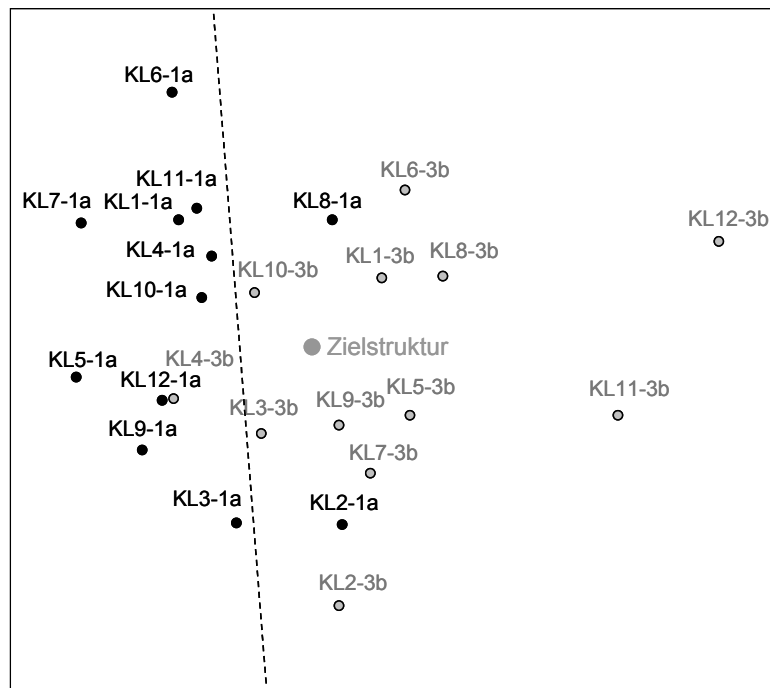


Abbildung 47: NMDS-Lösung der Ähnlichkeiten der individuellen Karten der Exkursionslerner vor dem ersten (d1a) und nach dem dritten Lerndurchgang (d3b)

Die NMDS-Lösung der individuellen Wissensstrukturen der Kärtchenlerner (Abbildung 48) stellt im wesentlichen eine Replikation des oben skizzierten Befundes dar, allerdings nicht in exakt derselben Deutlichkeit: Mit den Kärtchenlernern KL2 und KL8 haben sich bereits vor dem ersten Lernen zwei „heimliche Experten“ eingeschlichen, jedenfalls liegen ihre Wissenskarten verhältnismäßig dicht an der Zielstruktur (ohne dass beide allerdings im Fragebogen herausragend gut abgeschnitten hätten). Auf der anderen Seite bleibt Kärtchenlerner KL4 nach dem dritten Lerndurchgang ein wenig hinter seinen Kollegen zurück. Ohne diese drei Personen könnte man einfach wieder eine senkrechte Trennlinie zwischen d1a und d3b legen.



Stresswert NMDS: 0.232

Abbildung 48: NMDS-Lösung der Ähnlichkeiten der individuellen Karten der Kärtchenlerner vor dem ersten (d1a) und nach dem dritten Lerndurchgang (d3b)

Während sich die Gerichtetheit der Veränderungen also repliziert, nimmt die Konvergenz (anders als bei den Exkursionslernern) in der Gruppe der Kärtchenlerner nicht zu (durchschnittlicher AvgLoss zum Zeitpunkt d1a = 0.58 und zum Zeitpunkt d3b = 0.59; $t(22) = -0.51$, $p = 0.60$, zweiseitig). Numerisch liegt das allerdings am Zeitpunkt d1a, zu dem die Karten untereinander deutlich konsistenter sind als in der Gruppe der Exkursionslerner.

Ganz anders sieht die Befundlage bei den Kontrollpersonen aus (Abbildung 49): Hier hat es keine systematische Verschiebung der Wissensstrukturen gegeben. Im Gegenteil, die beiden Karten der jeweils selben Person sind ungefähr im gleichen Quadranten der Lösung wieder zu entdecken. Ohne das Lernen von Merkmalen bleibt die idiosynkratische Strukturierung der Vogelwelt weitgehend bestehen. Qualitativ bewegt sich nur eine einzige Kontrollperson (KP6) systematisch auf die Zielstruktur zu (wobei sich diese Person im Merkmalsfragebogen aber nicht gesteigert hatte). Alle anderen Verschiebungen verlaufen planlos im Raum. Auch eine Zunahme der Konsistenz der Karten untereinander ist nicht zu entdecken (durchschnittlicher AvgLoss zum Zeitpunkt d1a = 0.74 und zum Zeitpunkt d3b = 0.75; $t(16) = -0.44$, $p = 0.65$, zweiseitig).

Damit lässt sich eine systematische Verschiebung bei denjenigen und nur bei denjenigen Personen konstatieren, die die Merkmale der Vogelarten gelernt haben. Im Sinne der aufgestellten Hypothese ergibt sich daraus die Schlussfolgerung, dass das Lernen von Merkmalen die Bildung von Strukturwissen auf eine systematische Weise beeinflusst. Die erwartete Konvergenz der Karten war jedoch nur bei der Gruppe der Exkursionslerner zu beobachten.

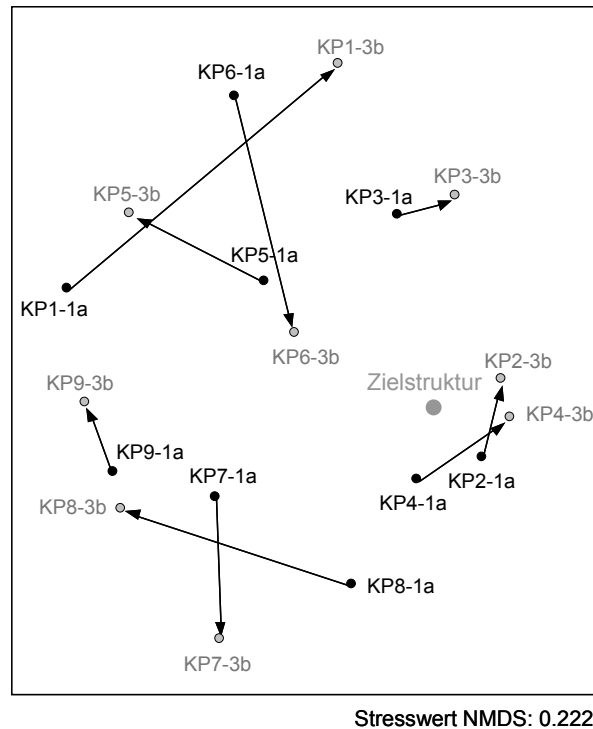


Abbildung 49: NMDS-Lösung der Ähnlichkeiten der individuellen Karten der Kontrollpersonen zum Zeitpunkt d1a und d3b

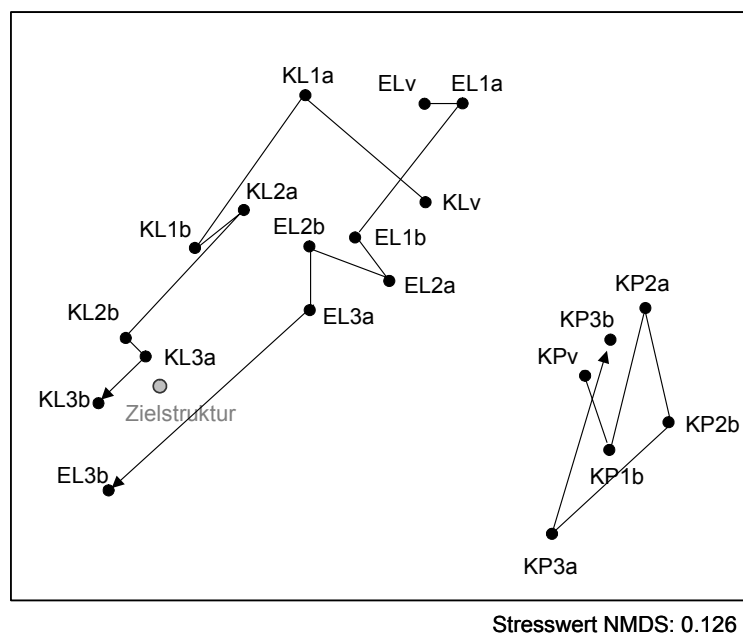


Abbildung 50: Veränderung der Kognitiven Karten (jeweils über die Stichproben gemittelte Proximitäten) über den Lernverlauf (Die Karte d1a bei der Kontrollgruppe degenerierte und ist deswegen nicht skaliert.)

Qualitative Verbesserung der Wissenskarten. Aber bedeutet diese systematische Veränderung auch eine systematische Verbesserung? In den vorausgehenden Abbildungen war bereits die Zielstruktur mit eingetragen, um schon einen ersten Hinweis auf die Beantwortung dieser Frage zu haben. Abbildung 50 zeigt dieselbe Auswertelogik für die NMDS-Karten aller gemittelten Ähnlichkeitsmatrizen über alle Zeitpunkte. In dieser Abbildung wird die Zielgerich-

tetheit besonders deutlich. Immer wenn gelernt wird, bewegt sich die resultierende Karte in Richtung unten links auf das Zielmodell zu.

Wie sich der Wissenszuwachs auf die individuellen Karten der Versuchspersonen auswirkt, ist in Abbildung 51 dargestellt. Die Graphik beinhaltet die strukturelle Differenz zwischen Versuchspersonen und Expertenmodell. Dabei wurde für jede Versuchsperson zu allen Zeitpunkten die Abweichung ihrer individuellen Karte zur Zielstruktur über Prokrustes-Transformation gerechnet (AvgLoss). In der Graphik sind die Mittelwerte der drei Gruppen zu den sechs Messzeitpunkten angegeben.

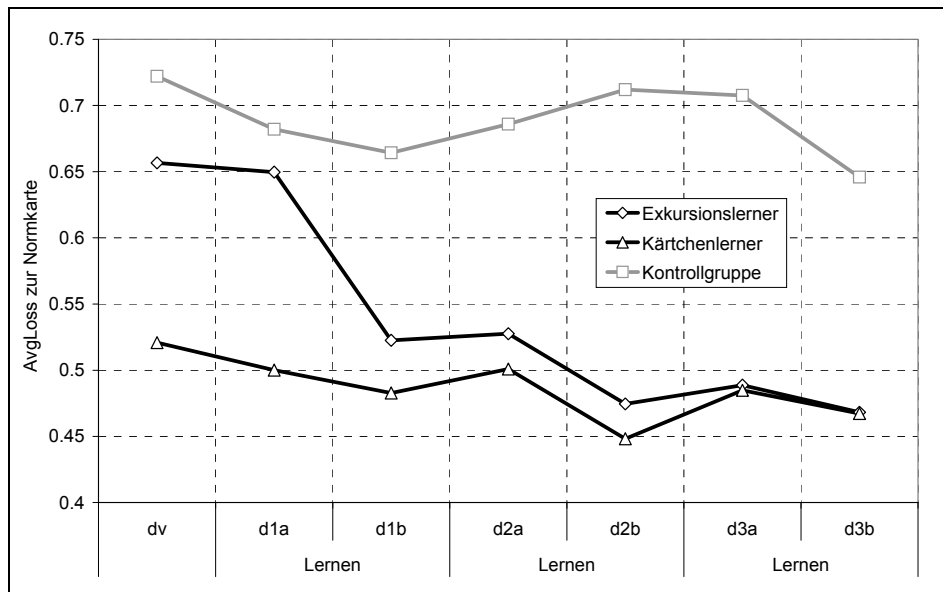


Abbildung 51: Mittlere Differenz (AvgLoss) der individuellen Wissenskarten zur Zielstruktur

Betrachtet man die Resultate vom Ende her, so besitzt Lernen eine klare positive Auswirkung auf die inhaltliche Qualität der individuellen Wissenskarten: Zum Zeitpunkt d3b, nachdem die beiden Experimentalgruppen drei Lernphasen durchlaufen hatten, sind ihre Strukturen der Expertenstruktur signifikant ähnlicher als die Strukturen der Kontrollpersonen (EL-KP: $t(19) = -3.17$, $p < 0.01$, zweiseitig; KL-KP: $t(19) = -2.88$, $p < 0.01$, zweiseitig). Die Karten der nicht lernenden Kontrollpersonen behalten zu allen Messzeitpunkten in etwa dieselbe Abweichung von der Zielstruktur. Zu den Zeitpunkten dv und d1a sind die Strukturen der Exkursionslerner praktisch genauso weit entfernt vom Expertenmodell wie diejenigen der Kontrollpersonen (Zeitpunkt dv: $t(19) = -1.4$, $p = 0.17$, zweiseitig; Zeitpunkt d1a: $t(19) = -0.52$, $p = 0.60$, zweiseitig). Im Unterschied zur Kontrollgruppe gleichen sich ihre Strukturen aber im Verlauf des Lernens nach und nach der Zielstruktur an. Einen grundsätzlich gleichen Lernverlauf kann man auch bei der Gruppe der Kärtchenlerner beobachten. Diese verfügen allerdings zu den Zeitpunkten dv und d1a – aus einem uns nicht erklärbaren Grund – bereits über Wissenskarten, deren AvgLosses signifikant besser sind als die der anderen beiden Gruppen (EL-KL zum Zeitpunkt dv: $t(22) = -4.09$, $p < 0.001$, zweiseitig und zum Zeitpunkt d1a: $t(22) = -3.14$, $p < 0.01$, zweiseitig). Unerklärlich ist dieser Befund deswegen, weil er in den Punktzahlen des Multiple-Choice-Tests keinerlei Entsprechung findet. Die genauere Analyse in Abbildung 52 macht deutlich, dass trotz der nur geringen Verbesserung im AvgLoss eine systematische Veränderung der individuellen Karten stattfindet.

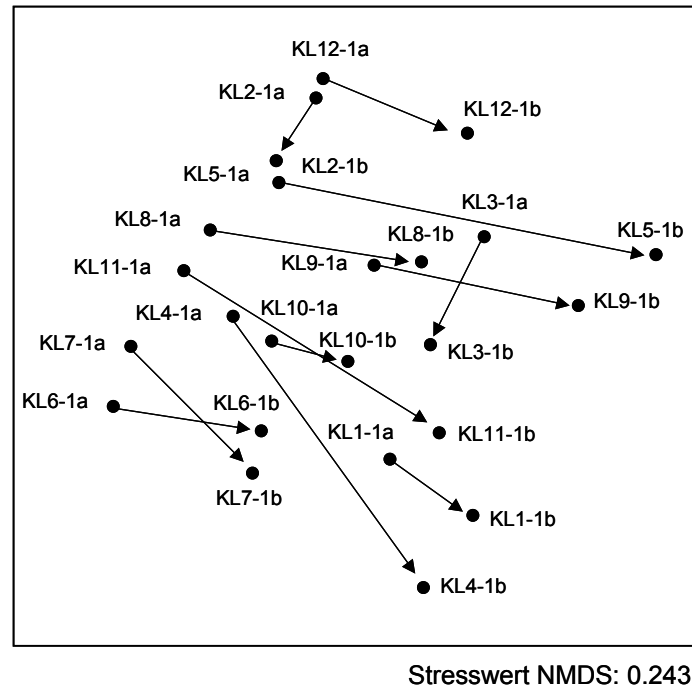


Abbildung 52: NMDS-Lösung der individuellen Karten der Kärtchenlerner zum Zeitpunkt d1a und d1b. Es gibt eine systematische Verschiebung in eine gemeinsame Richtung (nach unten rechts). Alle 12 Karten verschieben sich nach unten, 10 der 12 nach rechts. Obwohl die Karten vom AvgLoss her gesehen nicht nennenswert näher an die Expertenkarte herankommen, ist also trotz allem eine systematische Veränderung zu sehen.

Eine Verbesserung der AvgLoss-Werte fand (wie Abbildung 51 zeigt) jeweils beim Lernen statt, also zwischen der Diagnose d1a und d1b, zwischen d2a und d2b und zwischen d3a und d3b. Während der einwöchigen Pause (d1b zu d2a sowie d2b zu d3a) gab es in beiden Lerngruppen jeweils eine geringe Verschlechterung, die sich als „Vergessen“ interpretieren lässt. Dieser Gesamtbefund (Verbesserung der Karten direkt nach dem Lernen, Verschlechterung nach einer Woche Pause) ist mit einem Binomialtest auf dem 1%-Niveau gegen die Zufallserwartung von .50 sowohl für die Exkursionslerner als auch für die Kärtchenlerner abgesichert (jeweils $N = 60$, fünf prognostizierte „Abs“ und „Aufs“ für jeweils 12 Versuchspersonen gingen in die Verlaufsüberprüfung jeder Lerngruppe ein). Für die Kontrollpersonen ergibt sich kein signifikantes Ergebnis.

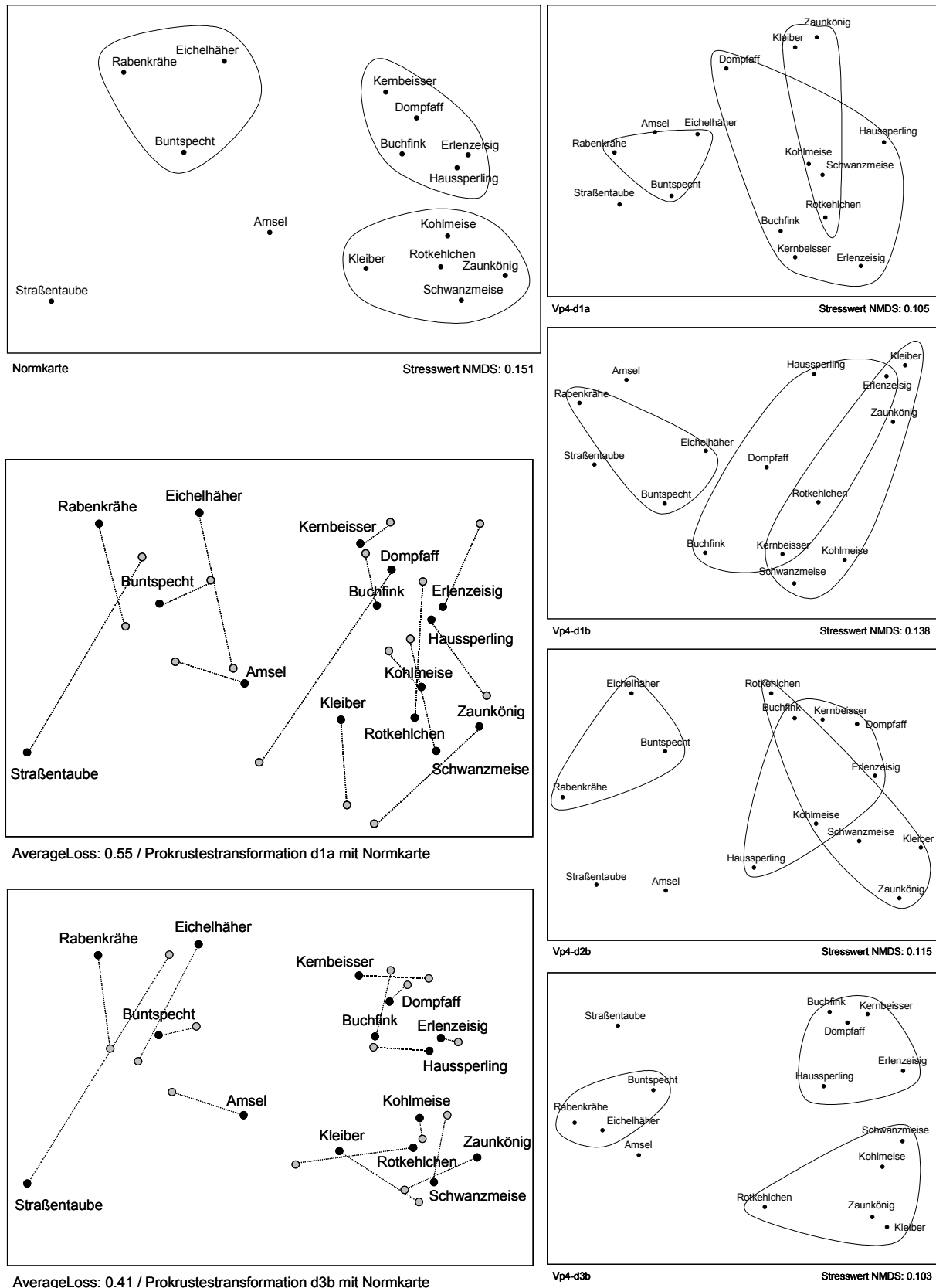


Abbildung 53: Lernfortschritt eines individuellen Lerners (d1a, d1b, d2b, d3b) mit eingezeichneten Expertenclustern und mit Prokrustes-Transformation der Karten d1a und d3b mit der Normkarte

Wirklich dramatisch ist als Einzeleffekt allerdings nur die Verbesserung der Kognitiven Karten der Exkursionslerner im ersten Lerndurchgang, die beim Mittelwertvergleich im t-Test

signifikant wird ($t(22) = 4.06$, $p < 0.001$, einseitig). Offensichtlich geschieht das wesentliche Lernen bereits im ersten Durchgang, und da die Kärtchenlerner bereits vorher recht gut waren (aus, wie gesagt, für uns nicht nachvollziehbaren Gründen), fällt bei ihnen dieser große Strukturverbesserungseffekt aus. Abbildung 53 zeigt eine die Karten einer durchschnittlichen Exkursionslernerin vor und nach dem Lernen. Das moderate Absinken des AvgLosses von 0.55 auf 0.41 geht mit einer sichtbaren Strukturverbesserung einher: Die zusammengehörenden Gruppen der Vogelarten finden zusammen, und die Abweichungen zum Normmodell wirken sich nach dem Lernen nur noch innerhalb der Cluster aus (Die alternative Auswertung, welche die Cluster aus der Normkarte in die Lernerkarte legt, zeigt sehr schön, wie sich die relevanten Gruppen finden). Die größte Abweichung besteht nach dem Lernen bei der Straßentaube, was aber keinen sonderlich hohen Stellenwert besitzt, da sie sowohl in der Normkarte als auch in der Lernerkarte eine Einzelposition bezieht.

7.2.4. Diskussion

Der Hauptbefund des Experiments entspricht den Erwartungen an die individuellen Wissenskarten: Sie verbessern sich mit dem Lernen systematisch auf die Zielstruktur zu, während in der einwöchigen Pause zwischen den Lernsitzen jeweils wieder etwas von der Qualität der Karten verloren geht. Dass die individuellen Karten das Zielmodell nach drei Lerndurchgängen nicht vollständig erreichen (und im Durchschnitt weiter von diesem weg bleiben als die befragten langjährigen Experten), lässt sich mit der Schwierigkeit der Materie erklären: Die Punktzahlen in den Multiple-Choice-Fragebögen signalisieren noch ausreichend Spielraum nach oben hin, so dass die verbleibende Diskrepanz zur Zielstruktur zumindest zum Teil auf ein noch mangelhaftes Wissen zurückgeführt werden kann. Dass jedoch auch die Experten individuell deutlich vom Zielmodell abweichen, spricht vielleicht für die Komplexität des gewählten Wissensgebietes oder aber dafür, dass dieses Wissensgebiet eben auch Raum für idiosynkratische Strukturvorstellungen lässt. Da vergleichende Untersuchungen bislang völlig fehlen, ist als Alternativerklärung aber auch nicht auszuschließen, dass das technische Verfahren (Ähnlichkeitsurteile, NMDS und Prokrustes-Transformation) gar keine wirklich gut aufeinander passenden individuellen Karten zu produzieren in der Lage ist. Deswegen sollte Experiment 2 versuchen, die ersten Erklärungen (Komplexität und unvollständige Konvergenz des Wissensgebietes) nach Möglichkeit auszuschalten, um die Grenze des durch das Verfahren technisch Machbaren etwas genauer auszuloten.

Dass sich die individuellen Wissenskarten zwar systematisch verbessern, dabei jedoch nur in der Gruppe der Exkursionslerner untereinander konvergieren, hängt mit dem unerklärlich guten Ausgangswert der Kärtchenlerner zusammen. Aber auch für die Exkursionslerner ist der Effekt trotz statistischer Signifikanz numerisch geringer als erwartet (von einem AvgLoss von 0.70 hin zu einem AvgLoss von 0.62). Das könnte damit zusammenhängen, dass die Versuchspersonen in der Regel sechs der 15 Arten bereits kannten und hinsichtlich der verbleibenden neun Arten zwar inadäquate, aber doch teilweise konvergierende Vorstellungen hatten. Auf diese Weise könnte es zur der beobachteten „Parallelverschiebung“ der individuellen Karten durch das Lernen gekommen sein. Natürlich ist das eine in den Details schwer zu überprüfende Interpretation, doch eine letztlich Klärung können die vorliegenden Daten wohl nicht bringen. Wir setzen damit unsere Hoffnungen hinsichtlich der noch zu zeigenden Konvergenz auf Experiment 2.

Der größte Sprung in der Verbesserung der Wissenskarten wird durch die erste Lerneinheit ausgelöst, besonders ausgeprägt zu sehen für die Gruppe der Exkursionslerner. Dieser Effekt kann zweierlei bedeuten: Entweder ist das „Aha“-Erlebnis, wie die einzelnen Vögel zueinander stehen, bereits durch das Kennen lernen der 15 Arten hinlänglich gegeben. Oder aber die Merkmalsinformationen werden im zweiten und dritten Lerndurchgang lediglich wieder aufgefrischt, es wird aber eigentlich nichts mehr dazugelernt. Beide Erklärungen widersprechen

sich nicht unbedingt und sind aus den vorliegenden Daten nicht überprüfbar. Es bleibt also eine offene Frage für zukünftige Forschung, wie viel Merkmalsinformation über die einzelnen Objekte bekannt sein muss, damit eine Person mit ihren Ähnlichkeitsurteilen eine in den Grundzügen stimmige Wissenskarte produzieren kann.

7.3. Experiment 2

7.3.1. Fragestellung und Hypothesen

Experiment 1 hat gezeigt, dass sich die Wissenskarten merkmalslernender Personen verändern und verbessern. Allerdings konnte dabei keine signifikante Konvergenz der Karten festgestellt werden, und wegen des doch recht anspruchsvollen Lernmaterials fand auch keine perfekte Angleichung der Karten an den Stand der Experten statt. Das zweite Experiment soll deswegen – parallel aufgebaut, jedoch mit einfacher gestricktem und perfekt konvergentem Lernmaterial – ausloten, wie stark die mögliche Annäherung an eine „perfekte“ Wissenskarte sein kann.

7.3.2. Methode und Durchführung

Aufbau des Experiments. In einer Längsschnittuntersuchung hatten die Versuchspersonen die Aufgabe, zu 15 Ländern die drei Merkmale Landesfläche, Einwohnerzahl und Bruttosozialprodukt so zu lernen, dass sie daraus eine Wissenskarte aufbauen konnten. In vier Sitzungen, die jeweils im Wochentakt stattfanden, erhielten sie dazu Informationen zu diesen 15 Objekten. Mittels paarweiser Ähnlichkeitsurteile unmittelbar vor und unmittelbar nach der Lerneinheit wurde dann die Auswirkung der Lerneinheiten auf die individuellen Wissenskarten der Versuchspersonen überprüft. Zu Beginn jeder Sitzung erhob ein computerbasierter Multiple-Choice-Fragebogen den aktuellen Wissensstand über die drei Merkmale zu jedem Land. In diesem Fragebogen hatten die Versuchspersonen die Länder einer Liste aller einzelnen Merkmalsausprägungen (plus Distraktoren) zuzuordnen. Zu jeder Angabe war zudem ein Konfidenzurteil abzugeben, so dass der Fragebogen zwischen lückenhaftem Wissen und falschem Wissen (die Angabe ist inkorrekt, die Person ist sich aber der Korrektheit subjektiv sicher) unterscheiden konnte.

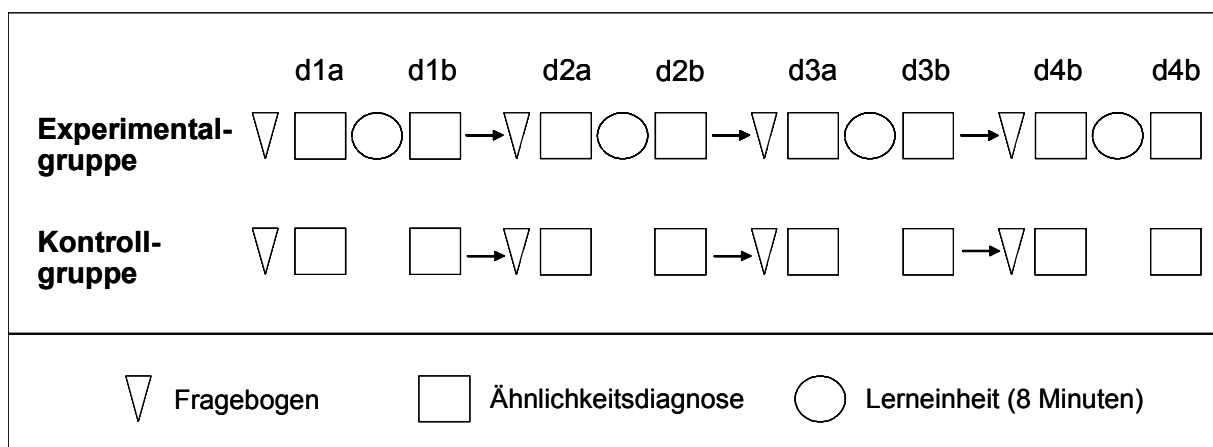


Abbildung 54: Schematische Darstellung des Ablaufs von Experiment 2

Die Versuchspersonen wurden vor dem Experiment zufällig in eine Experimentalgruppe und eine Kontrollgruppe aufgeteilt. Die Versuchspersonen der Kontrollgruppe lernten nicht, sondern führten lediglich das komplette Programm der Wissensdiagnostik durch. Der Aufbau des Experiments ist in Abbildung 54 skizziert.

Auswahl der Länder und Bestimmung der Zielstruktur. In einem Vorversuch mit insgesamt 46 Versuchspersonen hatte sich aus einer Vorauswahl von 26 Ländern der Pool von 15 Ländern als gut geeignet herausgestellt, der zur Grundlage der Zielstruktur (Abbildung 55) gemacht wurde. Außerdem hatten sich die drei intervallskalierten Merkmale Landesfläche, Einwohnerzahl und Bruttosozialprodukt pro Kopf als günstig für Lernen, Wissensdiagnose und Zielstruktur erwiesen. Die Zielstruktur errechnete sich über die Unähnlichkeiten der einzelnen Länder auf diesen drei Merkmalen. Dazu wurde folgende Formel verwendet:

$$U\ddot{A}_{(x,y)} = \frac{|(x_i - y_i) / (x_i + y_i)| + |(x_j - y_j) / (x_j + y_j)| + |(x_k - y_k) / (x_k + y_k)|}{3}$$

Wobei: $U\ddot{A}_{(x,y)}$: Unähnlichkeitsquotient eines Länderpaares
 x, y : Länderindex
 i : Einwohnerzahl
 j : Fläche in km²
 k : BSP pro Kopf in US\$

Dieser Quotient $U\ddot{A}_{(x,y)}$ beschreibt den Betrag des prozentualen Unterschiedes je zweier Länder aufgrund ihrer Merkmalsausprägungen. Die resultierende Dreiecksmatrix wurde mittels NMDS (RobuScal-Algorithmus) in eine zweidimensionale Karte transformiert, die für die Qualitätsabschätzung der individuellen Wissenskarten jeweils als Norm gilt (Abbildung 55).

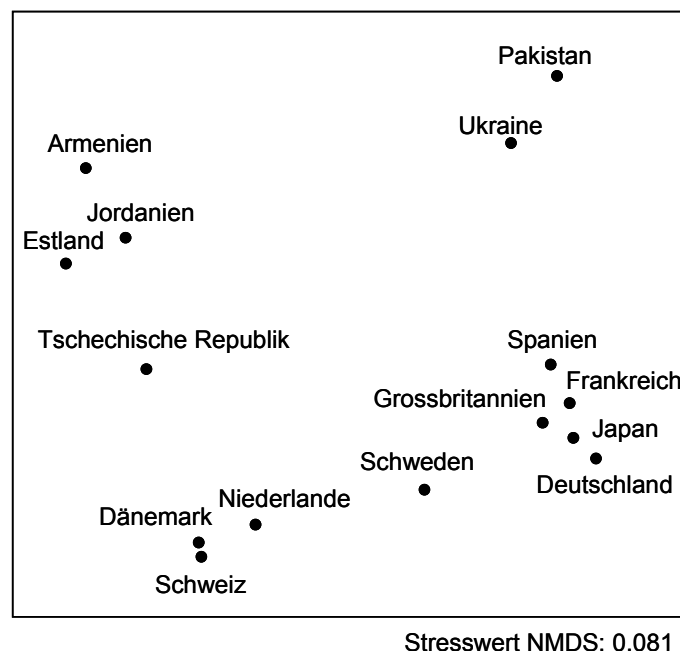


Abbildung 55: NMDS-Karte, die in Experiment 2 die Zielstruktur des Lernens ergibt.

Lernmaterialien. Das Lernmaterial bestand aus 15 Kärtchen, auf denen zu je einem Land die Angaben über Einwohnerzahl, Fläche und Bruttosozialprodukt standen. Die Versuchsperso-

nen waren frei in ihrer Lernstrategie; sie konnten die Lernkarten ordnen und Gruppen bilden, ausbreiten usw. Die Lerneinheit dauerte jeweils acht Minuten.

Versuchspersonen. 37 Personen (17 weiblich; 20 männlich) im Alter zwischen 19 und 33 Jahren nahmen an dem Experiment teil. Es handelte sich dabei hauptsächlich um Studentinnen und Studenten der Psychologie, Soziologie, Tiermedizin und Pharmazie. Hinzu kamen einige Versuchspersonen ohne Hochschulausbildung. 25 Personen wurden der Experimentalgruppe zugeordnet, 12 der Kontrollgruppe. Die Variablen Alter, Geschlecht und Schulbildung waren in beiden Gruppen gleichverteilt.

7.3.3. Resultate

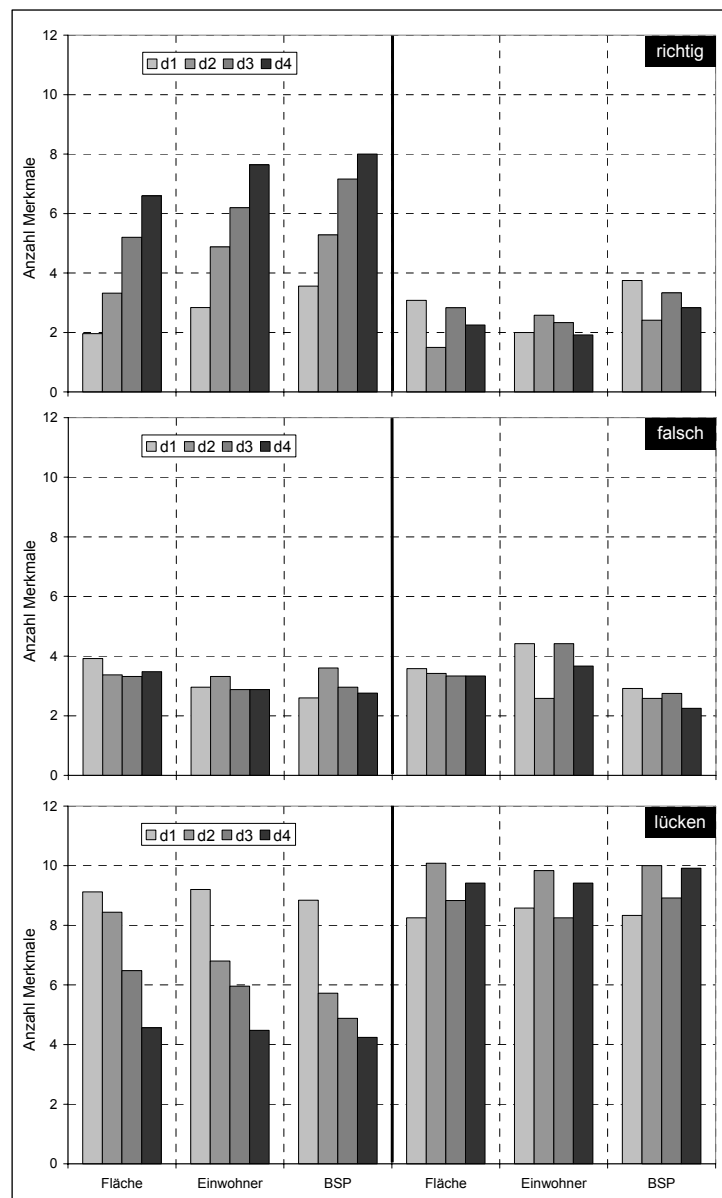


Abbildung 56: Verlauf des Merkmalswissens der Experimentalgruppe (links) und der Kontrollgruppe (rechts) in Experiment 2. Angegeben sind die durchschnittlichen Anzahlen pro Versuchsperson über alle 15 Länder; d1 – d4 = Erhebungszeitpunkte.

Merkmalswissen wird gelernt. Zu Beginn jeder der vier Sitzungen wurde das Merkmalswissen der Versuchspersonen mit einem Zuordnungs-Fragebogen geprüft. Dabei galt ein Datum als

gewusst, wenn die Angabe innerhalb einer 20%-Marke des tatsächlichen Wertes gemacht wurde. Abbildung 56 zeigt, dass für jedes der drei Merkmale in der Experimentalgruppe ein deutlicher Wissenszuwachs zu verzeichnen ist. Dabei basiert dieser Fortschritt vor allem auf einem „Auffüllen“ von Wissenslücken: Die Zunahme des korrekten Wissens und die gleichzeitige Abnahme der Wissenslücken zeigen für jedes der drei Merkmale in einer zweifaktoriellen Varianzanalyse (ANOVA) mit den zwei Intersubjektfaktoren Merkmal und Zeit signifikante Effekte des Merkmals, $F(1.78, 42.82) = 10.14$, $MSE = 86.30$, $p < .001$, und der Zeit, $F(2.0, 47.98) = 53.57$, $MSE = 457.44$, $p < .001$. Dabei wird von vornherein das Bruttosozialprodukt etwas besser gewusst als die Einwohnerzahl, und diese wiederum besser als die Fläche des jeweiligen Landes. Der Anteil falschen Wissens bleibt über alle Durchgänge und über alle Merkmale hinweg erstaunlich konstant bei ca. 20% des Gesamtwissens.

Bei der Kontrollgruppe ergibt sich keinerlei signifikante Veränderung. Diese Versuchspersonen steigen auf demselben Wissensniveau ein wie die Personen der Lerngruppe, verbleiben dort jedoch auch. Die Lerner sind bereits ab dem zweiten Durchgang signifikant besser, und diese Schere geht von Durchgang zu Durchgang weiter auf (d1: $t(109) = -0.33$, $p = 0.37$; d2: $t(91) = 4.79$, $p < 0.001$; d3: $t(99) = 6.23$, $p < 0.001$; d4: $t(106) = 9.14$, $p < 0.001$; einseitig).

Insgesamt zeigt Abbildung 56, dass die Versuchspersonen ihre Ähnlichkeitsurteile durchaus nicht auf dem Niveau perfekten Wissens abgeben (zumindest bis zum Zeitpunkt d4a, denn nach dem vierten Lerndurchgang wurden ja keine Fragebogendaten mehr erhoben). Die im folgenden zu betrachtenden Wissenskarten müssen also vor dem Hintergrund gelesen werden, dass den Versuchspersonen – selbst wenn sie ihr Wissen durch die Lerndurchgänge fast verdreifacht haben – noch immer ca. 30% der Informationen unbekannt sind und dass sie bei 20% des Gesamtwissens von inadäquaten Vorstellungen ausgehen.

Konvergente Veränderung der Wissenskarten. Wie in Experiment 1, so wurden auch hier alle individuellen Ähnlichkeitsmatrizen in zweidimensionale NMDS-Karten überführt und paarweise einer Prokrustes-Transformation unterzogen. Die Dreiecksmatrix der resultierenden AvgLoss-Werte zwischen allen individuellen Karten der Lernergruppe aus Durchgang d1a (vor dem ersten Lernen) und Durchgang d4b (nach dem vierten Lernen) wurde mit NMDS als zweidimensionaler Raum skaliert. Abbildung 57 zeigt die systematische Verschiebung der Wissensstrukturen nach dem Lernen in die rechte Region der Karte. Lediglich zwei (dunkelgrau markierte) Wissenskarten bleiben beim Niveau vor dem Lernen, und die Wissenskarten von vier Versuchspersonen (plus eine Kontrollperson) hatten sich bereits von Anfang an in der Zielregion befunden (zur Verdeutlichung dieses Befundes ist eine „Demarkationslinie“ in die Abbildung eingetragen). Die Kontrollpersonen positionieren sich in der Region der Experimentalgruppe und zeigen keine systematische Verschiebung.

Anders als in Experiment 1 zeigt sich neben dieser systematischen Verschiebung auch eine Konvergenz der individuellen Wissensstrukturen: Die mittlere Distanz zwischen allen Wissenskarten beträgt vor dem Lernen 0.50, nach dem Lernen 0.38. Dieser Unterschied ist im t-Test signifikant ($t(299) = 15.12$, $p < 0.001$). Für die Wissensstrukturen der Kontrollpersonen findet sich weder eine systematische Verschiebung noch eine Konvergenz. Damit bestätigt sich für das herangezogene Wissensgebiet der Länder, dass die individuellen Karten durch das Merkmalslernen systematisch konvergieren.

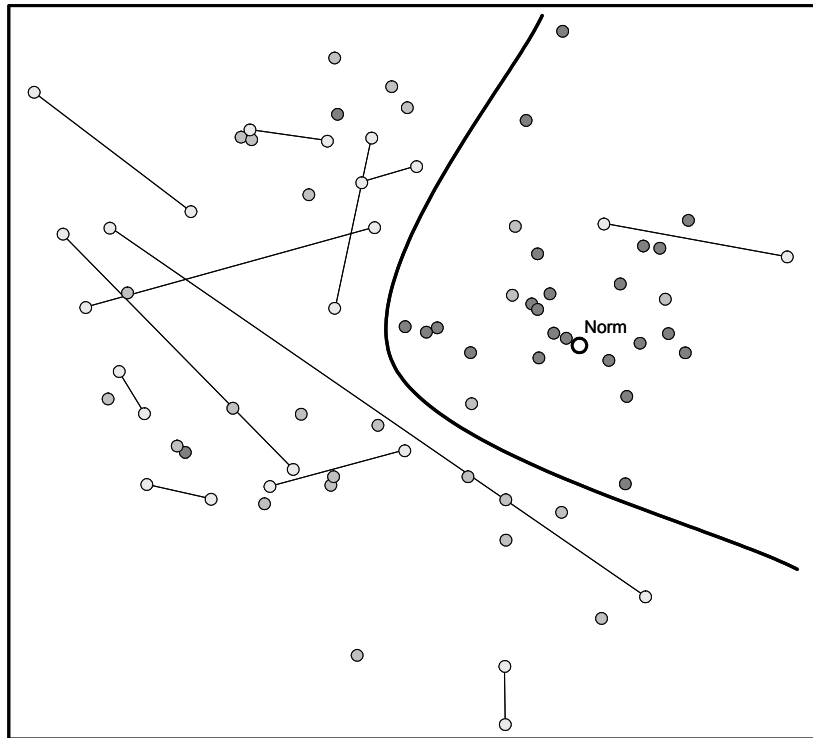


Abbildung 57: NMDS-Lösung mit allen individuellen Wissenskarten vor dem ersten Lernen (d1a = grau) und nach dem vierten Lernen (d4b = dunkelgrau). Zusätzlich ist zur Orientierung die Zielstruktur des Lernens (Norm) eingetragen. Die verbundenen, hellgrauen Punkte stehen für die Karten der Kontrollgruppe zum Zeitpunkt d1a und d4b.

Qualitative Verbesserung der Wissenskarten. In Abbildung 57 ist bereits die Zielstruktur des Lernens mit eingetragen. Aus der systematischen Verschiebung auf diese Zielstruktur hin ergibt sich bereits ein erster Hinweis auf die qualitative Verbesserung der Wissenskarten nach dem Lernen. Dieser wird durch eine Analyse aller AvgLosses zur Zielstruktur untermauert (Abbildung 58): Nach jeder der vier Lerneinheiten nähern sich die individuellen Wissenskarten unmittelbar der Zielstruktur an, so dass sie am Ende des Experiments bei einem AvgLoss von durchschnittlich 0.29 enden (der Unterschied zum Start des Experiments ist im t-Test auf dem 0,1%-Niveau signifikant; $t(24) = 6.26$, $p < 0.001$). Interessant ist auch, dass ein Teil der Verbesserung in den Wissenskarten in der Woche zwischen den Sitzungen jeweils wieder verloren geht. In Sitzung 3 war der Lerneffekt so gering, dass davon zu Beginn der vierten Sitzung nichts mehr übrig ist. Ansonsten überwiegt aber der Lerneffekt deutlich den Vergessenseffekt. Das Muster von Verbesserung der Wissenskarten durch Merkmalslernen und Verschlechterung durch eine Woche Pause wird im Binomialtest auf dem 0,1%-Niveau signifikant.

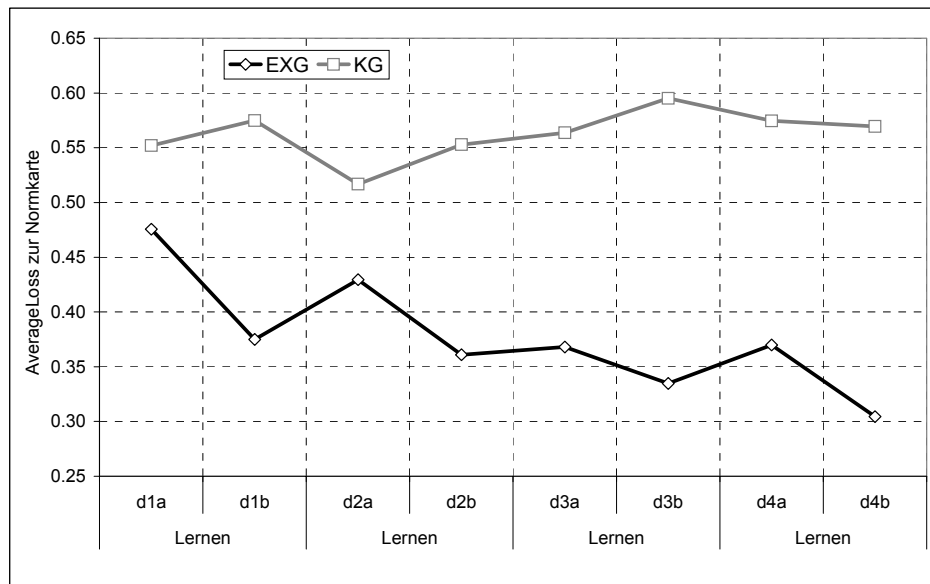
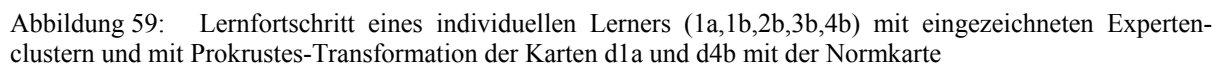


Abbildung 58: Gemittelte Abweichungen aller individuellen Wissenskarten von der Zielstruktur

Die Versuchspersonen der Kontrollgruppe hingegen bilden durchgängig schlechtere Karten als die Lernergruppe (einzig die Unterschiede beim Ausgangswert in d1a sind nicht signifikant; $t(35) = -1.53$, $p = 0.07$, einseitig). Die gemittelten AvgLosses schwanken unsystematisch um den Ausgangswert von 0.55, ohne sich systematisch zu verbessern oder zu verschlechtern (t-Test zwischen d1a und d4b: $t(11) = -0.38$, $p = 0.35$). Das mehrfache Abgeben von Ähnlichkeitsurteilen allein ergibt also keine Verbesserung in den Wissenskarten.

Der Lernfortschritt einer überdurchschnittlichen Person aus der Lernergruppe ist in Abbildung 59 illustriert (aufgebaut analog zu Abbildung 52). Die große Verbesserung im AvgLoss (von 0.59 auf nur noch 0.22) korrespondiert mit einer Karte, in der nur noch Japan maßgeblich von der Zielposition entfernt liegt, indem der Lerner dieses asiatische Land nicht vollständig in die Gruppe der ansonsten europäischen großen Industrienationen integriert. Das Einzeichnen der Cluster aus der Normkarte macht deutlich, dass der Lerner die grundsätzliche Zweiteilung bereits nach dem ersten Lernen (AvgLoss 0.39) und die weitere Unterteilung in vier Cluster bereits nach dem zweiten Lernen (AvgLoss 0.28) in seiner Wissenskarte repräsentiert hatte. Die qualitativen Beobachtungen bei der Interpretation der Karte korrespondieren damit sehr gut mit dem quantitativen Maß des AvgLoss und verdeutlichen die zielgerichtete Annäherung des strukturellen Wissens des Lernenden hin zur normativen Zielstruktur.



7.3.4. Diskussion

Auch Experiment 2 bestätigt die Hypothese, dass sich Wissenskarten durch das Lernen einzelner Merkmale substantiell verbessern. Zusätzlich ist es durch die Auswahl des Wissensgebietes und des Lernmaterials gelungen, den gewünschten Konvergenz-Effekt zwischen den individuellen Karten zu erzielen. Die Wissenskarten liegen außerdem – wie aufgrund des vereinfachten Materials angenommen – wesentlich dichter an der Zielstruktur als bei dem sehr komplexen Wissensfeld aus Experiment 1.

Gleichwohl bleibt eine mittlere AvgLoss-Distanz von 0.30 zur Zielstruktur bestehen. Ein Teil dieses Unterschiedes lässt sich sicherlich darauf zurückführen, dass die Versuchspersonen selbst vor dem vierten Lerndurchgang (d.h. im letzten der vier Fragebögen) nur über ein korrektes Merkmalswissen von ca. 50% verfügten. Zukünftige Forschung könnte hier also mit einem strengerem Kriteriumslernen noch weiter ausloten, wie nahe man Lerner durch reines Merkmalslernen an eine integrierte Zielstruktur heranführen kann. Als unsystematischer Hinweis dazu mag gelten, dass wir drei Personen einmal die Kärtchen mit den Merkmalen gegeben haben und sie baten, aufgrund dieser vorliegenden Informationen paarweise Ähnlichkeitsurteile zwischen den 15 Ländern zu bilden. Die entstandenen drei „Wissens“karten wichen durchschnittlich mit einem AvgLoss von 0.16 von der Zielstruktur ab. Aber hier handelt es sich um gegebene Information, und nicht um integriertes Wissen. Aus diesem Grund wären wir skeptisch, ob derart von Versuchspersonen konstruierte Ähnlichkeitsurteile auf dieselbe Weise zustande kämen. Trotzdem sollten die gemessenen Werte ermuntern, nach einer weiteren Annäherung an eine Zielstruktur zu suchen.

7.4. Gesamtdiskussion

Die zwei präsentierten Experimente haben aufgezeigt, dass sich das Lernen von Merkmalsinformationen auf die Ähnlichkeitseinschätzung von Objekten auswirkt und dieses wiederum einen systematischen strukturierenden Einfluss auf die Wissenskarten einer Person hat. Damit bietet sich das vorgeschlagene Verfahren – die Umsetzung einer Matrix von paarweisen Ähnlichkeitsurteilen in eine NMDS-Lösung und nachfolgende Prokrustes-Transformation dieser Karte mit einer Zielstruktur oder einem Expertenmodell – als generelles Werkzeug für die Messung des Wissensgrades einer Person in objektbasierten Inhaltsgebieten an. Hier kann nicht nur der AvgLoss als eine quantitative Größe für die qualitative Abweichung des Wissens einer Person von einer gesetzten Norm verwendet werden. Die dargestellten Einzelfallbeispiele untermauern auch die intuitive Interpretierbarkeit der strukturellen Details solcher NMDS-Lösungen.

Unseres Wissens handelt es sich bei den vorgelegten Experimenten um die erste systematische Studie, die den Einfluss von Lernen auf NMDS-Wissenskarten untersucht und dabei unterschiedliche Wissensstände zu unterschiedlichen Zeitpunkten miteinander vergleicht. Der bisherige Standard in der psychologischen und sozialwissenschaftlichen Forschung war der Einsatz dieses Skalierungsverfahrens für gemittelte Stichprobendaten. Das hatte, wie bereits in der Einleitung ausgeführt, seinen guten technischen Grund in der mangelhaften Robustheit (und damit mangelhaften Reliabilität) der vorliegenden Algorithmen. Durch den RobuScal-Algorithmus wird aber nun die Skalierung individueller Datensätze auch dann möglich, wenn diese möglicherweise mit Ausreißern behaftet sind. Das eröffnet dem Verfahren Perspektiven für all jene Wissensbereiche, die sich durch zueinander in Relation stehende Objekte definieren.

Dabei verstehen wir die Wissenskarten als Ausdruck einer Messung, und nicht als eine explizite Vorstellung, über die eine Person in dieser Form verfügt (Dass Letzterem so ist, drückt

sich in der Schwierigkeit aus, die eine Person hat, wenn man sie bittet, zu einem bestimmten Gebiet ihre Wissenskarte zu zeichnen – in der Regel gelingt dies entweder gar nicht oder nur sehr fehlerhaft.). Komplexe Wissenskarten können deswegen mit dem Verfahren von Ähnlichkeitsurteilen „gezeichnet“ werden, weil diese Paarurteile dasjenige sequentialisiert abfragen, was im Resultat dann parallel vorliegt. Derartig komplexes relationales Wissen bleibt für die Person selbst in der Regel etwas Implizites. Lediglich einfaches kategoriales Wissen (wie die Zuordnung von Tier- oder Pflanzenarten zu zoologischen oder botanischen Gruppen) oder klar dimensional strukturiertes relationales Wissen (wie z.B. die Links-Rechts-Achse für die Verortung politischer Parteien) ist in seiner Relationalität einer Person mit entsprechender Expertise auch bewusst. Nachdem die vorliegenden Experimente den qualitativen Zusammenhang dieser Karten mit dem Merkmalswissen gezeigt haben, ist die Möglichkeit des Explizitmachens strukturell sonst zumeist nicht bewusster Anteile des integrierten Wissens ein weiteres Argument, das vorgeschlagene Verfahren in der Wissensdiagnostik einzusetzen.

Ein drittes Argument schließlich ist die Form des Wissens, das gemessen wird. Durch das Abgeben von Ähnlichkeitsurteilen wird von den getesteten Personen eine Integrationsleistung ihres Merkmalswissens erwartet, die nicht erbracht werden kann, wenn Wissen lediglich stupide auswendig gelernt wurde. Insofern hat die Messung von Wissenskarten auch gegenüber gleich gut automatisierbaren Multiple-Choice-Tests einen Vorteil: Die paarweise Einschätzung von Ähnlichkeiten zwischen zwei Objekten stellt bereits die erste „Anwendung“ des erlernten Merkmalswissens dar, und diese Anwendung kann nur in der integrierten Form eines Gesamturteils korrekt erfolgen, wenn die einzelnen Daten auch verstanden wurden.

Diese Argumente sprechen für einen Einsatz von NMDS-Wissenskarten in der diagnostischen Praxis. Auch wenn die vorgelegten Experimente die grundsätzliche Abbildungsleistung von Merkmalswissen belegen konnten, bleibt dennoch weitere Forschung notwendig. Zu allererst sollte ausgetestet werden, wie weit sich die Interpretation der Kartengüte nicht nur auf die Gesamtqualität beziehen lässt, sondern auch auf Details in der Karte ausgeweitet werden kann. Ein Schritt in diese Richtung könnte der Abbildungsfehler der einzelnen Objekte in der Prokrustes-Transformation sein (Läge, Rüfenacht, Fichter & Streule, eingereicht). Des Weiteren sollte mit noch stärker kontrollierten Lernexperimenten gezeigt werden, wie dicht derartige individuelle Karten an eine Wissensnorm herangeführt werden können. Außerdem weist Läge (2004) darauf, dass es auch nicht-konvergente Wissensgebiete gibt, für die es nicht möglich sein dürfte, eine Wissensnorm in Form einer eindeutigen Karte zu formulieren. Hier müssten andere Mittel entwickelt werden, um die Güte der Karte zu testen. Dass sich individuelle Wissenskarten aber systematisch mit dem Stand von Merkmalswissen verändern, ist ein erster wichtiger Schritt auf dem Weg zu einem neuen Messverfahren in der Wissenspsychologie.

7.5. Literaturverzeichnis

- Aebli, H. (1976). *Psychologische Didaktik. Didaktische Auswertungen der Psychologie von Jean Piaget*. Stuttgart: Klett.
- Ausubel, D. P. (2000). *The acquisition and retention of knowledge*. Dodrecht: Kluwer Academic Publishers.
- Bonato, M. (1990). *Wissensstrukturierung mittels Struktur-Lege-Techniken. Eine graphentheoretische Analyse von Wissensnetzen*. Frankfurt a.M.: Lang.
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling – Theory and applications*. New York: Springer.

- Eckert, A. (1998). *Kognition und Wissensdiagnose. Die Entwicklung und empirische Überprüfung des computergestützten wissensdiagnostischen Instrumentariums „Netzwerk-Elaborierungs-Technik (NET)“*. Lengerich: Pabst.
- Eckert, A. (1999). Die „Mannheimer Netzwerk-Elaborierungs-Technik (MaNET)“ – Ein computergestütztes Instrument zur Analyse vernetzten Wissens. In W. Schulz (Hrsg.), *Aspekte und Probleme der didaktischen Wissensstrukturierung*. Frankfurt: Peter Lang. S. 93-111.
- Eckert, A. (2000). Die Netzwerk-Elaborierungs-Technik (NET) – Ein computerunterstütztes Verfahren zur Diagnose komplexer Wissensstrukturen. In H. Mandl & F. Fischer (Hrsg.), *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken*. Göttingen: Hogrefe. S. 137-157.
- Edelman, S. (1998). Representation is representation of similarities. *Behavioral and Brain Sciences*, 21, 449-498.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhardt.
- Gower, J. C. & Dijksterhuis, G. B. (2004). *Procrustes problems*. Oxford: Oxford University Press.
- Inhelder, B., De Caprona, D. & Cornu-Wells, A. (Eds.). (1987). *Piaget Today*. Sussex: Erlbaum.
- Inhelder, B., Sinclair, H. & Bovet, M. (1974). *Apprentissage et structures de la connaissance*. Presses Universitaires de France.
- Janetzko, D. & Strube, G. (2000). Knowledge Tracking – Eine neue Methode zur Diagnose von Wissensstrukturen. In H. Mandl & F. Fischer (Hrsg.), *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken*. Göttingen: Hogrefe. S. 199-217.
- Klauer, K. J. (1989). Allgemeine oder bereichsspezifische Transfereffekte eines Denktrainings. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 21, 185-200.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Läge, D. (2004). *Konvergente und nicht-konvergente Veränderung der Ähnlichkeitsurteile beim Merkmalslernen*. AKZ-Forschungsbericht Nr. 2. Universität Zürich: Angewandte Kognitionspsychologie.
- Läge, D., Daub, S., Bosia, L., Ryf, S. & Jäger, C. (in review). *Robustness in nonmetric multidimensional scaling*.
- Läge, D., Rüfenacht, I., Fichter, C. & Streule, R. (eingereicht). *Der ObjectLoss einer Prokrustes-Transformation als Maß für Wissen*.
- Lee, K., Byatt, G. & Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: testing the face-space framework. *Psychological Science*, 11, 379-385.
- Mandl, H. & Fischer, F. (Hrsg.). (2000). *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken*. Göttingen: Hogrefe.
- Mandl, H. & Spada, H. (1988). *Wissenspsychologie*. München: Psychologie Verlags Union.
- Marx, W. & Hejj, A. (1989). *Subjektive Strukturen*. Göttingen: Hogrefe.
- Marx, W. & Läge, D. (1995). *Der ideologische Ring*. Göttingen: Hogrefe.
- Mathar, R. (1997). *Multidimensionale Skalierung*. Stuttgart: Teubner.
- Montada, L. (1995). Die geistige Entwicklung aus der Sicht Piagets. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (3. Aufl.). Weinheim: Psychologie Verlags Union. S. 622-653.
- Piaget, J. & Inhelder, B. (1974). *Gedächtnis und Intelligenz*. Olten: Walter.
- Piaget, J. (1973). *Das Erwachen der Intelligenz beim Kinde*. Stuttgart: Klett.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Piaget, J. (2001). *Studies in reflecting abstraction*. Sussex: Psychology Press.

- Plötzner, R., Beller, S. & Härder, J. (2000). Wissensvermittlung, tutoriell unterstützte Wissensanwendung und Wissensdiagnose mit Begriffsnetzen. In H. Mandl & F. Fischer (Hrsg.), *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken*. Göttingen: Hogrefe. S. 180-198.
- Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing*. Cambridge, Mass.: MIT-Press. Pp. 227-270.
- Scheele, B. & Groeben, N. (1984). *Die Heidelberger Struktur-Lege-Technik (SLT). Eine Dialog-Konsens-Methode zur Erhebung Subjektiver Theorien mittlerer Reichweite*. Weinheim: Beltz.
- Steiner, G. (1996). *Lernen – Zwanzig Szenarien aus dem Alltag* (2. vollst. überarb. Aufl.). Bern: Huber.
- Weber, S. & Schumann, M. (2000). Concept Mapping Software Tool (COMASOTO) zur Diagnose strukturellen Wissens. In H. Mandl & F. Fischer (Hrsg.), *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken*. Göttingen: Hogrefe. S. 158-179.

8.

Der NMDS-Stresswert als Maß für Assimilation und Akkomodation beim Lernen von Wissensstrukturen

*Zusammenfassung*³⁶

Assimilation und Akkomodation werden von Piaget als Teilprozesse bei der Bildung eines kognitiven Gleichgewichts im Wissenserwerb verstanden. Für den Bereich mathematisch-logischen und naturwissenschaftlich-funktionalen Wissens sowie für propositional-sprachliches Wissen ist eine solche Äquilibration hinlänglich belegt. Bislang fehlte jedoch ein geeignetes Verfahren, um den Prozess der Veränderung einer Wissensstruktur auch bei merkmalsorientiertem Faktenwissen messen zu können. Der Stresswert einer Nonmetrischen Multidimensionalen Skalierung ist – neben einer Reihe anderer Parameter – auch wesentlich durch die innere Konsistenz der zugrunde liegenden Datenmatrix beeinflusst. Aus diesem Grund kann er als Schätzwert für den Grad der Integriertheit des Wissens derjenigen Person dienen, von der diese Daten stammen. Zwei Lernexperimente, in denen die Versuchspersonen das hinzugewonnene Merkmalswissen jeweils bei Ähnlichkeitsurteilen zwischen den Objekten einzusetzen hatten, überprüfen anhand des NMDS-Stresswerts gezielt die Hypothesen, dass Assimilation mit einer Verringerung der Gesamtkonsistenz der Wissensstruktur einhergeht, während zeitlich darauf folgende Akkomodationsprozesse diese Konsistenz wiederherstellen. Die Resultate bestätigen für Assimilation und für Akkomodation eine Effektgröße, die den gegenläufigen Übungseffekt (gemessen bei nicht lernenden Kontrollgruppen) nicht nur ausgleicht, sondern signifikant übertrifft.

Schlüsselwörter: Assimilation, Akkomodation, Äquilibration, Wissenspsychologie, Sachwissen, Merkmalswissen, Wissenskarte, Kognitive Struktur, Kognitive Karte, Lernexperiment, Ähnlichkeitsurteil, Nonmetrische Multidimensionale Skalierung, Stresswert, Prokrustes-Transformation

8.1. Einleitung

Das Merkmalswissen einer Person über eine Anzahl von Objekten lässt sich mit so genannten „Kognitiven Karten“ darstellen (Läge, 2001, Marx & Hejj, 1989; Marx & Läge, 1995). Als dimensionale Gebilde zeigen sie die Relationen, die eine Person zwischen diesen Objekten sieht, in Form von kleineren oder größeren Distanzen auf und lassen als Gesamtbild damit Rückschlüsse auf die von der Person als wesentlich erachteten Merkmalsbereiche zu. Gleichzeitig erlauben Kognitive Karten, das Wissen über ein spezifisches Objekt durch deren Position in der Gesamtstruktur nachzuvollziehen.

Abbildung 60 gibt ein Beispiel für eine solche Wissenskarte. Sie wurde mittels Nonmetrischer Multidimensionaler Skalierung (NMDS) aus einer Matrix paarweiser Ähnlichkeitswerte generiert. In diesem Fall handelt es sich – um eine weitere Abbildung einzusparen – um die aus den Merkmalsdaten von Experiment 2 errechnete Karte, aber man kann sich als Einführungsbeispiel auch vorstellen, dass diese NMDS-Karte auf der Basis eines Satzes individueller Ähnlichkeitsurteile entstanden ist. Durch Property Fitting (multiple lineare Regression; vgl. Läge, 2001) ließen sich im Nachhinein die drei Dimensionen „Bruttosozialprodukt“, „Landesfläche“ und „Einwohnerzahl“ mit genügend guten Korrelation in die Karte einpassen, so dass im Fall einer individuellen Wissenskarte interpretiert werden dürfte, dass die befragte Person bei ihren Urteilen vor allem auf diese Merkmale zurückgegriffen haben muss. Außerdem lässt

³⁶ Teile dieser Arbeit wurden am 43. Kongress der Deutschen Gesellschaft für Psychologie als Poster vorgestellt: Streule, R., Schlatter, K., & Läge, D. (2002). Assimilation und Akkomodation beim Wissenserwerb – gezeigt anhand Kognitiver Karten. In E. van der Meer, H. Hagendorf, R. Beyer, F. Krüger, A. Nuthmann, & S. Schulz (Hrsg.), *43. Kongress der Deutschen Gesellschaft für Psychologie 2002*. Lengerich: Pabst. S. 426.

sich bereits aus der Relation zu den anderen Objekten für ein beliebiges Land, z.B. für Frankreich, sagen, dass die Person hierunter ein relativ großflächiges, einwohnerreiches Land mit hohem Bruttosozialprodukt versteht.

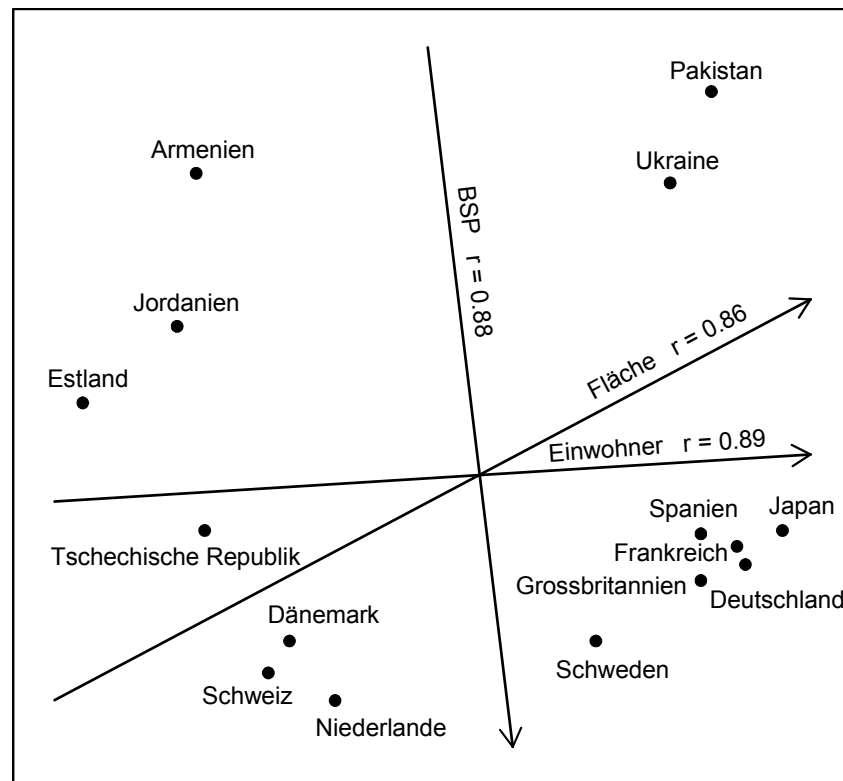


Abbildung 60: NMDS-Karte über 15 Länder und drei Merkmale, als Einführungsbeispiel und gleichzeitig als Zielstruktur des Lernens in Experiment 2

Die Genese derartiger Wissenskarten im Laufe von Lernprozessen ist bislang überraschend schlecht untersucht. Lediglich Roth (2000) macht einen leider aber isoliert bleibenden Vorschlag, die Veränderung von Wissen im Sinne der Piaget'schen Theorie mit artifiziellen Neuronalen Netzwerken abzubilden. Neuerdings werden allerdings normative Stimmen laut, Wissensvermittlung stärker an die bestehenden Kognitiven Strukturen anzupassen (z.B. Novak, 2002, Bezug nehmend auf Ausubel, 2000). Dazu wäre aber eine systematische Modellierung dieser Strukturen auf individuellem Niveau notwendig, um über die rein deskriptive Ebene hinausgehen zu können, die beispielsweise Venville & Treagust (1998) in ihrer multimedial sehr sorgfältig aufgebauten Studie über konzeptuelle Veränderungen während eines Genetik-kurses anbieten. Anwendungen Nonmetrischer Multidimensionaler Skalierung (oder verwandter Methoden) in der Psychologie, Soziologie oder Pädagogik beschränken sich bis anhin jedoch fast ausschließlich auf die Untersuchung der Sichtweise von Populationen und argumentieren deswegen auf der Basis von Proximitätenmatrizen, die über eine Stichprobe gemittelt wurden. Gelegentliche Einzelfallstudien, wie z.B. eine über das Dinosaurierwissen eines Kindes (Chi & Koeske, 1983) oder über die geographische Vorstellung eines Herrn B. (Gigerenzer, 1981), bilden die Ausnahme. Aber selbst sie zeichnen nicht einen Lernprozess nach, sondern begnügen sich mit einer Momentaufnahme in Form einer Statusdiagnose.

Dabei sind derartige Wissenskarten aufgrund ihres relationalen Charakters eigentlich geeignet, die schrittweise Verbesserung von objektbezogenem Sachwissen auf Strukturebene nachzuzeichnen und dabei – etwa in der Tradition Piaget'schen Denkens – in der Terminologie von Assimilations- und Akkomodationsleistungen zu beschreiben, denn diese beiden (ur-

sprünglich aus der Biologie entlehnten) Begriffe beschreiben ja kognitive Integrations- und Strukturierungsleistungen. Assimilation bedeutet Anpassung der Wirklichkeit oder der Umwelt an die eigene, aktuelle kognitive Organisation. Neue Information wird in bestehende Assimilationsschemata integriert. Akkomodation heißt Anpassung, Erweiterung oder Veränderung der kognitiven Schemata oder Strukturen in Richtung auf eine Angleichung an die Umweltanforderungen (für eine ausführliche Darstellung des Konzepts der kognitiven Adaptationen siehe Piaget, 1976).

Aus dieser Definition leitet Piaget seine zwei Grundpostulate ab, welche zentral für die Argumentationswege einer Gleichgewichtstheorie sind: „Erstes Postulat: Jedes Assimilationsschema hat die Tendenz zu wachsen, das heißt sich die Elemente einzuverleiben, die ihm äußerlich und mit seiner Natur verträglich sind. (...) Zweites Postulat: Jedes Assimilationsschema ist gezwungen, sich an die Elemente zu akkomodieren, die es assimiliert, das heißt sich entsprechend ihren Besonderheiten zu verändern, ohne deshalb seine Kontinuität (also seine Geschlossenheit als Zyklus interdependenter Prozesse) oder sein früheres Assimilationsvermögen zu verlieren...“ (Piaget, 1976, S. 14f.)

Das so entstehende Gleichgewicht (und auch die Probleme, die bei seiner Störung geschehen), sollte sich eigentlich in Wissenskarten gut detektieren lassen können, denn ihre anzunehmende Genese deckt sich praktisch eins zu eins mit den Definitionen und Postulaten Piagets: Eine Person wird neu gelernte Objekte so lange wie möglich in ihre bestehende Wissensstruktur integrieren. Erst wenn es durch die neuen Objekte (oder durch neues Merkmalswissen zu bereits bekannten Objekten) zu einem strukturellen Konflikt kommt, wird sie eine entsprechende Umstrukturierung ihrer kognitiven Wissensorganisation vornehmen. Dabei dürfte – ganz im Sinne des zweiten Postulates Piagets – die Kontinuität im Laufe des Prozesses erhalten bleiben.

Abbildung 61 zeigt ein Beispiel für einen Lerneffekt, der hauptsächlich auf einer Assimilationsleistung beruht: Fünf neue Länder sind hinzugekommen, welche in die bestehende Gesamtstruktur integriert werden konnten. Gleichzeitig ist diese Gesamtstruktur aber nicht ganz unverändert geblieben, sondern auch sie hat sich – um in Piagets Formulierung zu bleiben – „an die Elemente ... akkomodiert, die es assimiliert.“ Allerdings steht der Assimilationseffekt klar im Vordergrund, denn die Wissensstruktur hat ja keine nennenswerte qualitative Veränderung erfahren (der Abweichungswert aus der Prokrustes-Transformation ist gering), sondern wurde vielmehr erweitert.

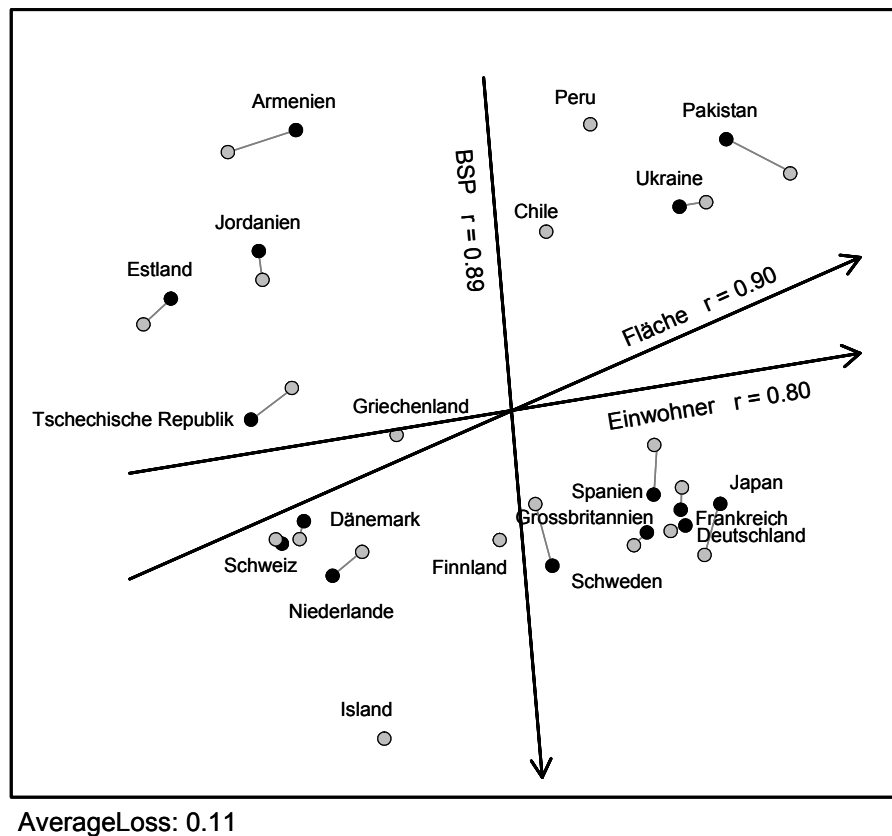


Abbildung 61: Erweiterung von Abbildung 1 um fünf weitere Länder (Chile, Finnland, Griechenland, Island, Peru). Die schwarzen Punkte stehen für die Normkarte aus Abbildung 1. Die grauen Punkte für die neu berechnete NMDS-Lösung mit 20 Objekten. Die Abweichung der Positionen mit 15 Objekten verglichen mit den Positionen aus der Karte mit 20 Objekten beträgt AvgLoss = 0.11.

Anders sieht dies aus, wenn zu den bereits bekannten Objekten andere Merkmalsinformation berücksichtigt wird. Deren Integration kann die aus den Ähnlichkeitsurteilen resultierende Wissenskarte gehörig auf den Kopf stellen. Abbildung 62 zeigt ein solches Beispiel, in dem die Merkmale durchschnittliche Jahrestemperatur und durchschnittliche jährliche Niederschlagsmenge die Merkmale Fläche und Einwohnerzahl ersetzen. In den Worten von Piaget's Theorie hat hier ein Umbau der kognitiven Struktur im Sinne einer Akkomodation stattgefunden: Die relationalen Beziehungen haben sich neu organisiert, und in der Folge verändern die schwarz markierten Objekte ihre Position im Vergleich zur vorausgehenden Karte. Die Person hat gewissermaßen ihr „Weltbild“ hinsichtlich der Geographie geändert, nachdem sich die neu gelernten Merkmale nicht mehr mit dem alten Schema in Einklang bringen ließen. Dass nicht ein noch größerer Umbau entstanden ist, hängt vor allem an der konstanten Verwendung des Merkmals Bruttosozialprodukt in beiden Karten.

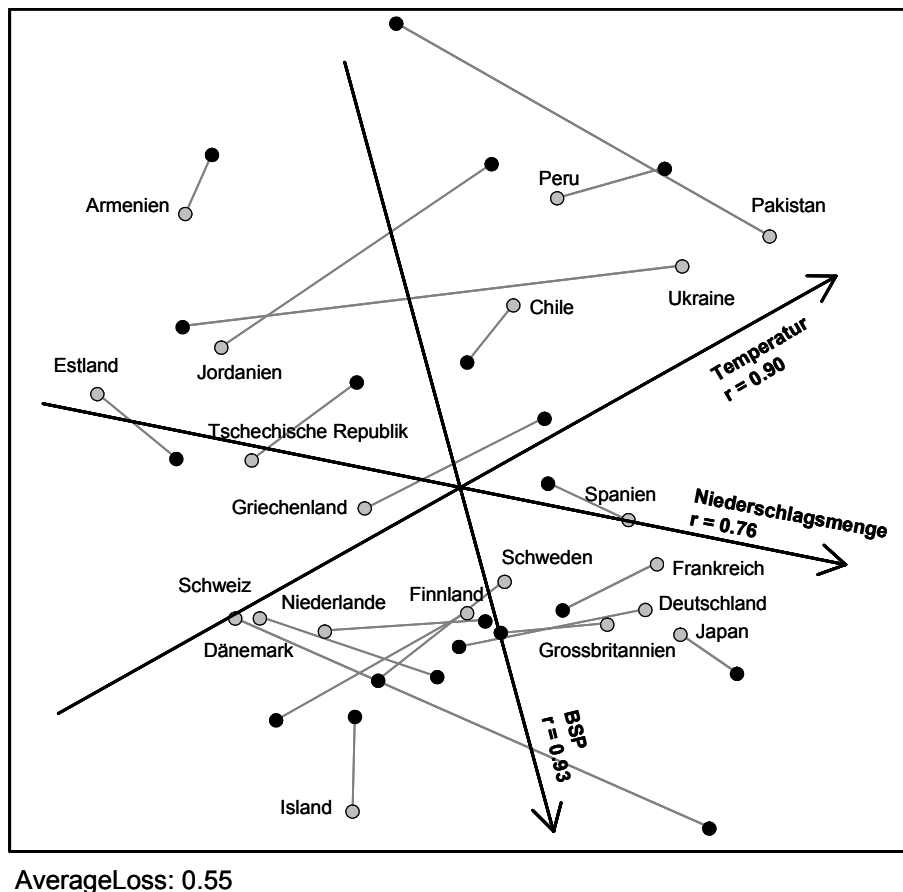


Abbildung 62: Veränderung der Karte im Sinne einer Akkomodation. Die grauen Punkte stehen für dieselben grauen Punkte in Abbildung 2. Die schwarze Punktekonfiguration basiert auf den Merkmalen BSP/Kopf, jährliche Durchschnittstemperatur und jährliche durchschnittliche Niederschlagsmenge. Ebenfalls sind die Regressionsgeraden für die schwarzen Punkte eingetragen. Deutlich zu erkennen ist der auf die unterschiedlichen Informationen rückführbare Umbau der Kartenstruktur mit einem AverageLoss von 0.55.

Die angeführten Beispiele sind idealtypisch und deswegen klar zu interpretieren. Sie zeigen anhand eines technischen Wertes (dem AverageLoss aus einer Prokrustes-Transformation) numerisch das Ausmaß der Veränderung der alten Struktur auf. Ist diese Veränderung für die ursprünglich bereits gekannten Objekte gering, so wurde das neue Wissen assimiliert, ist diese groß, so akkomodierte die Person. In der Realität gehen – wie Piaget betont – aber beide Prozesse Hand in Hand, und es dürfte schwierig sein, allein über einen (im Mittelbereich dann natürlich sehr unscharfen) Abweichungswert zwischen zwei Karten zwischen Akkomodation und Assimilation zu unterscheiden. Gleichwohl zeigen die Beispiele, dass es reizvoll wäre, für den Bereich solcher Wissenskarten ein geeignetes Messverfahren zu entwickeln. Gelingt das nämlich, dann könnte sich eine wesentliche Lücke schließen lassen, die bislang durch die starke Asymmetrie in der Erforschung regelbasierten Wissens auf der einen und merkmalsbasierten Sachwissens auf der anderen Seite besteht (Montada, 1995, liefert dazu einen hervorragenden Überblick). Diese Asymmetrie zugunsten des regelbasierten Wissens (für Felder wie Physik oder Mathematik liegt eine enorme Dichte aussagekräftiger Forschungsbefunde vor) dürfte zum einen dadurch geleitet sein, dass für die Messung physikalischer oder mathematischer Vorstellungen ein reichhaltiger Satz an Experimental- und Rechenaufgaben vorliegt, um das Wissen in jeder Altersgruppe zu untersuchen. Zum anderen mag aber auch eine Rolle spielen, dass das Faktenwissen bislang in einer stark an der traditionellen, regelbasierten K.I.-Forschung orientierten Kognitionpsychologie nur als unverbundenes „lexikalisches Wissen“ angesehen wurde. Strube (1996) fasst in seinem Handbuchartikel zu *Wissen & Wissensrepräsentation* diese Sichtweise prägnant zusammen, wenn er folgerichtig von „dem pri-

mitivsten Format zur Wissensrepräsentation“ spricht (S. 806). Betrachtet man Merkmalswissen hingegen als in einem Wissensfeld integriert, dann lässt sich auch das Lernen von „trivialen“ Fakten in seiner Integration in das Gesamtwissen einer Person nachzeichnen.

Das ist aber leichter gefordert als getan. Das Betrachten einer einzelnen Punktkonfiguration allein führt nicht weiter. Der Vergleich zeitlich aufeinander folgender Wissenskarten (z.B. durch Prokrustes-Transformation) könnte ein Ansatz sein. Doch dieser Vergleich einer Konfiguration mit dem vorherigen Wissensstand lässt assimilative oder akkomodative Aspekte wohl nur im Extremfall (wie in den beiden gewählten Beispielen) klar zu Tage treten. Eine Kognitive Karte besteht ja nur aus Relationen der Objekte zueinander, und so fehlt der archimedische Punkte, an dem sich der Unterschied zwischen assimilativer Einpassung von neuem relationalen Wissen in die alte Struktur und globaler Veränderung der alten Struktur durch Akkomodationsleistungen festmachen ließe.

Wir schlagen deswegen ein Maß für die innere Konsistenz abgegebener Ähnlichkeitsurteile vor. Ganz im Sinne Piagets gehen wir nämlich davon aus, dass neues Wissen solange assimiliert wird, solange dadurch das bestehende Weltbild als ganzes in sich stimmig bleibt. In seinem zweiten Postulat bezeichnet Piaget dieses jedoch als einen Grenzfall, wenn er davon spricht, dass ein Assimilationsschema immer auch zu akkomodieren habe. Assimilation als Prozess erzeugt damit also so etwas wie „Reibung“ beim schon Bestehenden: kleine Inkonsistenzen treten zu Tage, werden aber solange es geht toleriert. Akkomodation hingegen bewirkt das Auflösen dieser Inkonsistenzen durch Restrukturierung des Gesamtfeldes.

Beziehen wir diesen Gedanken auf Ähnlichkeitsurteile, so ergibt sich folgendes: Befragen wir eine Person unmittelbar nach einem Lerndurchgang, so sollte sie versuchen, die neu gelernte Merkmalsinformation zusätzlich zum bisherigen Wissen in jedes dieser Urteile einfließen zu lassen. In der Summe sollte das dann zu einer qualitativ verbesserten Wissenskarte führen (gegeben einmal, die gelernte Merkmalsinformation ist funktional und sachrichtig). Gleichzeitig sollte aber eine gewisse Beharrungstendenz hinsichtlich des Alten bestehen bleiben. Die einzelnen Ähnlichkeitsurteile werden dadurch zwar in der Summe qualitativ besser, verlieren jedoch an innerer Konsistenz. Assimilation als Integration des neuen Wissens in die alte Struktur erzeugt damit die oben beschriebene „Reibung“. Erst ein gegenläufiger Prozess, nämlich die Integration des alten in die strukturbildenden Fakten des neuen Wissens (auch so ließe sich „Akkomodation“ beschreiben) löst diese „Reibung“ auf. Ähnlichkeitsurteile sollten nach einem Akkomodationsprozess in sich stimmiger sein als vorher.

Wann findet Assimilation statt, wann Akkomodation? Laut Piaget wird ein Gleichgewicht hergestellt durch die Elaboration bestehender Strukturen und deren Vergleich mit den neu hinzugekommenen Informationen (Piaget, 1976, S. 167). Daraus folgt, dass diese regulierenden Kompensationen kognitive Ressourcen beanspruchen, welche abhängig von den entstehenden Widerständen/Störungen zwischen den zuvor neu erworbenen Informationen und den alten Strukturen unterschiedliche zeitliche Beanspruchungen erfordern. Zeitlich gesehen geht Assimilation der Akkomodation damit stets voraus: Zuerst wird eine neue Information an bestehende Strukturen angehängt; im Fall des Merkmalslernens werden die dem Wissensraum zugrunde liegenden Merkmalsdimensionen mit Werten belegt. Das Faktenwissen gewinnt somit an Qualität. Eine kognitive Struktur oder ein Schema in einem Gleichgewichtszustand zeichnet sich aber darüber hinaus durch eine in sich stabile Relationalität zwischen den Merkmalsausprägungen der Objekte aus. Eine Person kann zwar über objektiv richtiges Faktenwissen auf der Ebene von Einzelinformationen verfügen, es gelingt ihr aber nicht, die Beziehungen dieser Fakten zu anderen Inhalten konsistent zu strukturieren und in kongruente Beziehungen zu setzen. Dieser Zustand ist derjenige kurz nach dem Lernen von neuem Wissen. Die Neuaufnahme von Wissensinhalten erfordert immer eine Restrukturierung, eine Regulierung der Beziehungen zwischen den Objekten, da eben eine Erweiterung einer kognitiven Struktur nicht nur die alleinige Anbindung beinhaltet, sondern zudem ein Gesamtbild

konstruiert werden muss. Dieser akkomodative Aspekt der Verarbeitung setzt nach diesen Überlegungen später ein und erfordert vermutlich auch mehr Zeit.

Assimilation von Wissen bedeutet also Verringerung der Konsistenz bei gleichzeitiger qualitativer Verbesserung der Gesamtstruktur. Akkomodation ist der nachfolgende Prozess der Verringerung der Inkonsistenzen.

Ein Maß für die innere Konsistenz von Ähnlichkeitsurteilen ist deren geometrische Umsetzbarkeit in eine Karte gegebener Dimensionalität. Die Nonmetrische Multidimensionale Skalierung gibt für jede Lösung den Stresswert an, also denjenigen Anteil bezogen auf die Gesamtheit aller Distanzen, um den einzelne Distanzen in der Konfiguration hinsichtlich der gegebenen Ordnung in den Daten zu groß oder zu klein ausfallen. Abbildung 63 gibt ein Beispiel dafür, wie sich der Stresswert aus einzelnen Einpassungsfehlern zusammensetzt. Die grundsätzliche Idee, den Stresswert für jeden Punkt der Konfiguration zu betrachten, hat bereits Borg, 1978 formuliert). Dieser Stresswert ist ein wichtiges technisches Maß der NMDS, denn die iterativen Approximationsalgorithmen sind so gebaut, dass sie beim Finden der endgültigen Lösung diesen Abweichungswert zwischen Daten und Resultat minimieren (vgl. Borg & Groenen, 1997, für die Logik des Verfahrens und zur mathematischen Herleitung Mathar, 1997). Außerdem gibt der Stresswert auch einen Hinweis darauf, ob die skalierten Daten überhaupt Struktur enthalten (für eine Monte-Carlo-Studie zur zufallskritischen Absicherung des Stresswertes siehe Gigerenzer, 1981).

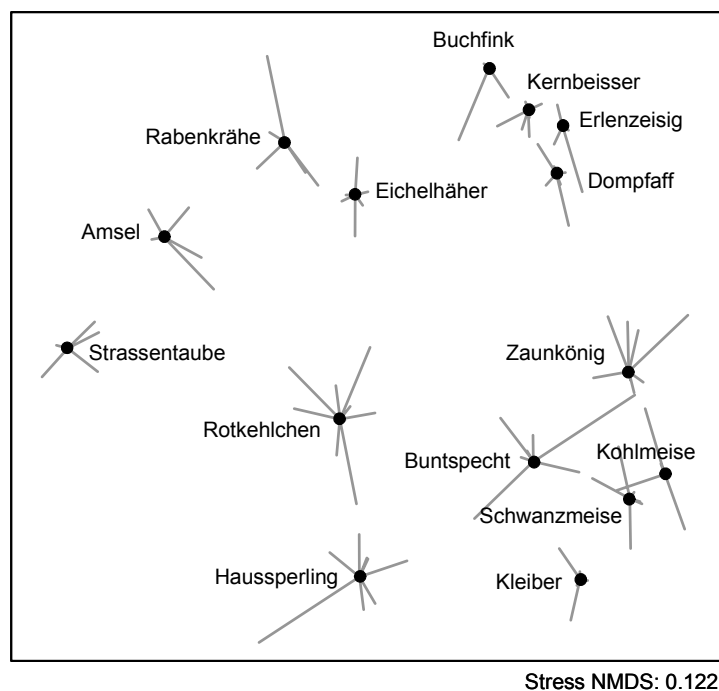


Abbildung 63: Eine individuelle Karte aus Experiment 1 mit eingezeichnetem Distanzstress (symmetrischer Stress pro Objekt = graue Striche) und Gesamtstress = 0.12.

In diesen Stresswert fließen eine Reihe von Einflussgrößen ein: Neben der Dimensionalität der Lösung und der Anzahl der zu skalierenden Objekte auch die Wahl der Ausgangskonfiguration, die Art der Behandlung von Rangbindungen in den Daten, das gewählte Verfahren zur Berechnung von Zieldistanzen, der gewählte Verschiebealgorithmus, die Stärke der resultierenden Clusterung der Lösung, die Konsistenz der Proximitäten für die „Eckpunkte“ der Lösung und vieles mehr. Die innere Konsistenz der zu skalierenden Daten ist dabei zwar ein wichtiger, aber eben nur ein Einflussfaktor unter vielen.

NMDS-Algorithmen sind normalerweise so gebaut, dass sie als einziges Ziel den Gesamtstress minimieren. Die Frage der inneren Konsistenz der Daten spielt dabei nur eine untergeordnete Rolle. Lediglich ein robuster Algorithmus würde versuchen, in den Daten das „Signal“ (das heißt den zueinander passenden Teil der Proximitäten) zu entdecken und möglichst fehlerfrei darzustellen. Was an Inkonsistenzen übrig bleibt, geht dann als „Rauschen“ (und bei groben Inkonsistenzen gegebenenfalls als Ausreißer) in den Stresswert ein. Mit RobuScal (Läge, 2001) liegt ein solch robuster Algorithmus vor. Er gibt Anlass zu der Hoffnung, dass über Stresswertunterschiede von NMDS-Lösungen auch eine verlässliche Aussage über die innere Konsistenz einer Proximitätenmatrix möglich ist (Läge, Daub, Bosia, Ryf & Jäger, in review). Angenommen, es gäbe einen perfekt robusten NMDS-Algorithmus, der es erlauben würde, Strukturanteil und Rauschen in den Ähnlichkeitsurteilen exakt auseinanderzudividieren (ein Wunschtraum!), dann würde ein assimilativer Prozess ein erhöhtes Rauschen und ein akkomodativer Prozess reduziertes Rauschen bedeuten. Leider kann man sich dieser Wunschvorstellung nach perfekter Zerlegung einer Ähnlichkeitsmatrix in Struktur und Rauschen mathematisch nur unvollkommen annähern, auch mit einem (hinsichtlich der anderen existierenden Algorithmen) so robusten Verfahren wie RobuScal. Trotzdem soll im Folgenden versucht werden, über den Stresswert der NMDS-Lösungen eine Schätzung für Akkomodation und Assimilation vorzunehmen.

Aus lerntheoretischer Sicht wäre die Verbindung der Überlegungen in der Denktradition Piagets mit der technischen Aspekten der Skalierung folgende: Unmittelbar nach einer Lerneinheit wird eine Person versuchen, das gerade erlernte Merkmalswissen in Verbindung zum bestehenden Wissen zu bringen, ohne das Bestehende stärker anzutasten als nötig. Sie hat zum Zeitpunkt der auf das Lernen folgenden Ähnlichkeitsurteile noch keine Gelegenheit gehabt, sich grundsätzlichere Gedanken über einen notwendigen Umbau der Gesamtstruktur zu machen. Deswegen sind unmittelbar nach dem Dazulernen von neuer Information zunächst einmal assimilative Prozesse zu erwarten. Diese erhöhen das Rauschen in der Ähnlichkeitsmatrix, was sich in einer robusten NMDS als erhöhter Stresswert auswirkt – selbst wenn sich inhaltlich die Lösung verbessert.

Nach einer gewissen Zeit (in den präsentierten Experimenten wird das jeweils eine Woche sein) hat sich das neue Wissen „gesetzt“ (und ein Teil davon wurde vielleicht auch vergessen). Wenn die Person jetzt Ähnlichkeitsurteile abgeben soll, dann hat sie weniger Mühe, altes und neues Wissen zusammenzubringen. Das integrierte neue Wissen wird sich konsistenter auf alle Ähnlichkeitsurteile auswirken, und das wird das Rauschen in der Urteilmatrix senken. Eine robuste NMDS produziert daraus eine Karte mit geringerem Stresswert, selbst wenn diese nicht unbedingt inhaltlich besser (d.h. näher an einem Normmodell gelegen) ist als die Karte vor einer Woche (ein Teil der Fakten wird ja vermutlich inzwischen vergessen sein). Das nach einer Woche resultierende geringere Rauschen würde für Umbau der Struktur in Richtung auf ein konsistentes Gesamtbild, also für Akkomodation sprechen.

Vorhergesagt wird also eine Erhöhung des Stresswertes unmittelbar nach dem Lernen (bei gleichzeitiger qualitativer Verbesserung der Wissenskarte) und eine Verringerung nach einer Woche Latenz. Wegen der vielen Faktoren, die auf den Stresswert einer NMDS-Lösung einwirken, ist hier die Wahrscheinlichkeit besonders groß, einen Beta-Fehler zu begehen, d.h. aufgrund der vielen intermittierenden Einflüsse auf den Stresswert einen Effekt von Assimilation und Akkomodation nicht zu finden, obwohl er eigentlich da wäre. Anders gewendet: Sollte sich der vorhergesagte Effekt finden lassen, dann ist er eher verborgen im Gestrüpp zu erwarten als auf dem Präsentierteller.

Gleichzeitig sollte man bei Vorhersagen hinsichtlich der Stresswerte auch berücksichtigen, was üblicherweise bei mehrfach hintereinander abgegebenen Ähnlichkeitsurteilen geschehen sollte (wenn nicht gelernt wird): Liegt wenig Zeit zwischen den Urteilen, dann sollten sich die herangezogenen Kriterien verfestigen. Zu erwarten ist also eine Art Übungseffekt, der mit ei-

ner Erhöhung der Urteilkonsistenz und somit mit einer Verringerung des Stresswertes einhergeht. (Qualitativ sollte sich eine Karte dadurch natürlich nicht verbessern, sie sollte nur konsistenter werden.) Liegt hingegen viel Zeit zwischen der Abgabe von Ähnlichkeitsurteilen, dann sollte sich dieser Übungseffekt nicht einstellen. Hat es ihn zuvor gegeben, dann lässt sich sogar erwarten, dass er nach einiger Zeit wieder ganz oder teilweise verschwindet.

Diese Annahmen hinsichtlich des Abgebens von Ähnlichkeitsurteilen bei gleich bleibendem Wissen kontrastieren scharf zu denen bei Assimilation und Akkomodation. Da sie genau gegenläufig sind, und da ein Übungseffekt im Abgeben von Ähnlichkeitsurteilen natürlich auch dann eintritt, wenn dazugelernt wird, könnte dieser Übungseffekt den erwarteten Effekt durch Assimilation und Akkomodation auch auslöschen. In jedem Fall muss letzterer Effekt größer sein als der Übungseffekt, um in den Resultaten sichtbar werden zu können.

Im Folgenden werden dazu zwei Lernexperimente präsentiert. Das erste basiert auf recht anspruchsvollem, von den Merkmalsinformationen her nicht so leicht zu integrierendem Wissen (Ornithologie). Im zweiten Experiment ist das Lernmaterial wesentlich einfacher gestrickt (drei intervallskalierte Merkmale zu 15 Ländern). In beiden Experimenten sollten sich für Assimilations- und Akkomodationsprozesse aber vergleichbare Befunde ergeben. Das Lernmaterial ist jeweils so konstruiert, dass auch eine Personengruppe ohne spezielles Lernprogramm mit ganz normalem Vorwissen bereits halbwegs vernünftig strukturierte Ähnlichkeitsurteile abzugeben in der Lage sein sollte (damit der angesprochene Übungseffekt gemessen werden kann).

8.2. Methode

8.2.1. Experiment 1

Aufbau des Experiments. In einer Längsschnittuntersuchung hatten die Versuchspersonen die Aufgabe, 15 Vogelarten anhand vier charakteristischer Merkmalsbereiche zu lernen. In drei Sitzungen, die jeweils im Wochentakt stattfanden, erhielten sie dazu Informationen zu diesen 15 Objekten. Mittels paarweiser Ähnlichkeitsurteile wurde dann die Auswirkung der Lerneinheiten auf die individuellen Wissenskarten der Versuchspersonen überprüft.

Die Versuchspersonen wurden vor dem Experiment zufällig in drei Gruppen aufgeteilt: zwei Experimentalgruppen mit unterschiedlichem Treatment (Exkursionslerner und Kärtchenlerner) und eine Kontrollgruppe. Die Kärtchenlerner erhielten in allen drei Lernphasen Kärtchen der 15 Vogelarten mit merkmalsstrukturierten Informationen. Die Exkursionslerner nahmen im ersten Lerndurchgang mit einem Ornithologieexperten an einer Exkursion teil, an der sie alle 15 Arten beobachten konnten. Die beiden folgenden Lernphasen unterschieden sich nicht von denen der Kärtchenlerner. Die Kontrollgruppe durchlief nur die Wissensdiagnose, nahm jedoch nicht an den Lerneinheiten teil. Der Aufbau des Experiments ist in Abbildung 64 skizziert.

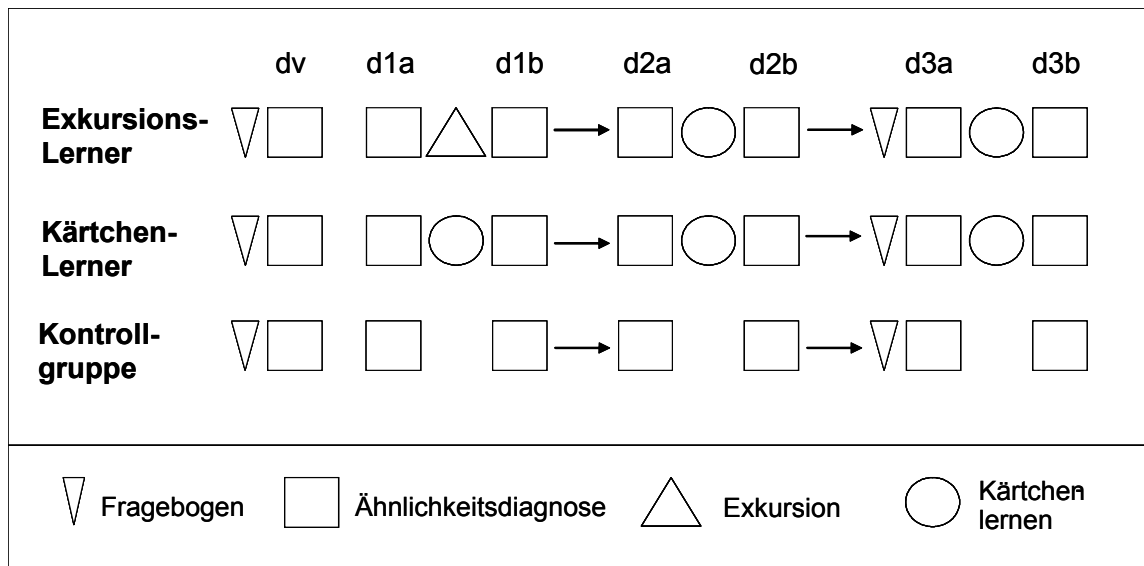


Abbildung 64: Schematische Darstellung des Ablaufs von Experiment 1. Die erste Ähnlichkeitsdiagnose in Experiment 1 wird im Resultateteil nicht berichtet, die zweite Diagnose mit d1a (Diagnose der ersten Sitzung vor dem Lernen, die dritte Diagnose mit d1b (= Diagnose der ersten Sitzung nach Lernen), die vierte Diagnose mit d2a (= Diagnose der zweiten Sitzung vor Lernen) usw. markiert sein.

Vor der ersten und vor der dritten Lernphase durchliefen alle Versuchspersonen einen merkmalsorientierten Sachwissenstest, indem ein Multiple-Choice-Fragebogen zu allen 15 Vogelarten die vier Merkmalsbereiche aus den Lerneinheiten testete. (Der zweite Test fand mit Absicht nicht unmittelbar nach einer Lerneinheit statt, um Kurzzeiteffekte zu vermeiden. Vor der dritten Lerneinheit lag eine Woche lernfreie Zeit, so dass nur der langfristige Erfolg aus den beiden ersten Durchgängen gemessen wurde.) Vor und nach jeder Lerneinheit wurde der aktuelle Wissensstand aller Versuchspersonen mittels paarweiser Ähnlichkeitsurteile auf der Strukturebene erhoben. Vor der ersten Lerneinheit gaben die Versuchspersonen den Satz von 105 Ähnlichkeitsurteilen zweimal ab. Damit lassen sich allfällige Übungseffekte beim Umgang mit diesem Verfahren zeigen (Der zweite der beiden Datensätze (d1a) wird dann bei der Auswertung als Messung des Vorwissens angesehen. Der erste Datensatz (dv) wird nicht berichtet, da keine Effekte aufgetreten sind). Die Kontrollgruppe durchlief dieselbe Wissensdiagnostik.

Auswahl der Vogelarten und Bestimmung der Zielstruktur. Die Auswahl der 15 Vogelarten, so wie sie sich im Expertenmodell (Abbildung 6) präsentiert, hatte einigen Kriterien zu folgen. (1) Für die Exkursion mussten alle Arten im Raum Zürich im selben Habitat leben und dort leicht zu finden sein, und zwar auch im Winter, als dieses Experiment durchgeführt wurde. (2) Einzelne Arten sollten nicht dermaßen von allen anderen Vögeln verschieden sein, dass auswertungstechnisch betrachtet die Gefahr degenerierter NMDS-Lösungen bestanden hätte (was zu unrealistisch niedrigen Stresswerten führen würde). Auch sollten die Arten nicht nur wenigen, scharf getrennten Verwandtschaftsgruppen angehören, um übermäßig kategoriale Lösungen zu vermeiden. Das war technisch wichtig, weil sich sonst ein möglicher Lernfortschritt nicht numerisch fair durch Prokrustes-Transformation von NMDS-Karten messen lassen würde. (3) Ca. sechs Arten sollten Laien (zumindest oberflächlich) bekannt sein, damit eine gewisse Grundstruktur vorhanden war. Mindestens sechs Arten sollten aber den Versuchspersonen völlig unbekannt sein, um ausreichenden Spielraum für das Lernen zu besitzen.

Die ersten beiden Restriktionen führten zu einer eingegrenzten Liste möglicher Arten. Diese wurden in einem Vorversuch 30 Studierenden der Universität Zürich in einem Fragebogen

vorgelegt, um Bekanntheitsgrad und Vorwissen einschätzen zu können. Daraus erfolgte dann die Auswahl der 15 am geeignetsten erscheinenden Arten.

Auf Strukturebene benötigen wir eine Wissensnorm, mit der jede Lernerkarte zu jedem Zeitpunkt evaluiert werden kann. Da diese Karte für die numerische Bestimmung des Lernerfolgs sehr wichtig ist, wurde große Mühe darauf verwendet, sie so akkurat wie möglich zu erheben. Deshalb wurden die 15 ausgewählten Arten acht erfahrenen Ornithologie-Experten vorgelegt, die jeweils die paarweisen Ähnlichkeiten einzuschätzen hatten. Bei vier von diesen Experten konvergierten die resultierenden NMDS-Karten in sehr hohem Masse, so dass ihre gemittelten Urteile mit einem Gewicht von 50% zur Grundlage für die Normkarte herangezogen wurden. Um sicherzustellen, dass das Normmodell auch wirklich die zu lernenden Merkmalsbereiche abdeckte, wurden die individuellen Ähnlichkeitsurteile nach dem letzten Lernen für jede der beiden Treatment-Gruppen gemittelt. Jede dieser beiden Ähnlichkeitsmatrizen floss mit einem Gewicht von 25% in die durch NMDS zu bildende Normkarte ein (Abbildung 65).

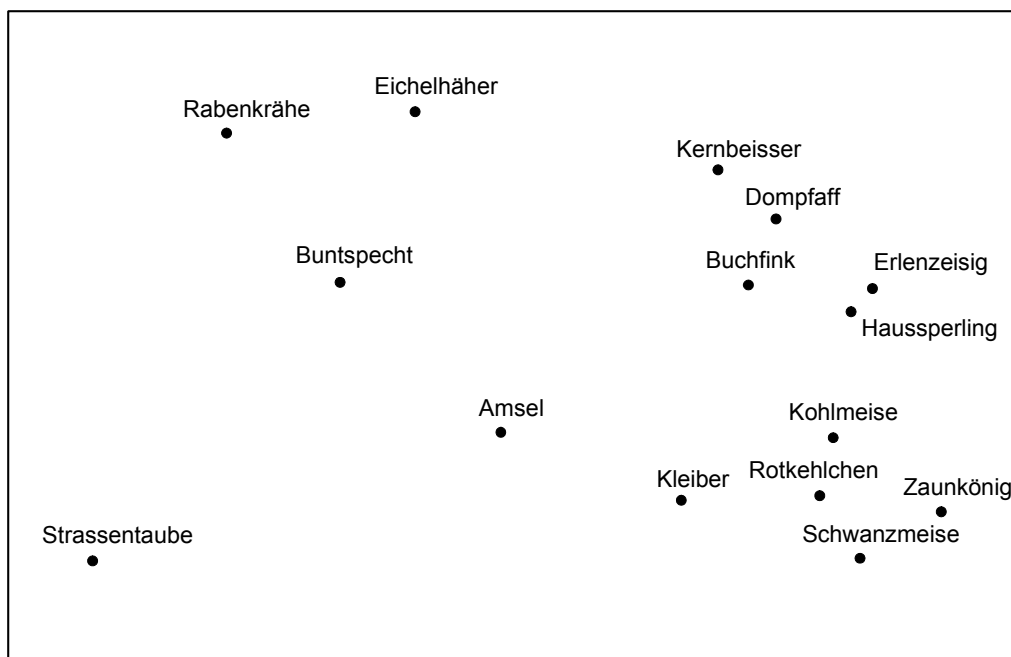


Abbildung 65: NMDS-Karte, die in Experiment 1 die Zielstruktur des Lernens ergibt.

Lernmaterialien. Die Exkursion fand in zwei Sechsergruppen statt und dauerte ca. eineinhalb Stunden. Die Lerner konnten alle 15 Arten unter Anleitung und begleitenden Kommentaren eines erfahrenen Exkursionsleiters per Fernglas beobachten. Zusätzlich zu seinen Auskünften hatten die Exkursionsteilnehmer die Lernkärtchen in der Hand und konnten sich bei Bedarf daran orientieren.

Diese Lernkarten setzten sich je Vogelart aus drei Teilen zusammen: aus a) einer farbigen Abbildung der Vogelart, aus b) einem kurzen, beschreibenden Text, der neben Besonderheiten und leicht eingänglichen Kuriositäten über die Art eine Merkmalsbeschreibung zu Lebensraum, Nahrung und Nistgewohnheiten enthielt (so wie man sie in Standard-Vogelbüchern findet), sowie aus c) einer tabellarischen Merkmalsliste, die alle wichtigen Merkmale aus dem beschreibenden Text tabellarisch aufführte.

In den individuell durchgeführten Lerndurchgängen erhielten die Versuchspersonen jeweils dieselben Lernmaterialien wieder vorgelegt. Sie waren in der Organisation ihres Lernens weitgehend frei, hatten sich aber in den beiden ersten Sitzungen jeweils 30 Minuten und in

der dritten Sitzung 20 Minuten damit zu beschäftigen. Als Kriterien wurde ihnen an die Hand gegeben, dass sie versuchen sollten, aufgrund der Abbildungen die Arten zu benennen und dass sie in der verbleibenden Zeit alle Kärtchen mindestens einmal durchgelesen haben sollten. Sie wussten, dass sie Ähnlichkeitsurteile über diese Arten abzugeben hätten und konnten bereits beim Lernen die Kärtchen strukturiert auf einem Tisch ausbreiten.

Versuchspersonen. An der Hauptuntersuchung nahmen 31 Studierende der Universität Zürich im Alter von 22–27 Jahren teil. Sie besaßen keine spezifischen ornithologischen Vorkenntnisse. In den beiden Experimentalgruppen befanden sich jeweils 4 Frauen und 8 Männer, die Kontrollgruppe bestand aus 5 Frauen und 4 Männern. Die Gruppenzuteilung erfolgte vor dem Experiment und zufällig.

8.2.2. Experiment 2

Aufbau des Experiments. In einer Längsschnittuntersuchung hatten die Versuchspersonen die Aufgabe, zu 15 Ländern die drei Merkmale Landesfläche, Einwohnerzahl und Bruttosozialprodukt so zu lernen, dass sie daraus eine Wissenskarte aufbauen konnten. In vier Sitzungen, die jeweils im Wochentakt stattfanden, erhielten sie dazu die relevanten Informationen zu diesen 15 Objekten. Mittels paarweiser Ähnlichkeitsurteile unmittelbar vor und unmittelbar nach der Lerneinheit wurde dann die Auswirkung der Lerneinheiten auf die individuellen Wissenskarten der Versuchspersonen überprüft. Zu Beginn jeder Sitzung erhob ein computerisierter Multiple-Choice-Fragebogen den aktuellen Wissensstand über die drei Merkmale zu jedem Land. In diesem Fragebogen hatten die Versuchspersonen die Länder einer Liste aller einzelnen Merkmalsausprägungen (plus Distraktoren) zuzuordnen.

Die Versuchspersonen wurden vor dem Experiment zufällig in eine Experimentalgruppe und eine Kontrollgruppe aufgeteilt. Die Versuchspersonen der Kontrollgruppe lernten nicht, sondern führten lediglich das komplette Programm der Wissensdiagnostik durch. Der Aufbau des Experiments ist in Abbildung 66 skizziert.

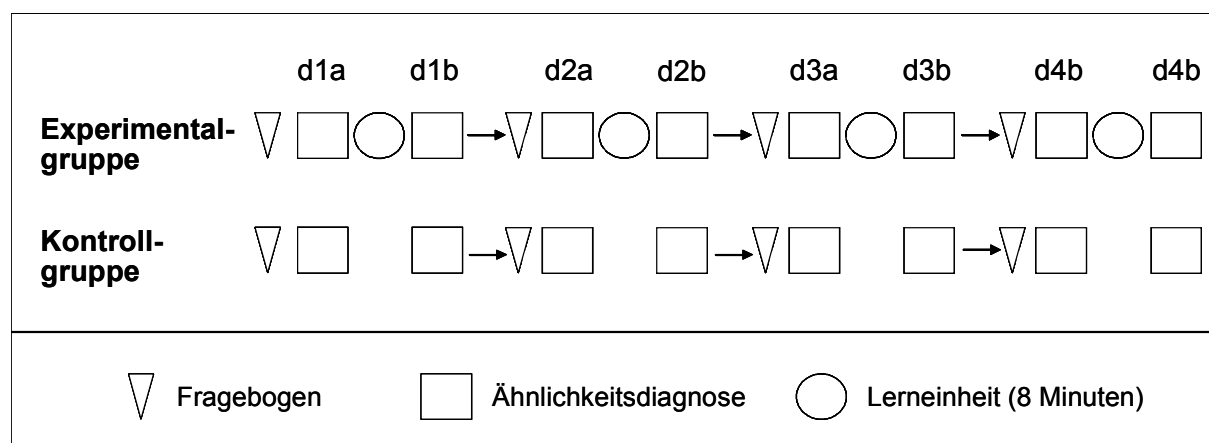


Abbildung 66: Schematische Darstellung des Ablaufs von Experiment 2

Auswahl der Länder und Bestimmung der Zielstruktur. In einem Vorversuch mit insgesamt 46 Versuchspersonen hatte sich aus einer Vorauswahl von 26 Ländern der Pool von 15 Ländern als gut geeignet herausgestellt, der zur Grundlage der Zielstruktur gemacht wurde. Außerdem hatten sich die drei intervallskalierten Merkmale Landesfläche, Einwohnerzahl und Bruttosozialprodukt pro Kopf als günstig für Lernen, Wissensdiagnose und Zielstruktur erwiesen. Die Zielstruktur errechnete sich über die Unähnlichkeiten der einzelnen Länder auf diesen drei Merkmalen. Dazu wurde folgende Formel verwendet:

$$U\check{A}_{(x,y)} = \frac{|(x_i - y_i) / (x_i + y_i)| + |(x_j - y_j) / (x_j + y_j)| + |(x_k - y_k) / (x_k + y_k)|}{3}$$

<i>Wobei:</i>	$U\check{A}_{(x,y)}$:	<i>Unähnlichkeitsquotient eines Länderpaares</i>
	x, y :	<i>Länderindex</i>
	i :	<i>Einwohnerzahl</i>
	j :	<i>Fläche in qkm</i>
	k :	<i>BSP pro Kopf in US\$</i>

Dieser Quotient $U\check{A}_{(x,y)}$ beschreibt den Betrag des prozentualen Unterschiedes je zweier Länder aufgrund ihrer Merkmalsausprägungen. Die resultierende Dreiecksmatrix wurde mittels NMDS (RobuScal-Algorithmus; Läge et al., in review) in eine zweidimensionale Karte transformiert, die für die Qualitätsabschätzung der individuellen Wissenskarten jeweils als Norm gilt.

Lernmaterialien. Das Lernmaterial bestand aus 15 Kärtchen, auf denen zu je einem Land die Angaben über Einwohnerzahl, Fläche und Bruttosozialprodukt standen. Die Versuchspersonen waren frei in ihrer Lernstrategie; sie konnten die Lernkarten ordnen und Gruppen bilden, ausbreiten usw. Eine Lerneinheit dauerte jeweils acht Minuten.

Versuchspersonen. 37 Personen (17 weiblich; 20 männlich) im Alter zwischen 19 und 33 Jahren nahmen an dem Experiment teil. Es handelte sich dabei hauptsächlich um Studentinnen und Studenten der Psychologie, Soziologie, Tiermedizin und Pharmazie der Universität Zürich. Hinzu kamen einige Versuchspersonen ohne Hochschulausbildung. 25 Personen wurden der Experimentalgruppe zugeordnet, 12 der Kontrollgruppe. Die Verteilung auf die beiden Stichproben erfolgte im Einzelfall zufällig, insgesamt gesehen aber so, dass die Variablen Alter, Geschlecht und Schulbildung in etwa gleichverteilt waren.

8.2.3. Auswertung der Daten und erwartete Befunde

Auswertung der Fragebögen: erwartete Verbesserung des Merkmalswissens. Voraussetzung für Assimilations- und Akkomodationsleistungen auf der Ebene der Wissenskarte ist die Verbesserung des Merkmalswissens. Deswegen werden zunächst die Fragebögen ausgewertet. Die Versuchspersonen der Lerngruppen sollten sich signifikant verbessert haben, die Versuchspersonen der Kontrollgruppen nicht.

Auswertung jeder einzelnen Ähnlichkeitsmatrix. Jede Ähnlichkeitsmatrix jeder Versuchsperson zu jedem Diagnosezeitpunkt wurde mit dem RobuScal-Algorithmus in eine zweidimensionale NMDS-Karte überführt. Diese Karte wurde dann mittels Prokrustes-Transformation auf das Normmodell für das jeweilige Wissensgebiet gelegt. Als numerisches Resultat für die Wissensgüte diente der AverageLoss, als Maß für die Konsistenz der Karte der Stresswert aus der NMDS.

Verlauf der Qualität der Wissenskarten. Für die Wissenskarten wird erwartet, dass die AverageLosses in der Prokrustes-Transformation jeweils nach einer Lerneinheit kleiner sind als in der Karte vor dieser Lerneinheit. Außerdem wird angenommen, dass in der Woche nach dem Lernen ein Teil des Zugewinns durch Vergessen wieder verloren geht. Die Wissenskarten der Personen aus den Kontrollgruppen sollten sich nicht systematisch verbessern. Am Ende des Experiments sollten die Versuchspersonen der Experimentalgruppen die klar besseren Karten aufweisen (Diese Hypothesen sind die minimal aufzustellenden, um von „Lernen“ auf Strukturebene sprechen zu können – der notwendigen Voraussetzung, dass es sich überhaupt lohnt, nach Assimilations- und Akkomodationsprozessen bei der Veränderung der Strukturen zu suchen. Eine wesentlich detailliertere Analyse der Verbesserung der individuellen Wissenskar-

ten beim Merkmalslernen folgt einer völlig anderen Fragestellung als der hier präsentierten und wird bei Läge, Streule & Schlatter (eingereicht) behandelt.).

Übungseffekt bei wiederholter Abgabe von Ähnlichkeitsurteilen. Für die Personen der Kontrollgruppen wird erwartet, dass sie bei wiederholter Abgabe von Ähnlichkeitsurteilen innerhalb einer Sitzung konsistenter werden, was sich in einem systematisch niedrigeren Stresswert der individuellen Wissenskarten äußert. Dieser Übungseffekt sollte in der einwöchigen Pause zwischen den Sitzungen ganz oder teilweise wieder verschwinden.

Erhöhung der Inkonsistenzen durch Assimilationsprozesse. Für die Versuchspersonen der Experimentalgruppen erwarten wir, dass die unmittelbar nach dem Lernen erhobenen Karten zwar inhaltlich besser, dafür strukturell aber inkonsistenter sind als vor dem Lernen. Sie sollten damit einen erhöhten Stresswert aufweisen.

Verringerung der Inkonsistenzen in den Lernpausen. In der Zeit zwischen zwei Sitzungen sollte sich das dazu gewonnene Wissen strukturell „setzen“. Wir erwarten, dass die Wissenskarten der Versuchspersonen in der Ähnlichkeitsdiagnose nach einer Woche Pause durch diese Akkomodationsprozesse niedrigere Stresswerte aufweisen als unmittelbar nach dem vorausgegangenen Lernen.

8.3. Resultate

Merkmalswissen wird gelernt. In Experiment 1 durchliefen alle Versuchspersonen vor der ersten und vor der dritten Lerneinheit einen allgemeinen Wissenstest. Zu allen 15 Vogelarten wurde dabei das Merkmalswissen über vier komplexe, mehrere Einzelmerkmale enthaltende Fragen in Form eines Multiple-Choice-Tests erhoben. Maximal konnten in dem Test 60 Punkte erzielt werden. Die Versuchspersonen der beiden Experimentalgruppen verbesserten sich von Test 1 zu Test 2 von durchschnittlich 32 auf 45.25 Punkte. Diese Verbesserung ist im t-Test auf dem 0.1%-Niveau signifikant ($t(23) = -11.19, p < 0.001$). Auch die Kontrollpersonen zeigen eine Zunahme der Punkte im Fragebogen (28 auf 30.78 Punkte), im Vergleich zu der Lernergruppe ist dieser Effekt jedoch gering und nicht signifikant ($t(8) = -2.04, p = 0.076$).

In Experiment 2 wurde das Merkmalswissen der Versuchspersonen zu Beginn jeder der vier Sitzungen mit einem Zuordnungs-Fragebogen geprüft. Dabei galt ein Datum als gewusst, wenn die Angabe innerhalb einer 20%-Marke des tatsächlichen Wertes gemacht wurde. In der ersten Sitzung erzielten die Personen der Lerngruppe durchschnittlich 8 Punkte (von 45 möglichen), die der Kontrollgruppe durchschnittlich 9 Punkte. Während die Kontrollgruppe auf diesem Level verblieb (im zweiten Fragebogen sank der Durchschnittswert sogar vorübergehend auf 7 Punkte), verbesserten sich die Personen der Experimentalgruppe fast linear über 13 und 19 auf 22 Punkte. Diese Verbesserung verteilt sich gleichmäßig auf die drei gelernten Merkmale und zeigte in einer zweifaktoriellen Varianzanalyse (ANOVA) mit den zwei Inter-subjektfaktoren Merkmal und Zeit signifikante Effekte des Merkmals, $F(1.78, 42.82) = 10.14$, $MSE = 86.30, p < .001$, und der Zeit, $F(2.0, 47.98) = 53.57$, $MSE = 457.44, p < .001$. Trotz dieser Verbesserung beherrschten die Versuchspersonen vor Beginn des vierten Lerndurchgangs aber erst ca. 50% des möglichen Gesamtwissens.

Die erzielten Punktzahlen zeigen deutlich an, dass die Personen der Experimentalgruppen tatsächlich an Merkmalswissen hinzugewonnen hatten, nicht aber die Kontrollpersonen. Damit ist die erste der zwei Grundvoraussetzungen für die Messung von Assimilations- und Akkomodationsleistungen in den beiden Experimenten gegeben.

Qualitative Verbesserung der Wissenskarten. Die zweite Voraussetzung besteht darin, dass sich der Zugewinn an Merkmalswissen auch auf die Wissenskarten auswirkt. Diese wird

durch eine Analyse aller AverageLosses zur Zielstruktur überprüft: Nach jeder der Lerneinheiten nähern sich die individuellen Wissenskarten unmittelbar der Zielstruktur an, so dass sie am Ende des Experiments 1 bei einem AverageLoss von durchschnittlich 0.47 und in Experiment 2 bei einem AverageLoss von 0.30 enden (Abbildung 67). Diese Verbesserung ist im t-Test in Experiment 1 auf dem 1%- und im Experiment 2 auf dem 0,1%-Niveau signifikant (Experiment 1: $t(23) = 3.22$, $p < 0.01$; Experiment 2: $t(24) = 4.15$, $p < 0.001$). Interessant ist auch, dass ein Teil der Verbesserung in den Wissenskarten in der Woche zwischen den Sitzungen jeweils wieder verloren geht. Teilweise war der im AverageLoss sichtbare Lerneffekt bereits so gering, dass davon zu Beginn der nächsten Sitzung nichts mehr übrig ist. Zumeist überwiegt aber der Lerneffekt deutlich den Vergessenseffekt und fällt vor allem nach dem ersten Lerndurchgang in beiden Experimenten am größten aus. Das Muster von Verbesserung der Wissenskarten durch Merkmalslernen und Verschlechterung durch eine Woche Pause wird im Binomialtest auf dem 5%- (Experiment 1) und dem 1%-Niveau (Experiment 2) signifikant.

Insgesamt zeigt sich ein Niveauunterschied der Wissenskarten zwischen den beiden Experimenten: Das zum Thema Ornithologie erweist sich als wesentlich anspruchsvoller, und die AverageLosses liegen konsequent um ca. 0.15 höher als im klarer strukturierten Länderexperiment. Auch gelingt es den Versuchspersonen der Kontrollgruppe im Experiment nicht, eine qualitativ anständige Karte zu produzieren. Zwar sind die gemessenen AverageLosses zufallskritisch gegen das Übereinanderlegen strukturfremder Karten abgesichert (der kritische Wert würde auf dem 1%-Niveau bei ca. 0.95 liegen), doch ist die Abweichung eines Wertes um die 0.70 doch beträchtlich. Insgesamt bildeten die Versuchspersonen der Kontrollgruppen in beiden Experimenten durchgängig schlechtere Karten und verbesserten sich zu keinem Zeitpunkt signifikant. Das mehrfache Abgeben von Ähnlichkeitsurteilen allein ergibt also keine Verbesserung in den Wissenskarten.

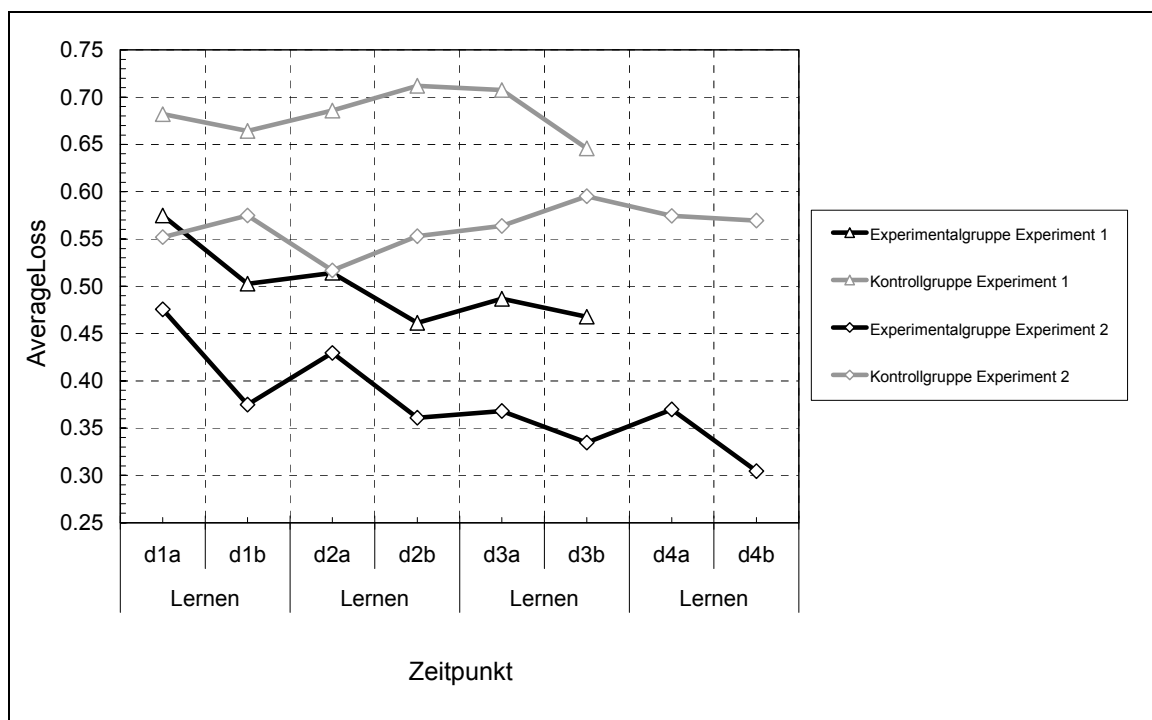


Abbildung 67: Gemittelte Abweichungen aller individuellen Wissenskarten von der Zielstruktur aus Experiment 1 und 2.

Übungseffekt bei wiederholter Abgabe von Ähnlichkeitsurteilen. Trotzdem machen die Kontrollpersonen nicht einfach „irgendetwas“. Dass sie sich bei den Ähnlichkeitsurteilen redlich Mühe geben, zeigt der jeweils erzielte Übungseffekt. Mit Ausnahme der vierten Lerneinheit in Experiment 2 verbessern sich die Versuchspersonen innerhalb einer Sitzung jeweils in der Konsistenz ihrer Urteile (Abbildung 68). Während der einwöchigen Pause geht dann (ebenfalls mit Ausnahme der letzten Sitzung von Experiment 2) der Konsistenzgewinn wie prognostiziert jeweils wieder zu einem guten Teil verloren. Das resultierende Muster entspricht – mit Ausnahme des in Experiment 2 zu niedrigen Mittelwertes für d4a – exakt den Vorhersagen. Zufallskritisch ist diese Übereinstimmung im Binomialtest auf dem 1%-Niveau als signifikant abgesichert (11 von 13 Vorhersagen korrekt bei einer Zufallserwartung von 50% pro Vorhersage). Auch hier erweist sich das Ornithologie-Experiment als das deutlich schwierigere der beiden. Nicht nur dass die Karten weiter vom Normmodell weg liegen, sie sind auch in sich um eine ganze Größenordnung inkonsistenter.

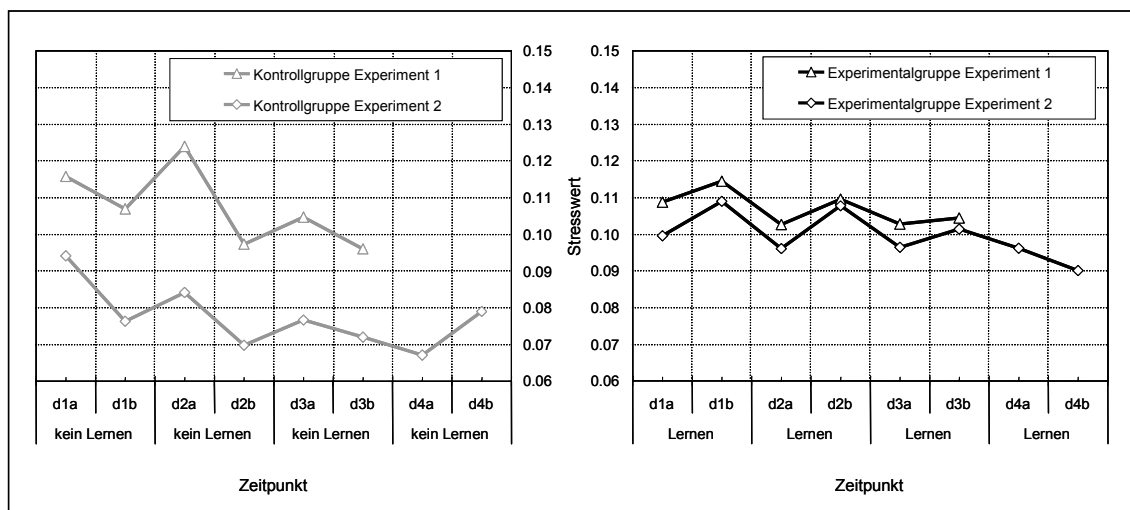


Abbildung 68: Stresswertverlauf der Kontrollgruppen (links) und der Experimentalgruppen (rechts) in Experiment 1 und 2.

Erhöhung der Inkonsistenzen durch Assimilationsprozesse. Für die Versuchspersonen der Experimentalgruppen haben wir genau gegenteilige Hypothesen hinsichtlich des Stresswertverlaufes aufgestellt. Ihre Karten sollten unmittelbar nach dem Lernen jeweils einen erhöhten Stresswert aufweisen. Das ist (wieder mit Ausnahme der vierten Sitzung (d4a, d4b) in Experiment 2) auch durchgängig der Fall (Abbildung 68).

Interessant an dem Vergleich der Kurven für die beiden Experimente ist, dass sie viel enger zusammen liegen als bei der Kontrollgruppe. Offensichtlich bietet das Lernmaterial aus Experiment 2 den Versuchspersonen die Chance, eine für sie konsistente Vorstellung über das Wissensgebiet aufzustellen (im Gegensatz zu den Kontrollpersonen, die ja nur auf ihr unzureichendes Vorwissen zurückgreifen können). Im Gegensatz dazu haben die Kontrollpersonen anscheinend in Experiment 2 einen Konsistenzvorteil. Sie können die Länder so strukturieren, wie sie das aufgrund ihres Vorwissens möchten, und müssen sich nicht mit der Aufgabe herumschlagen, das Lernmaterial zu integrieren. In der Summe ergibt sich dann der große Unterschied bei der Kontrollgruppe, aber nur ein minimaler Unterschied bei den Lernern.

Verringerung der Inkonsistenzen in den Lernpausen. Ebenfalls aus Abbildung 68 zu entnehmen ist die jeweilige Verbesserung der Konsistenz der Karten in der Woche zwischen den Sitzungen. Dies deckt sich mit der Annahme eine Gesamtintegration des hinzugewonnenen Wissens durch akkomodative Prozesse. Wie bereits bei den Personen der Kontrollgruppe, ist

auch die Übereinstimmung des Datenverlaufs in den Experimentalgruppen (Abbildung 68) im Binomialtest auf dem 1%-Niveau abgesichert (12 von 13 Vorhersagen korrekt bei einer Zufallserwartung von 50% pro Vorhersage).

Stichprobe	Effekt	p(t)	d
Experiment 1 (Experimentalgruppe)	Assim	n.s.	0.14
	Akkom	s.	0.27
Experiment 2 (Experimentalgruppe)	Assim	s.	0.25
	Akkom	s.	0.35
Experiment 1 & 2 (Experimentalgruppe)	Assim	s.	0.20
	Akkom	s.	0.32
Experiment 1 (Kontrollgruppe)	Übung	s.	0.42
	Vergessen	n.s.	0.34
Experiment 2 (Kontrollgruppe)	Übung	s.	0.28
	Vergessen	n.s.	0.15
Experiment 1 & 2 (Kontrollgruppe)	Übung	s.	0.31
	Vergessen	s.	0.20

Tabelle 4: Überblick über Assimilations- und Akkomodationseffekte (Experimentalgruppe) und Übungs- und Vergessenseffekte (Kontrollgruppe) im Stresswert. Signifikanz der Mittelwertsvergleiche ($\alpha = 0.05$) = p(t); Effektstärke = d.

Bei der Betrachtung der einzelnen Vergleiche zwischen den Zeitpunkten, aufgeteilt nach Assimilation (Vergleich der Stresswerte bevor und nach dem Lernen) und Akkomodation (Vergleich der Stresswerte direkt nach dem Lernen und nach der Lernpause) zeigt sich dasselbe Bild (Tabelle 4). Bei der (den) Experimentalgruppe(n) wird bis auf eine Ausnahme jeder Einzelvergleich signifikant, wenngleich sich die Effektgrößen auf dem Niveau eines kleinen ($d = 0.2$) bis mittleren Effektes ($d = 0.5$). Bei der Kontrollgruppe wirkt sich die Übung positiv auf die interne Konsistenz der Urteile aus.

Schätzung der Absolutgröße des Assimilationseffektes und des Akkomodationseffektes. Die beiden in der Mittelwertstendenz aufzufindenden Effekte (die längst nicht für die Stresswerte jeder individuellen Versuchsperson gelten!) geben den prognostizierten Hinweis auf Assimilation und Akkomodation. Einleitend hatten wir festgestellt, dass die Erwartungslinie der Stresswerte in diesem Experiment allerdings nicht das Beibehalten des Stresswertes ist, sondern der Übungseffekt bzw. seine Auflösung bei Personen, die nur Ähnlichkeitsurteile abgeben, aber nicht lernen. Aus diesem Grund präsentiert Abbildung 69 eine Schätzung des „Netto-Effektes“, indem jeweils die zu überwindende gegenläufige Tendenz hinzuaddiert wird. Dabei ist allerdings in der Interpretation die gleiche Vorsicht angeraten wie beispielsweise bei hirnpfysiologischen Messungen (wo ja ebenfalls eine entsprechende Differenzbildung zum Aufdecken von Effekten betrieben wird), weil Stresswerte zwar im Prinzip intervallskaliert sind, aber nicht mit Sicherheit eine vollständig lineare Verlaufsfunktion über eine beliebige Anzahl erzeugter Stresswerte haben (es gibt unseres Wissens dazu keine systematischen Berechnungen oder Simulationen). Trotzdem kann man sich den Differenzeffekt ja einmal anschauen, denn er ergibt doch beachtliche Werte (mit Ausnahme für die vierte Sitzung in Experiment 2, in der die Versuchspersonen beider Gruppen, wie wir gesehen haben, systematisch vom ansonsten durchgängigen Muster abweichen).

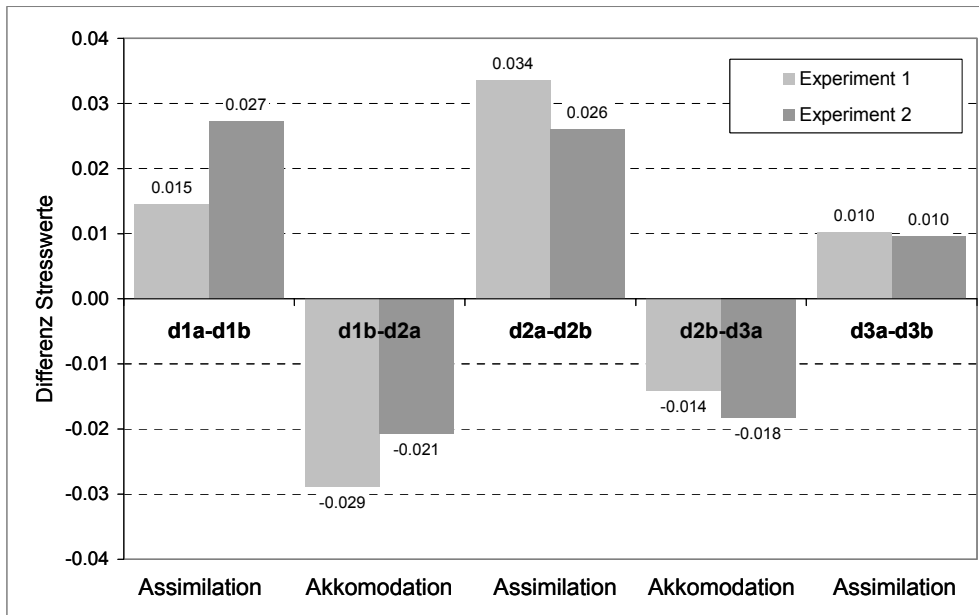


Abbildung 69: „Netto-Effekte“ der Assimilation und Akkomodation für Experiment 1 und 2 über die Zeit (Für Experiment 2 fehlt der letzte Akkomodations- (d3b-d4a) und der letzte Assimilationseffekt (d4a-d4b) aufgrund der fehlenden Berechnungsberechtigung; nähere Erläuterungen im Text). Absolute Assimilations- und Akkomodationseffekte wurden von Übungs- und Vergessenseffekten bereinigt.

8.4. Diskussion

Die Resultate der beiden Experimente zeigen einen mit den Vorhersagen erstaunlich gut übereinstimmenden Verlauf, und zwar sowohl in den Lernergruppen als auch in den Kontrollgruppen. Das spricht für einen systematischen Einfluss der Lerneinheiten auf die innere Konsistenz der Ähnlichkeitsurteile, und zwar ganz in der Richtung, wie dies durch die Prozesse von Assimilation und Akkomodation theoretisch beschrieben wird (bei gleichzeitig feststellbaren Lern- und Vergessenseffekten im Vergleich zu einer Wissensnorm!). Der eingesetzte robuste NMDS-Algorithmus scheint also die gewünschte „Filterfunktion“ zu erbringen, denn es bleibt jeweils das vorhergesagte unterschiedliche Ausmaß an Stress als nicht-skalierbares Rauschen der Daten in diesem „Filter“ hängen. Damit konnte dieses technische Maß zum ersten Mal für die psychologische Forschung auch inhaltlich sinnvoll eingesetzt werden.

Die einzige Unregelmäßigkeit im Muster des Stresswert-Verlaufs tauchte im letzten Lerndurchgang von Experiment 2 auf. Hier erwies sich der mittlere Stresswert nach dem Lernen als ungewöhnlich niedrig, was den erwarteten Effekt konterkarierte. Um die Stabilität respektive die Zufälligkeit des Stresswertverlaufs zu testen, haben wir die Lernergruppe von Experiment 2 in zwei Teilstichproben zerlegt. Beide Teilstichproben zeigen exakt einen parallelen Verlauf der Stresswerte über alle Zeitpunkte hinweg, inklusive eben dem letzten. Hier muss also etwas Systematisches im Aufbau des Experiments passiert sein, was nicht nur auf einige wenige Personen zurückzuführen sein kann (zumal ja auch die Resultate der Kontrollgruppe in diesem Durchgang genau gegen das sonst durchgängige Befundmuster laufen). Spekulieren könnte man über motivationale Probleme mit einem doch zum Schluss recht repetitiv werdenden Experiment. Aber dagegen spricht, dass die AverageLosses für die Experimentalgruppe einen sehr ordentlichen Lerneffekt aufweisen. Es gibt eben Resultate, die sich einer vernünftigen Interpretation zu verschließen scheinen.

Für die Gesamtlage der Befunde kommt den Kontrollgruppen eine wichtige Funktion zu: *Ers-*
tens zeigen sie, dass der „normale“ Prozess der Konsistenz von Ähnlichkeitsurteilen gerade
umgekehrt ist: Die Wiederholung innerhalb einer Sitzung fördert die Konsistenz, die Woche
Pause nimmt Konsistenz weg. Damit ist *zweitens* aber ein additiver Effekt hinsichtlich des
Lernens gegeben: Die Stresswertverläufe der Experimentalgruppen wiegen nicht nur den
normalen Übungseffekt auf, sondern sie übertreffen ihn sogar in einem signifikant werdenden
Masse. Wenn man will, dann kann man die Beträge (Lernergruppe + Kontrollgruppe) also ru-
hig addieren, um eine Schätzung zu haben, wie stark sich das Lernen bzw. die einwöchige
Pause danach auf Assimilation und Akkomodation auswirkt. Allerdings sind die vorgelegten
Befunde zunächst einmal eher als Resultate zu verstehen, die die Existenz eines systemati-
schen Stresswert-Effekts als solches belegen. Sie können keine endgültige quantitative Be-
zugsgröße liefern und sollen das auch gar nicht. Die Berechnung eines „Netto-Effekts“ ist
deswegen lediglich als ein Verfahrensvorschlag einzuordnen, um den kontrastierenden Ü-
bungseffekt korrekterweise mit ins Kalkül einzubeziehen. Hier wäre noch weitere systemati-
sche Forschung notwendig, welche die Größe von Assimilations- und Akkomodationseffekten
als unabhängige Variable definiert und dann den NMDS-Stresswert als abhängige Variable
quantitativ auslotet.

Die hier vorgelegten Befunde wurden auch lediglich mit einem statistisch eher schwachen
Test statistisch abgesichert: Die aus der Theorie abgeleiteten Prognosen hinsichtlich Vergrö-
ßerung und Verringerung hätten unter Zufallsbedingungen eine Eintretenswahrscheinlichkeit
von je 50% gehabt, so dass sich dieses Prognoseschema mittels Binomialtest auf dem 1%-
Niveau vom Zufall abheben ließ. Allerdings liegen die Mittelwerts-Unterschiede zwischen
zwei aufeinander folgenden Zeitpunkten liegen jeweils unter einer Standardabweichung der
beiden Verteilungen. Die Effektstärke ist also wie erwartet gering bis mittel; der Stresswert
wird halt durch die Summe der anderen Faktoren stärker beeinflusst als durch Akkomodation
und Assimilation.

Nun ist, wie in der Einleitung bereits ausgeführt, der Stresswert in allererster Linie ein inter-
nes technisches Maß für den Skalierungsprozess. Seine Höhe in einer NMDS-Lösung ist
durch die Restriktionen bei der iterativen ordinalen Darstellung der Proximitäten zwischen n
Objekten in einem niedrigdimensionalen Raum bedingt. Deswegen darf nicht damit gerechnet
werden, dass sich auf Einzelfallebene individueller Karten bereits vorhersagbare Unterschiede
zeigen. Insofern ist der Stresswert als psychologisches Maß vielleicht am besten mit Reakti-
onszeiten oder mit hirnpfysiologischen Massen zu vergleichen: Auch hier sind die einzelnen
Messungen von vielen Gegebenheiten abhängig, von denen die unabhängige Variable des Ex-
periments in der Regel nur einen geringen Teil ausmacht. Ein einzelnes Messdatum allein sagt
deswegen herzlich wenig aus und wird von der Forschung auch gar nicht beachtet. Erst wenn
man viele Messungen übereinander legt, zeigt sich in den Mittelwerten ein Effekt im Sinne
eines in seiner Größe stabilen und statistisch signifikanten Unterschiedes zu einer Ver-
gleichsmessung.

Obwohl in der Psychologie solche Form von Messqualität üblich ist, wäre an dieser Stelle ei-
ne Art „medizinischer Test“ natürlich wünschenswerter. Ein Test, der mit klar definierter
Sensitivität und Spezifität für den Einzelfall eine Diagnose über Assimilation und Akkomoda-
tion zuließe, könnte nämlich Aussagen über den individuellen kognitiven Prozess einer ler-
nenden Person machen. Die vorgelegten Ergebnisse müssen jedoch Zweifel daran aufkom-
men lassen, ob es möglich sein wird, aus der Konsistenz von Urteilen einen solchen Einzel-

fall-Test abzuleiten. Mit dem NMDS-Stresswert scheint uns das Limit auf Stichprobenebene bereits erreicht zu sein³⁷.

Gleichwohl öffnet dieses Maß den Zugang zur Abschätzung von Assimilations- und Akkomodationsprozessen auf einer neuen Ebene, auf der des objektbezogenen Faktenwissens nämlich. Damit wird die bisherige Forschung, die sich hauptsächlich auf a) die Entwicklung mathematisch-naturwissenschaftlichen Wissens (für einen Überblick bzw. für instruktive Beispiele s. Inhelder, De Caprona, & Cornu-Wells, 1987; Inhelder, Sinclair, & Bovet, 1974; Sodian, 1995; Venville & Treagust, 1998; Vosniadou, 1991), auf b) die Modellierung und Veränderung propositionalen Sprachwissens (Aebli, 1976; Steiner, 1996) oder auf c) die Beschreibung der Genese sozialer Kognition (Eckensberger & Silbereisen, 1980; Freitag, 1984; Furth, 1980;), Interaktionskompetenz (Bardis, 1978, 1979; Edelstein & Habermas, 1984) und Integration (Chiswick, 2002; Demorgon & Molz, 1996; Ghuman, 1994; Kao & Tienda, 1995; Mayades & Elliott, 1992; Straub, 1999; Wang, 1991; Weinberg, 2001) bezog, um ein viertes Feld erweitert, auf das sich das Konzept Piagets ebenfalls sehr gut anwenden lässt.

Inhaltlich in jedem Fall interessant sind die Interpretationsperspektiven der hiermit bereits vorgelegten Resultate hinsichtlich der Wissensintegration beim Lernen von Sachwissen. In Verbindung mit der qualitativen Verbesserung der individuellen Wissenskarte zeigt der Assimilationseffekt direkt nach dem Lernen, dass die Personen die hinzugewonnenen Fakten zwar gewinnbringend einsetzen können, dass dieser Einsatz aber noch nicht in sich stimmig ist. Nachdem sie sich dann eine Woche nicht bewusst mit dieser Materie befasst haben, ist ein Teil des erworbenen Wissens wieder vergessen (qualitativ verschlechtern sich die Wissenskarten im Vergleich zum Normmodell wieder). Aber das, was die Personen behalten haben, ist jetzt als Gedächtnisspur offenbar so gut mit dem zuvor bereits bestehenden Wissen verschmolzen, dass die Gesamtkonsistenz der abgegebenen Ähnlichkeitsurteile merklich ansteigt. Hier bietet sich auch von gedächtnispsychologischer Seite her ein lohnenswert erscheinender Ansatzpunkt, die Integrationsprozesse des Faktenwissens noch systematischer unter die Lupe zu nehmen als das in den zwei in dieser Arbeit präsentierten Experimenten möglich war.

Ein Festhalten an den Ähnlichkeitsurteilen als Instrument der Datenerhebung scheint dazu nicht der schlechteste Weg zu sein. Klauer (1989) hat das Bilden solcher Urteile als induktives Denken beschrieben und – im Duktus der von ihm herausgearbeiteten Transfereffekte – das Feststellen von Gleichheit und Verschiedenheit als Ausgangspunkt für das Entdecken von Regelmäßigkeiten benannt. Für den experimentellen Bereich ist dazu dieser Konnekt zu den Transferleistungen nützlich, aber nicht einmal unbedingt notwendig. Es reicht aus, mit Klauer festzustellen, dass bei Ähnlichkeitsurteilen „... nicht Objekte als Ganze und global, sondern Merkmale von Objekten beziehungsweise Relationen zwischen Objekten miteinander verglichen werden.“ (S. 187). Mit einem gesunden Schuss Optimismus könnte man demnach das

³⁷ Eine Alternative könnte man darin sehen, auf Matricebene die ordinale Transitivität zu überprüfen. Diese ist aber – abgesehen von Spezialfällen – prinzipiell gegeben. Und die Spezialfälle allein eröffnen kein ausreichendes Feld für eine numerische Auswertung. Natürlich muss eine Person, um konsistent zu sein, in ihren Urteilen auch transitiv sein: Wenn die Ähnlichkeit zwischen A und B größer ist als zwischen C und D und die Ähnlichkeit zwischen C und D größer ist als zwischen E und F, dann muss auch die Ähnlichkeit zwischen A und B größer sein als zwischen E und F. Das ist aber immer der Fall, weil die Vp ja paarweise Urteile abgibt und nicht Paare miteinander vergleicht. Für einen Transitivitätsvergleich müsste eine Vp also immer Ähnlichkeitsurteile zwischen zwei Paaren von Objekten abgeben. Aber das sprengt kombinatorisch sofort die Grenze des praktisch Machbaren, weil bei 15 Objekten bereits mehr als 5000 Paare von Paaren möglich sind. Kaufmann (1999) und Läge (2001) legen einen Kurztest zur Abschätzung von Transitivität bei Paaren von Paaren vor, der allerdings für die feinen Unterschiede, so wie sie zwischen assimilativen und akkomodativen Prozessen zu erwarten sind, viel zu grob misst.

Ähnlichkeitsurteil als eine Art „Königsweg zur Relationalität des Merkmalswissens“ bezeichnen.

Die mittels solcher Ähnlichkeitsurteile vorgelegten Resultate zeigen denn auch auf, dass die relational strukturierenden Prozesse von Assimilation und Akkomodation – so wie dies von Piaget angenommen wird – zwei (zeitlich voneinander verschiedene) Teilprozesse einer Gesamtfunktion sind, nämlich der, ein neues Gleichgewicht zwischen alten und neuen Wissens-elementen zu bilden. Vor diesem Hintergrund muss darauf verwiesen werden, dass die differentialdiagnostische Begriffsverwendung „assimilativer“ oder „akkomodativer“ Lernstile von Personen (z.B. Lynch, Woelfl, Steele, & Hanssen, 1998) maximal metaphorisch gesehen werden sollte. Unter gedächtnispsychologischer Perspektive hängen beide sehr eng miteinander zusammen, weil sie sich unwillkürlich, aber zeitversetzt aus demselben Lernen ergeben. Insofern bestätigen die Befunde Piagets Theorie der Äquilibration auf einer Ebene, auf der bislang wenig nach Struktur gesucht wurde. Das lässt eine weitere Erforschung der kognitiven Prozesse beim Lernen von Faktenwissen als lohnend erscheinen.

8.5. Literaturverzeichnis

- Aebli, H. (1976). *Psychologische Didaktik. Didaktische Auswertungen der Psychologie von Jean Piaget*. Stuttgart: Klett.
- Ausubel, D. P. (2000). *The acquisition and retention of knowledge*. Dordrecht: Kluwer Academic Publishers.
- Bardis, P. D. (1978). Social interaction and social processes. *International Behavioural Scientist*, 10, 7-32.
- Bardis, P. D. (1979). Social interaction and social processes. *Social Science*, 54 (3), 147-167.
- Borg, I. (1978). PAL: Point-wise alienation coefficients in multidimensional scaling. *Journal of Marketing Research*, 15, 478-479.
- Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling – Theory and applications*. New York: Springer.
- Chi, M. T. H., & Koeske, R. D. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology*, 19, 29-39.
- Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2002). *Longitudinal analysis of immigrant occupational mobility: a test of the immigrant assimilation hypothesis*. Bonn: IZA.
- Demorgon, J., & Molz, M. (1996). Bedingungen und Auswirkungen der Analyse von Kultur(en) und interkulturellen Interaktionen. In A. Thomas (Hrsg.), *Psychologie interkulturellen Handelns*. Göttingen: Hogrefe. S. 43-86.
- Eckensberger, L. H., & Silbereisen, R. K. (Hrsg.). (1980). *Entwicklung sozialer Kognitionen: Modelle, Theorien, Methoden, Anwendung*. Stuttgart: Klett-Cotta.
- Edelstein, W., & Habermas, J. (Hrsg.). (1984). *Soziale Interaktion und soziales Verstehen. Beiträge zur Entwicklung der Interaktionskompetenz*. Frankfurt/M.: Suhrkamp.
- Freitag, B. (1984). *Der Aufbau kindlicher Bewusstseinsstrukturen im gesellschaftlichen Kontext*. München: Fink.
- Furth, H. G. (1980). *The world of grown-ups. Childrens conceptions of society* (Chapt. 9: Piagets theory of social relation). New York: Elsevier.
- Ghuman, P. A. S. (1994). Canadian or Indo-Canadian: A study of South Asian adolescents. *International Journal of Adolescence and Youth*, 4, 229-243.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhardt.

- Inhelder, B., De Caprona, D., & Cornu-Wells, A. (Eds.). (1987). *Piaget Today*. Sussex: Erlbaum.
- Inhelder, B., Sinclair, H., & Bovet, M. (1974). *Apprentissage et structures de la connaissance*. Paris: Presses Universitaires de France.
- Kao, G., & Tienda, M. (1995). Optimism and achievement: the educational performance of immigrant youth. *Social Science Quarterly*, 76, 1-19.
- Klauer, K. J. (1989). Allgemeine oder bereichsspezifische Transfereffekte eines Denktrainings. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 21, 185-200.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Läge, D., Daub, S., Bosia, L., Ryf, S., & Jäger, C. (in review). *Robustness in nonmetric multidimensional scaling*.
- Läge, D., Streule, R., & Schlatter, K. (eingereicht). *Die Veränderung von Wissenskarten beim Merkmalslernen*.
- Lynch, T. G., Woelfl, N. N., Steele, D. J., & Hanssen, C. S. (1998). Learning style influences student examination performance. *The American Journal of Surgery*, 176, 62-66.
- Marx, W., & Hejj, A. (1989). *Subjektive Strukturen*. Göttingen: Hogrefe.
- Marx, W., & Läge, D. (1995). *Der ideologische Ring*. Göttingen: Hogrefe.
- Mathar, R. (1997). *Multidimensionale Skalierung*. Stuttgart: Teubner.
- Mayadas, N. S., & Elliott, D. (1992). Integration and xenophobia: an inherent conflict in international migration. *Journal of Multicultural Social Work*, 2, 47-62.
- Montada, L. (1995). Die geistige Entwicklung aus der Sicht Piagets. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (3. Aufl.). Weinheim: Psychologie Verlags Union. Kap. 11, S. 622-653.
- Novak, J. D. (2002). Meaningful learning: the essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education*, 86 (4), 548-571.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Roth, W.-M. (2000). Artificial neural networks for modeling knowing and learning in science. *Journal of Research in Science Teaching*, 37 (1), 63-80.
- Sodian, B. (1995). *Entwicklung bereichsspezifischen Wissens*. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (3. Aufl.). Weinheim: Psychologie Verlags Union. Kap. 13, S. 622-653.
- Steiner, G. (1996). *Lernen – Zwanzig Szenarien aus dem Alltag* (2. vollst. überarb. Aufl.). Bern: Huber.
- Straub, J. (1999). *Verstehen, Kritik, Anerkennung. Das Eigene und das Fremde in der Erkenntnisbildung interpretativer Wissenschaften* (Essener Kulturwissenschaftliche Vorträge, Band 4). Göttingen: Wallstein.
- Strube, G. (1996). Wissen & Wissensrepräsentation. In G. Strube et al. (Hrsg.), *Wörterbuch der Kognitionswissenschaft*. Stuttgart: Klett-Cotta. S. 799-815.
- Venville, G. J., & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretic framework. *Journal of Research in Science Teaching*, 35 (9), 1031-1055.
- Vosniadou, S. (1991). Conceptual development in astronomy. In S. M. Glynn, R. H. Yeany, & B. K. Britton (Eds.), *The psychology of learning science*. Hillsdale, NJ: Erlbaum. Pp. 149-178.
- Wang, L. L. (1991). Roots and changing identity of the Chinese in the United States. *Daedalus*, 120 (2), 181-206.

Weinberg, N. (2001). Mass migration and labor market incorporation: Soviet immigrants in Israel. *Research in Social Stratification and Mobility*, 18, 249-274.

9.
**A Structure-Based Expert Model of the ICD-10 Men-
tal Disorders**

Abstract

Many authors have highlighted that a classificatory system such as the ICD-10 or the DSM-IV should attempt to integrate categorical as well as dimensional aspects rather than relying only on a categorical distinction. In the current study, a method is presented that allows both perspectives to be considered. Based on their clinical experience, twenty therapists were asked to rate a selection of ICD-10 mental disorders in terms of their similarity. The resulting data was processed by nonmetric multidimensional scaling (NMDS) on individual basis as well as mean basis and a cluster analysis. From a categorical point of view, the resulting mean cognitive map and the cluster analytic groups showed a good concordance with the ICD-10 F groups. Additionally, analogies to the critical observations with regard to the classification systems could be observed. From the dimensional point of view, there were two dimensions found (cognition and psychosis) which showed high correlations in the mean map. In conclusion, the method which is presented in this study offers an opportunity to account for dimensional as well as categorical aspects of classification by exploring the structure of the ICD-10 mental disorders directly from the point of view of therapists and thus creating the basis for an expert model.

Keywords: ICD-10, mental disorder, expert model, cognitive structure, cognitive map, similarity judgment, nonmetric multidimensional scaling (NMDS)

9.1. Introduction

One of the main challenges of a classification system in psychopathology such as the ICD-10 Classification of Mental Disorders (World Health Organization, 1992) or the DSM-IV (American Psychiatric Association, 1994) that has repeatedly been highlighted is the difficulty of determining clear-cut boundaries between different illness entities in general (Clark, Watson & Reynolds, 1995; Dittmann, 1996; Widiger & Clark, 2000) and between various diagnostic categories in particular. For instance, Maser & Cloninger (1990) argued in “Comorbidity of Mood and Anxiety Disorders” that “In other words, psychopathology does not appear to be comprised of discrete, mutually exclusive disorders, as Kraepelin originally proposed”. Another prominent diagnostic category which has recently gained greater attention is the category of personality disorders. In this diagnostic area, too, the suggestions for an adaptation of the classification tend away from a pure categorical model towards a combination of a categorical and a dimensional model (Mundt, 2002; Oldham & Skodol, 2000; Skodol, Gunderson, Pfohl, Widiger, Livesley & Siever, 2002). Indeed, the debate of whether a categorical or a dimensional system is to be favored, which generally results in suggestions for a combination of the two, has been going on for many years (Blashfield, 1990; Clark, Watson & Reynolds, 1995). The aim of the present study is to integrate both perspectives by employing multivariate structure analyses, especially nonmetric multidimensional scaling (NMDS). These methods allow the ICD-10 classification of mental disorders to be looked at from the point of view of therapists (i.e. psychiatrists and psychotherapists) with the aim of establishing the basis for an expert model of mental disorders.

9.2. Methods

Subjects. The sample consisted of 20 practicing therapists (8 female and 12 male) in Zurich, Switzerland. Eight of the therapists hold a medical degree, eleven hold a psychological degree

and one holds both a medical and a psychological degree. Their mean clinical experience was $M = 12.75$ ($SD = 9.8$) years (min = 1.5 years, max = 30 years). In their clinical practice they apply cognitive behavioural $N = 3$, psychodynamic -/analytic $N = 2$, integrative/eclectic $N = 13$, and other $N = 2$ therapeutic approaches.

Tasks. The experiment consisted of two tasks: First the subjects had to rate 210 presented pairs of 21 disorders (e.g. paranoid schizophrenia and dysthymia) based on their similarity on a scale from 1 = minimal similarity to 9 = maximal similarity (similarity judgments: Coxon, 1982; Daub, 2001; Laege, 2001; Smith, Shoben & Rips, 1974). Subjects were specifically instructed not to attempt to reproduce the categorization of the ICD-10, but to make judgments according to their clinical experience. Second, subjects were required to separate the same selection of disorders into groups based on their similarity (parallel sorting task: Laege, 2001) in reference to (a) phenomenology and (b) aetiology. Additionally, after completing the two tasks participants answered several questions regarding their education and their professional experience, and completed ratings concerning a) the subjective importance of the ICD-10 in their clinical everyday life (1 = not important at all to 9 = very important), b) their subjective expertise in the ICD-10 (1 = nearly no knowledge to 9 = very good knowledge), c) the subjective positive attitude towards the ICD-10 (1 = not very positive to 9 = very positive), and d) their potential preference of another classification system. The sequence of the presented pairs of disorders in the first task, the two sorting criteria in the second task, and the questions regarding the ICD-10 were randomized. All tasks were presented on a computer screen with the INTUS[®] data collection software (Ryf, Daub & Laege, 2003) (except for the four ratings, which were given using paper and pencil).

Selected Disorders. The 21 disorders which had to be rated were selected based on (a) representativeness and variability with regard to the field of knowledge (i.e. the psychopathological disorders as defined by the ICD-10) and (b) clinical prevalence (based on a common ranking of three studies with regard to clinical prevalence of mental disorder diagnoses (Lange, Munk-Jorgensen, Bertelsen, Schurmann, Michels, Malchow & Dilling, 2002; Mussigbrodt, Michels, Malchow, Dilling, Munk-Jorgensen & Bertelsen, 2000; Neuenschwander, Meyer & Hell, 1999)). Mental retardation (F7), disorders of psychological development (F8), and behavioural and emotional disorders with onset usually occurring in childhood and adolescence (F9) were excluded from this experiment. Ultimately, three disorders from each category from F0 to F6 were chosen, which added up to the total of the following 21 disorders: Dementia in Alzheimer's disease with early onset (F00.1); organic delusional (schizophrenia-like) disorder (F06.2); organic personality disorder (F07.0); mental and behavioural disorder due to the use of alcohol, withdrawal state (F10.3); mental and behavioural disorder due to the use of cocaine, psychotic disorder (F14.5); mental and behavioural disorder due to multiple drug use and use of other psychoactive substances, dependence syndrome (F19.2); paranoid schizophrenia (F20.0); hebephrenic schizophrenia (F20.1); schizoaffective disorder, manic type (F25.0); bipolar affective disorder, current episode manic without psychotic symptoms (F31.1); severe depressive episode without psychotic symptoms (F32.2); dysthymia (F34.1); agoraphobia (F40.0); mixed anxiety and depressive disorder (F41.2); adjustment disorder (F43.2); anorexia nervosa (F50.0); bulimia nervosa (F50.2); nonorganic insomnia (F51.0); paranoid personality disorder (F60.0); emotionally unstable personality disorder (F60.3) and enduring personality change after catastrophic experience (F62.0).

Analysis. The data of the similarity judgments and of the parallel sorting task were translated into matrices and then analyzed by robust nonmetric multidimensional scaling (NMDS by means of the ROBUSCAL algorithm; Laege, Daub, Bosia, Ryf & Jaeger, in review) and a hierarchical cluster analytic method (average model; Bortz, 1999; Johnson, 1967; Wiedenbeck & Zuell, 2001). The NMDS analysis resulted in cognitive maps which were calculated based on individual matrices as well as on the mean matrix of the similarity judgments. The cluster

groups resulting from the cluster analysis were calculated based on the mean similarity matrix as well as on the distance matrix, which was extracted from the cognitive map calculated based on the mean similarity matrix. The cognitive maps depict the cognitive relational structure of the subjects in the form of a Euclidian space, where a small distance between the objects (i.e. disorders) corresponds to a high similarity and vice versa. This Euclidian metrics significantly differs from a city-block metrics of a factor analysis. In a factor analysis, the distance between two items can only be described as an arithmetic operation of the values of these points on the corresponding factors/ dimensions. In contrast, the Euclidian space constructed by the NMDS allows the relational position of the items to be directly described and interpreted. Additionally, the dimensions in this space do not need to be orthogonal to each other, as is compulsory for the dimensions/ factors of a factor analysis. The individual maps were compared to each other as well as to the map calculated based on mean matrix by means of a multiple procrustean transformation (Borg & Groenen, 1997; Coxon, 1982; Hurley & Cattell, 1962) (in this study the specific PROPERSCAL method was employed; Laege, 2001). These comparisons resulted in deviation measures (average losses; Laege, 2001) between the maps. Based on the deviations of the individual maps from each other, a two-dimensional map was calculated. This map represents the relational position of the therapists based on their individual cognitive structures (i.e. their cognitive maps). In the resulting map the data with regard to professional experience, therapeutic approach, academic training and attitude towards the ICD-10 was examined for structure explicative value (meaningful clusters and multiple regressions; Property Fitting; Laege, 2001).

In addition, we conducted a content analysis (comparative, semantic -semantic category analysis) (Merten, 1995) of the ICD-10 clinical descriptions and diagnostic guidelines for each of the selected disorders. In a first step, three dimensions/ categories (Cognition, Psychosis and Affect) were defined, derived from the AMDP System (Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie, 2000). In a second step, the frequency of terms occurring in the text associated with each of the dimensions was determined. This frequency served as an indicator of the emphasis of each disorder on the corresponding dimension. In a third step, these indicators were fitted into the two dimensional cognitive map (based on the mean similarity matrix) by Property Fitting. All statistical analyses were performed with Microsoft[®] Excel[®] (Microsoft Corporation, 2002), SPSS[®] (SPSS Inc., 2001), INTUS[®] (Ryf, Daub & Laege, 2003) and ASSO[®] (Schnyder, Daub, Jaeger, Ryf, Christen & Laege, 2003).

Evaluation. In a second meeting, the cognitive maps and the cluster groups were presented to those therapists who could be reached again for the evaluation (13/20). They were asked to interpret the cognitive maps based on individual and mean matrices and the cluster groups based on similarity and distance matrices with regard to clinical meaning and their subjective preference. The therapists were blinded to the calculation basis of the cognitive maps and the cluster groups, and the sequence of the presentation was randomized.

9.3. Results

The core result of this investigation is represented by the cognitive map based on the mean similarity matrix in Figure 1. This map depicts the therapists' mean cognitive relational structure of the selected mental disorders. As will be demonstrated, it enables categorical as well as dimensional structural aspects to be interpreted. Before looking at the map with regard to content, it is necessary to explain an internal quality measure of the ROBUSCAL algorithm: The standardized stress value (Laege, Daub, Bosia, Ryf & Jaeger, in review) is a measure of how well the algorithm was able to translate the similarity judgments into an n-dimensional

map. Thus it is also an indicator of the level of consistency of the similarity judgments given by the participants throughout the experiment, and consequently provides a marker of the interpretability of the cognitive map. According to the literature (Borg & Groenen, 1997; Gigerenzer, 1981) the stress value of this mean map of 0.17 as well as the stress values of the individual two-dimensional cognitive maps ($M = 0.16$, $SD = 0.02$) are more than acceptable, and thus interpretable.

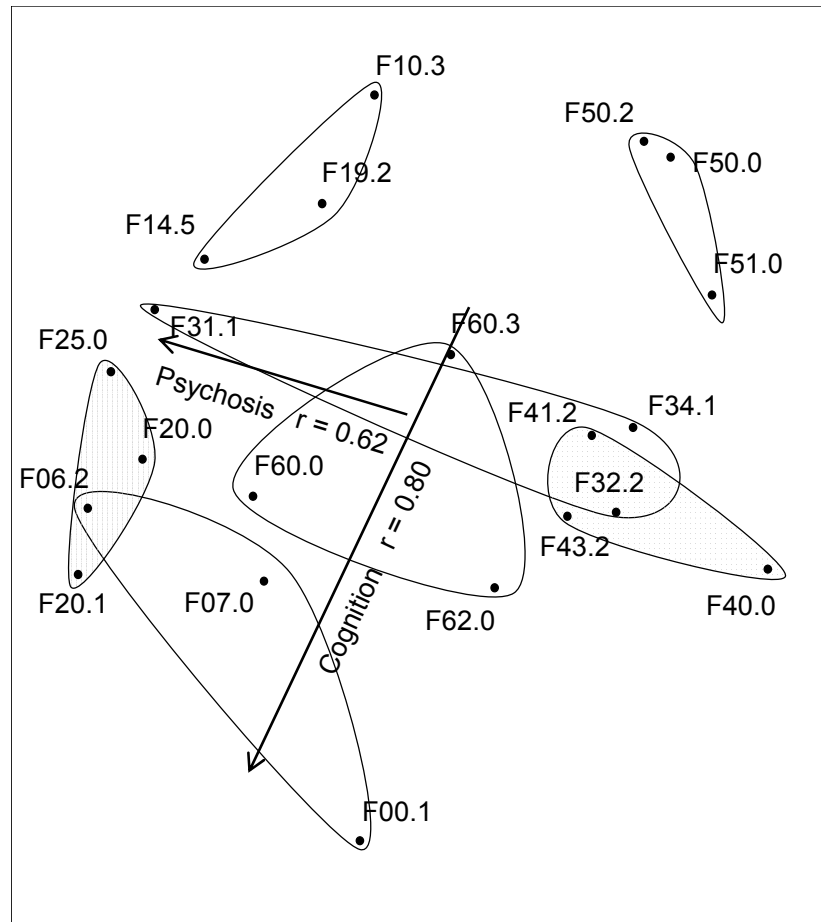


Figure 1: Two-dimensional cognitive map based on the mean similarity matrix (The plotted clusters depict the ICD-10 F categories; Dimensions: Cognition, $r = 0.80$, $R^2 = 0.64$; Psychosis, $r = 0.61$, $R^2 = 0.37$)

As mentioned above it is possible to look at the structure of the cognitive map in Figure 1 from a categorical as well as from a dimensional perspective: The plotted delineations comprise those disorders that are subsumed in the same F category according to the ICD-10, and thereby represent a traditional categorical point of view. As described in the analysis section, the cognitive maps allow the researcher to observe the Euclidian distances between the objects (i.e. disorders). Consequently it is also possible to interpret the relative positions of the disorders in reference to each other, which opens up a dimensional perspective. The examination of the cognitive maps for explicative dimensions was carried out by means of multiple regressions (Property Fitting). The values used in the Property Fitting resulted from the content analysis described in the analysis section. The two arrows in Figure 1, which are labelled “psychosis” and “cognition”, depict the orientation of the corresponding dimensions in the map. The emphasis in the ICD-10 of a given disorder on the corresponding dimension increases with its orthogonal projection on the dimension in the direction of the arrow. The following disorders showed a high correlation ($r = 0.80$, $R^2 = 0.64$) on (a) the cognition dimension (Dementia in Alzheimer’s disease with early onset (F00.1); organic delusional (schizo-

phrenia -like) disorder (F06.2); organic personality disorder (F07.0); mental and behavioural disorder due to multiple drug use and use of other psychoactive substances, dependence syndrome (F19.2); paranoid schizophrenia (F20.0); hebephrenic schizophrenia (F20.1); bipolar affective disorder, current episode manic without psychotic symptoms (F31.1); severe depressive episode without psychotic symptoms (F32.2); agoraphobia (F40.0); mixed anxiety and depressive disorder (F41.2); nonorganic insomnia (F51.0); paranoid personality disorder (F60.0) and emotionally unstable personality disorder (F60.3)), and a fairly high correlation ($r = 0.61$, $R^2 = 0.37$) on (b) the psychosis dimension (organic delusional (schizophrenia -like) disorder (F06.2); organic personality disorder (F07.0); mental and behavioural disorder due to the use of cocaine, psychotic disorder (F14.5); mental and behavioural disorder due to multiple drug use and use of other psychoactive substances, dependence syndrome (F19.2); paranoid schizophrenia (F20.0); hebephrenic schizophrenia (F20.1); schizoaffective disorder, manic type (F25.0) and paranoid personality disorder (F60.0)). The selected disorders showed only a low correlation on the affect dimension ($r = 0.33$, $R^2 = 0.11$), which is consequently not presented in the map.

In addition to the possibility of plotting the delineations of the disorders in the cognitive maps, which is described above, there is a further method for looking at the similarity matrices from a categorical point of view: The cluster analysis subsumes homogeneous (i.e. similar) objects in clusters in such a way that objects of different clusters are as heterogeneous as possible (Bortz, 1999). The cluster analysis of the mean similarity matrix resulted in six groups of similar disorders (Table 1). The detailed procedure has been described elsewhere (Bortz, 1999; Wiedenbeck & Zuell, 2001). This reanalysis of the similarity matrices by cluster analysis allows the categories found by the NMDS analysis in the cognitive maps to be confirmed or refuted.

F40.0 Agoraphobia
F62.0 Enduring personality change after catastrophic experience F43.2 adjustment disorder F41.2 mixed anxiety and depressive disorder F34.1 dysthymia F32.2 severe depressive episode without psychotic symptoms F51.0 nonorganic insomnia
F50.0 anorexia nervosa F50.2 bulimia nervosa F60.3 emotionally unstable personality disorder
F31.1 bipolar affective disorder, current episode manic without psychotic symptoms F25.0 schizoaffective disorder, manic type F20.1 hebephrenic schizophrenia F06.2 organic delusional (schizophrenia like) disorder F20.0 paranoid schizophrenia F60.0 paranoid personality disorder
F10.3 mental and behavioral disorder due to the use of alcohol, withdrawal state F14.5 mental and behavioral disorder due to the use of cocaine, psychotic disorder F19.2 mental and beh. dis. due to mult. drug use and use of other psychoactive subst., depend. Syndr.
F00.1 Dementia in Alzheimer's disease with early onset F07.0 organic personality disorder

Table 1: Groups resulting from the cluster analysis based on the mean similarity matrix

Although we have so far only been looking at the analyses of the mean results, the NMDS also allows cognitive maps to be calculated and interpreted on an individual level. This allows an examination of whether the mental structures of the therapists expressed as cognitive maps rather converge or diverge from each other. Figure 2 shows the result of the comparison (Procrustean transformation) of the individual cognitive maps of two therapists chosen as an example. The connecting lines between the disorders represent their deviation of their spatial positions in the cognitive maps of the corresponding therapists. The mean average loss of the comparisons of the individual maps to each other ($M = 0.59$, $SD = 0.09$) as well as the average losses of the comparisons of the individual maps to the map based on the mean similarity matrix ($M = 0.46$, $SD = 0.10$) were relatively high.

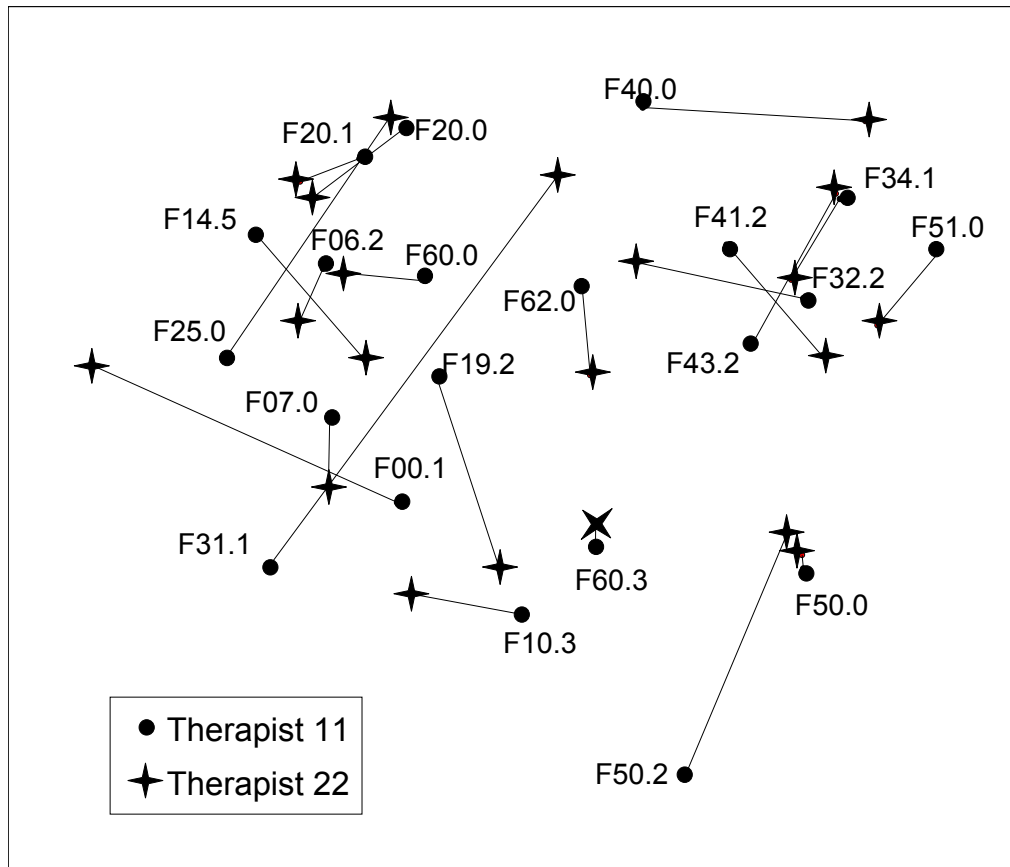


Figure 2: Procrustean Transformation of the cognitive maps of therapist 11 and therapist 22

The average losses of the comparisons of the individual maps to each other also form the basis for the map presented in Figure 3. Consequently the items in this map represent the relational position of the therapists based on the deviation of their individual cognitive structures from each other (based on only 19/20 therapists, as one outlier was excluded from this analysis). The combination of the additional variables with this map allows it to be examined whether they are associated with the relative positions of the therapists: The abbreviation above each position represents the academic training (psych. = psychology, med. = medical science, both = both degrees), the one below represents the therapeutic approach (CBT = cognitive behavioural, dynam. / analyt. = psychodynamic- / analytic, int. / eclectic = integrative / eclectic and other). In this map, none of the data with regard to professional experience (number of years) and attitude towards the ICD-10 showed significant explicative value (meaningful clusters or significant correlations in the multiple regressions).

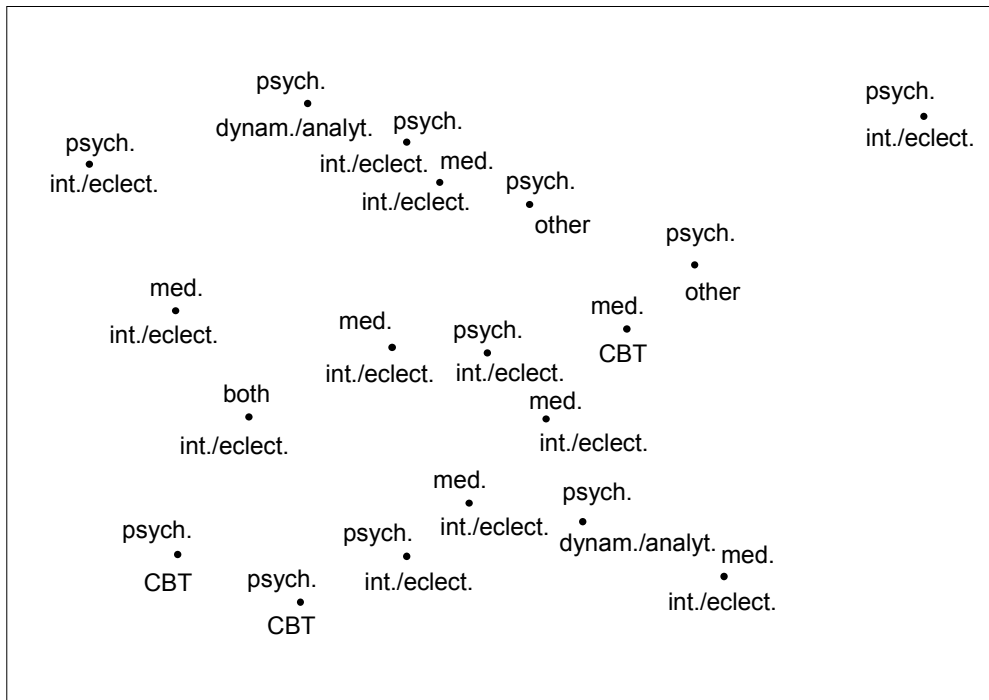


Figure 3: Relational positions of the therapists based on the deviation of their cognitive maps (Academic training: psych. = psychology, med. = medical science, both = both degrees; therapeutic approach: CBT = cognitive behavioral, dynam./analyt. = psychodynamic-/analytic, int./eclect. = integrative/eclectic and other)

Consistent with the observation of the therapists that two main aspects in the judgment of similarities of the mental disorders were phenomenology and aetiology, this issue was further investigated. Hence after completing the similarity judgments, they were asked to sort the disorders in groups of maximal similarity according to the criteria mentioned above. Figure 4 shows the map resulting from this parallel sorting task with regard to aetiology (abbr. “a”) and phenomenology (abbr. “p”). Each sorting task resulted in a 0/1 matrix (0 = disorders located in different groups, 1 = disorders located in the same group). The summarized differences of these matrices were used as distance measures to calculate the map (by NMDS). Therefore items which are positioned in spatial proximity of each other represent similar groupings and vice versa. The item labelled “ICD” represents the result which corresponds to the grouping of the disorders according to the ICD-10 F clusters. The delineations subsume all groupings with regard to phenomenology and aetiology respectively.

After having measured the cognitive structures of the mental disorders from the perspective of the therapists by way of similarity judgments and parallel sorting tasks, we verified the clinical plausibility of our analysis by confronting the therapists with the results. This evaluation of the maps and the cluster groups revealed that 8/13 therapists judged the cognitive maps to better represent the structure of the presented mental disorders than the cluster analytic groups. Among the cognitive maps, 10/13 therapists preferred the map based on the mean similarity judgments to the map based on their own individual similarity judgments. The preferred map received a good average rating ($M = 4.6$, $SD = 1.0$) on a Swiss school grade scale (1 = worst, 6 = best) with regard to acceptance by the therapists.

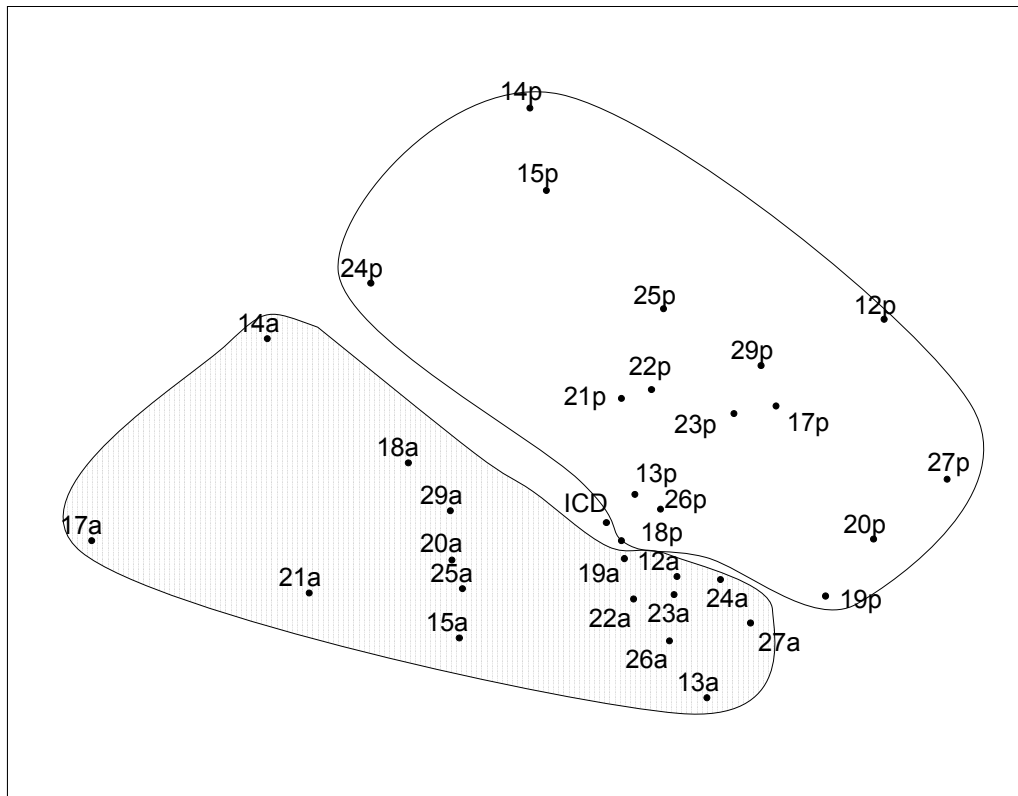


Figure 4: Results of the parallel Sorting task with regard to etiology (a) and phenomenology (p) (ICD = Grouping of the disorders according to the ICD-10 F clusters)

9.4. Discussion

As repeatedly highlighted in earlier studies, the question of whether to look at the classification of mental disorders from a categorical or from a dimensional point of view cannot really be answered exclusively in favour of one of the alternatives (Clark, Watson & Reynolds, 1995; Mundt, 2002; Widiger & Clark, 2000). The present study explored a method which attempts to bear in mind both perspectives from the point of view of therapists, i.e. the individuals who actually have to interpret the classification system. Their structural mental representation of the selected disorders was measured by similarity judgments which were translated into two-dimensional cognitive maps as well as into cluster analytic groups. Considering that the instruction was to judge similarities between the disorders based on their personal clinical experience and not to attempt to reproduce the ICD-10 structure, the spatial representation of the disorders in Figure 1 presents a meaningful configuration.

The plotted clusters depict the ICD-10 F categories, and therefore allow categorical aspects of the structure to be considered. It is noticeable that these clusters were quite exactly reproduced, as they can be clearly delineated from each other, with the exception of three overlapping areas. These areas for their part can be hypothesized to illustrate several of the critical observations regarding the ICD-10 and the DSM-IV classification, which have accumulated in the past years and which are evidently also perceived by the participating therapists. First, the overlap of F3 (mood (affective) disorders) and F4 (neurotic stress-related and somatoform disorders) was already similarly described much earlier by Maser & Cloninger (1990) as the difficulty of defining a clear-cut boundary between the diagnostic entities of mood and anxiety disorders. Second, the overlap of F6 (disorders of adult personality and behaviour) and F3 is mainly due to the spatial positions of emotionally unstable personality disorder (F60.3) and

bipolar affective disorder, current episode manic without psychotic symptoms (F31.1): F60.3 is positioned between the categories of F1 (mental and behavioural disorders due to psychoactive substance use), F3 and F5 (behavioural syndromes associated with physiological disturbances and physical factors – here especially the eating disorders have to be taken into account). It is well known that mental disorders from these three categories are frequently observed as comorbid diagnoses of F60.3 (Gunderson, 2001). F31.1 is located across from the dysthymia (F34.1) and the severe depressive episode without psychotic symptoms (F32.2), and in addition is closely located to schizoaffective disorder, manic type (F25.0) with which it shares most of its symptoms according to the ICD-10. Third, looking at the location of the organic delusional (schizophrenia-like) disorder (F06.2) in the F2 (schizophrenia, schizotypal and delusional disorders) cluster, this position also fits from the point of view of its phenomenological description, as F06.2 is only located in the F0 (organic, including symptomatic, mental disorders) category due to its aetiology.

The structure of the cognitive maps can be interpreted not only with regard to meaningful clusters as demonstrated above, but also in terms of the dimensions: One possibility for this is to collect additional data on external scales with regard to the criteria which are hypothesized to have an explicative value in the explored structure, as was done in a related study by Purkhardt & Stockdale (1993). A considerable number of the selected disorders in our experiment showed a high (cognition) and a fairly high (psychosis) correlation on two of the three dimensions that emerged from the content analysis. In contrast to the dimensions / factors of a factor analysis, these dimensions in the cognitive map calculated by NMDS can be directly attributed to the Euclidian distances between the items / disorders. Thus the spatial position of these disorders in the cognitive map can be partially explained by the underlying dimensions of cognition and psychosis. On the cognition dimension, Dementia in Alzheimer's disease occupies the most extreme position. Therefore this disorder shows the strongest emphasis on this dimension (within the structure of the cognitive map) in the ICD-10 among the selected disorders. The same observation applies to the schizoaffective disorder, which can be found at the most extreme position of the psychosis dimension.

As mentioned in the results section, the reanalysis of the similarity matrix by cluster analysis supports the verification of the clusters found in the cognitive map. The three overlaps which could be found in the map can for the most part also be observed in the results of the cluster analysis (Table 1): Adjustment disorder (F43.2), mixed anxiety and depressive disorder (F41.2), dysthymia (F34.1) and severe depressive episode without psychotic symptoms (F32.2) are subsumed in the same group, and thereby emphasize the overlap of the F3 and the F4 categories. Emotionally unstable personality disorder (F60.3) is placed in the same cluster as the eating disorders, and the organic delusional (schizophrenia-like) disorder (F06.2) is located together with the other schizophrenic disorders. A further interesting congruence of the results of the cluster analysis and the cognitive map can be observed in terms of the spatial extension of the clusters: For instance, the F6 (disorders of adult personality and behaviour) cluster, which shows a large circumference in the map, can be found in three different groups in the cluster analysis. The F1 (mental and behavioural disorders due to psychoactive substance use) cluster on the other hand, which subsumes its disorders in a small area in the cognitive map, can also be found in only one group in the cluster analysis. These findings could be interpreted as an indicator that the selected disorders belonging to a large cluster can be less clearly distinguished from other clusters than disorders belonging to a small cluster. Additionally, according to Blashfield (1990), large clusters are an indicator for a dimensional rather than a categorical structure and vice versa. The latter observation with regard to the F6 cluster is also supported by suggestions in the literature for a dimensional structure of personality disorders (Mundt, 2002; Oldham & Skodol, 2000; Skodol, Gunderson, Pfohl, Widiger, Livesley & Siever, 2002).

After having discussed the cognitive map based on the mean similarity matrix it is also interesting to look at the individual maps. The rather high average losses (i.e. the extent of the deviation) of the individual maps compared to each other as well as compared to the map based on the mean similarity judgments, suggest that the therapists' cognitive structures were highly individual. Cognitive maps of experts in other domains of expertise (e.g. ornithology) showed lower average losses (approx. 0.4) when compared to each other (Laege, 2001). Figure 2 illustrates the deviating relative positions of the corresponding disorders in the cognitive maps of the two therapists 11 and 22 chosen as an example. It is evident that some disorders show only a small distance between each other (e.g. the personality disorders) and are therefore conceptualized similarly by the therapists. Other disorders (e.g. bipolar affective disorder, current episode manic without psychotic symptoms (F31.1) and Dementia in Alzheimer's disease with early onset (F00.1)) show a larger discrepancy, which can be interpreted in terms of a larger divergence of these disorders in the cognitive structure of the corresponding therapist compared with the other therapist. Despite the considerable divergence of the subjects in our sample with regard to academic training, clinical experience and therapeutic approach (as described in the methods section), these variables did not show any explicative value in the map representing the relational position of the therapists (i.e. their cognitive maps). In this regard it is noteworthy that neither the often presumed gap between psychologists and psychiatrists, nor a grouping of the therapists in terms of their therapeutic approach could be confirmed in this study (Figure 3).

Another reason for the considerable discrepancies of the individual structures might be found in the variability of the similarity judgments. As they were based on personal clinical experience, they may involve multiple criteria, e.g. aetiology, phenomenology, comorbidity etc. This argument was explored in a successive experiment by constricting the similarity judgments to a specific criterion (i.e. phenomenology). This resulted in a smaller variance and a lower level of the average losses between the maps (Streule, Egli, Schlatter & Laege, submitted). The ability of the therapists to constrict their judgment to a single criterion (i.e. aetiology and phenomenology) was also examined in the parallel sorting task. The map presented in Figure 4 illustrates that the therapists were able to differentiate the two tasks well, as the corresponding results of the phenomenology task (abbr. "p") can clearly be delineated from the results of the aetiology task (abbr. "a"). Interestingly, the result based on the ICD-10 grouping is not located in the phenomenology cluster as could be expected (Stieglitz, Freyberger & Mombour, 2002) but rather between the phenomenology and the aetiology clusters. Thus this map illustrates the conflict of classification between these two criteria, which is evidently also perceived by the therapists who participated in our experiment.

In conclusion, the cognitive maps based on the similarity judgments offer an opportunity to empirically explore and visualize the structure of the mental disorders directly from the point of view of the therapists, who are the individuals who ultimately have to interpret and use a classification system such as the ICD-10. As we have demonstrated, a cognitive map allows the integration of a categorical as well as a dimensional perspective. The evaluation by the therapists showed that the majority favoured this presentation over a pure categorical presentation such as cluster analysis, and deemed it to be of good quality. The map based on the mean similarity matrix was preferred over the maps based on the individual matrices, and the individual maps showed a smaller divergence when compared to the mean map than compared to each other. These observations indicate that a common structure of the disorders of the participating therapists is a plausible assumption and this common configuration of the disorders as a cognitive map can therefore serve as an expert model. Such an expert model can be used for instance to structure and monitor learning progress in the field of psychopathology, where it is already being tested and included in an extensive e-learning project. Additionally, it could open up opportunities to explore local differences in perceptions of therapists

regarding a common diagnostic system (thereby serving as a new approach to determine inter-rater reliabilities already on the level of the perception of the classification system) as well as to refine the discrepancies and convergences between related diagnostic systems such as the ICD and the DSM. Furthermore profiles of patients (e.g. descriptions of case examples) could be fitted into a cognitive map of mental disorders and help to clarify issues of comorbidity and multiple diagnoses. Combined with the arguments mentioned above this method might help to build a bridge across the gap between theoretical concepts and clinical practice in the field of psychopathology.

9.5. References

- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, ed. 4 (DSMIV)*. Washington: American Psychiatric Association.
- Arbeitsgemeinschaft fuer Methodik und Dokumentation in der Psychiatrie (2000). *Das AMDP-System Manual zur Dokumentation psychiatrischer Befunde (7. Aufl.)*. Goettingen: Hogrefe.
- Blashfield, R. K. (1990). Comorbidity and Classification. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of Mood and Anxiety Disorders*. Washington: American Psychiatric Press. Pp. 61-82
- Borg, I. & Groenen, P. (1997). *Modern Multidimensional Scaling – Theory and Applications*. New York: Springer.
- Bortz, J. (1999). *Statistik fuer Sozialwissenschaftler (5. Aufl.)*. Berlin: Springer.
- Clark, L. A., Watson, D. & Reynolds, S. (1995). Diagnosis and Classification of Psychopathology: Challenges to the Current System and Future Directions. *Annual Review of Psychology*, 46, 121-153.
- Coxon, A. P. M. (1982). *The Users Guide to Multidimensional Scaling*. London: Heinemann.
- Daub, S (2001). *Similarity Simulation - wie man den Code des Globalurteils knackt*. Dissertation, Philosophische Fakultät, Universität Zuerich. Zuerich.
- Dittmann, V. (1996). Grundprinzipien der psychiatrischen Diagnostik und Klassifikation. In B. Neundorfer (Hrsg.), *Atlas der Nervenheilkunde Neurologie und Psychiatrie in Bild und Wort*. Karlsruhe: Braun Fachverlage. S. 282-285.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. Muenchen: Reinhardt.
- Gunderson, J. G. (2001). *Borderline Personality Disorder – A Clinical Guide*. Washington: American Psychiatric Press.
- Hurley, J.R. & Cattell, R. B. (1962). The Procrustes Program: Producing Direct Rotation to Test a Hypothesized Factor Structure. *Behavioral Science*, 7, 258-262.
- Johnson, S. C. (1967). Hierarchical Clustering Schemes. *Psychometrika*, 32, 241-254.
- Laege, D. (2001). *Aehnlichkeitsbasierte Diagnostik von Sachwissen. Habilitationsschrift, Philosophische Fakultät, Universität Zuerich*. Zuerich.
- Läge, D., Daub, S., Bosia, L., Ryf, S. & Jaeger, C. (in review). *Robustness in Nonmetric Multidimensional Scaling*.
- Lange, W., Munk-Jorgensen, P., Bertelsen, A., Schurmann, A., Michels, R., Malchow, C. P. & Dilling, H. (2002). *Comparison of Psychiatric ICD-10 Diagnoses in Denmark and Germany*. *Psychopathology*, 35, 36-47.
- Maser, J. D. & Cloninger, C. R. (Eds.) (1990). *Comorbidity of Mood and Anxiety Disorders*. Washington: American Psychiatric Press.
- Merten, K. (1995). *Inhaltsanalyse: Einfuehrung in Theorie, Methode und Praxis (2. Aufl.)*. Opladen: Westdeutscher Verlag.

- Microsoft Corporation (2002). *Microsoft Excel v 2002*. Microsoft Corporation.
- Mundt, C. (2002). Psychological Perspectives for the Development of Future Diagnostic Systems. *Psychopathology*, 35, 145-151.
- Mussigbrodt, H., Michels, R., Malchow, C.P., Dilling, H., Munk-Jorgensen, P. & Bertelsen, A. (2000). Use of the ICD-10 Classification in Psychiatry: An International Survey. *Psychopathology*, 33, 94-99.
- Neuenschwander, M., Meyer, P. C. & Hell, D. (1999). Tabellenband 1998: *stationaere Behandlungen in psychiatrischen Kliniken, Psychotherapiestationen und Institutionen fuer Suchtkranke des Kantons Zuerich*. Psychiatrische Universitaetsklinik Zuerich. Zuerich.
- Oldham, J. M. & Skodol, A. E. (2000). Charting the Future of Axis II. *Journal of Personality Disorders*, 14, 17-29.
- Purkhardt, S. C. & Stockdale, E. (1993). Multidimensional scaling as a Technique for the Exploration and Description of a Social Representation. In G. M. Breakwell & D. V. Canter (Eds.), *Empirical Approaches to Social Representation*. Oxford: Clarendon Press. Pp. 272-297.
- Ryf, S., Daub, S. & Laege, D. (2003). *INTUS - Computergestuetzte Erfassung von relationalen Urteilen*, v 2.4.4. Allgemeine Psychologie, Psychologisches Institut, Universitaet Zuerich. Zuerich.
- Schnyder, F., Daub, S., Jaeger, C., Ryf, S., Christen, S. & Laege, D. (2003). *ASSO - Statistische Auswertung von Assoziationsdaten*, v 3.2.3. Allgemeine Psychologie, Psychologisches Institut, Universitaet Zuerich. Zuerich.
- Skodol, A. E., Gunderson, J. G., Pfohl, B., Widiger, T. A., Livesley, W. J. & Siever, L. J. (2002). The Borderline Diagnosis I: Psychopathology, Comorbidity and Personality Structure. *Biological Psychiatry*, 51, 936-950.
- Smith, E. E., Shoben, E. J. & Ribs, L. J. (1974). Structure and Process in Semantic Memory. A Featural Model for Semantic Decisions. *Psychological Review*, 81, 214-241.
- SPSS Inc. (2001). *SPSS for Windows v 11.0.1*. SPSS Inc.
- Stieglitz, R. D., Freyberger, H. J. & Mombour, W. (2002). Klassifikation und diagnostischer Prozess. In H. J. Freyberger, W. Schneider & R. D. Stieglitz (Hrsg.), *Kompendium Psychiatrie Psychotherapie Psychosomatische Medizin (11. vollst. ern. u. erw. Aufl.)*. Basel: Karger. S. 17-31.
- Streule, R., Egli, S., Schlatter, K. & Laege, D. (submitted). *Das Zürcher Wissensstrukturmodell für Psychopathologie ZWP – Eine Pilotstudie*.
- Widiger, T. A. & Clark, L. A. (2000). Toward DSM-V and the Classification of Psychopathology. *Psychological Bulletin*, 126, 946-963.
- Wiedenbeck, M. & Zuell, C. (2001). Klassifikation mit Clusteranalyse: Grundlegende Techniken hierarchischer und K-means-Verfahren. *ZUMA How-to-Reihe*, 10.
- World Health Organization (1992). *ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization.

10.

**Das Zürcher
Wissensstrukturmodell für Psychopathologie (ZWP)
—
eine Pilotstudie**

Zusammenfassung³⁸

Theoretischer Hintergrund: Es besteht ein Bedarf an objektiven Kriterien zur Evaluation von psychotherapeutischen Aus-, Weiter- und Fortbildungsprogrammen.

Fragestellung: Überprüfung der Verfahrenslogik „Das Zürcher Wissensstrukturmodell für Psychopathologie (ZWP)“ als Qualitätssicherungsinstrument.

Methode: Darstellung der strukturellen Aspekte des individuellen deklarativen therapeutischen Sachwissens (psychopathologische Störungsbilder) in Wissenskarten über die Nonmetrische Multidimensionale Skalierung und Vergleich mit einem Normmodell zur Abschätzung der momentanen Wissensqualität mit 26 Kandidaten aus den vier Jahrgängen der postgradualen psychotherapeutischen Weiterbildung (Verhaltensmedizin und Kognitive Verhaltenstherapie) der Universität Zürich.

Ergebnisse: Mit zunehmender Expertise unterscheiden sich die Wissenskarten der Auszubildenden immer weniger von einem Expertenmodell, und auch die Unterschiede zwischen den Karten der Auszubildenden nehmen ab.

Schlussfolgerung: Die Resultate sprechen für den Einsatz von ZWP als komplementäres Evaluationsinstrument in der psychotherapeutischen Aus-, Weiter- und Fortbildung.

Schlüsselwörter: Nonmetrische Multidimensionale Skalierung, Prokrustes-Transformation, Qualitätssicherung, Bildungsprogramme, Evaluationsinstrument, Wissensmessung, Psychotherapie-Training, Psychopathologie

10.1. Einleitung

„Die Qualität der therapeutischen Versorgung kann nur so gut sein wie die Ausbildung der Therapeuten.“ lautet der einleitende Satz des Beitrags von Reinecker & Schindler (2000, S. 655). Diese Feststellung führt zur Forderung nach einer qualitativ hochwertigen, sich ständig weiterentwickelnden und an neue Erkenntnisse aus der psychotherapeutischen Forschung anpassenden Aus-, Weiter- und Fortbildung. Gleichzeitig ist damit aber auch die Frage der Evaluation der eingesetzten Lehr- und Lerninhalte sowie der didaktischen Methoden angesprochen. Trotz der Forderung nach effektiven Psychotherapie-Bildungsprogrammen fehlt es vielfach an überzeugenden Evaluationsinventaren.

Kenntnisse (Wissen um theoretische Grundlagen), Fertigkeiten (konkreter Umgang mit Patienten) und Erfahrung (Aspekte der Interaktion eigener persönlichen Merkmale mit der therapeutischen Situation) bilden die drei Grundpfeiler der Kompetenz qualifizierter Psychotherapeuten (Buchkremer & Schneider, 1995; Reinecker & Schindler, 2000). Diese zu erreichen ist das explizit zu fordernde Ergebnis eines Bildungsprogramms (Willutzki & Botermans, 1997). Dabei wird diese Qualifikations-Charakteristik üblicherweise als gestufter Prozess betrachtet: Viele Aus-, Weiter- und Fortbildungen in Psychotherapie vermitteln zu Beginn des Lernprozesses grundlegendes Wissen über Störungsbilder und Handlungsmodelle. Erst dann folgen praktische Anwendungen, welche Fertigkeiten und die notwendige (Selbst)Erfahrung vermitteln. In diesem gestuften Lernprozess soll die Umsetzung abstrakter Wissensinhalte in automatisierte und adäquate Handlungsabläufe das Ziel der psychotherapeutischen Bildung sein

³⁸ An dieser Stelle möchten wir uns bei Frau Prof. Dr. Ulrike Ehlert und Frau Dr. Nora Kaiser für die Möglichkeit zur Durchführung der Studie im Rahmen ihres Weiterbildungskurses sowie für die anregenden Gespräche zum Einsatz des Verfahrens und zur Einordnung der Resultate bedanken.

(Eckert & Westhof, 2000, Freedheim & Overholser, 1998). Grundlage und zwingend notwendige Voraussetzung für ein an die jeweilige Situation angepasstes therapeutisches Verhalten ist damit das Vorhandensein deklarativer, kognitiver Wissensstrukturen. Erst die differenzierte Verarbeitung, das Erkennen von Gemeinsamkeiten und Unterschieden, die Integration und Elaboration einzelner Wissensinhalte in ein Gesamtbild in der Ausbildung ermöglichen den anschließenden Wissenstransfer in die praktische Situation hinein. Das dürfte der Grund sein, warum auch in der praktisch orientierten Therapieausbildung die Vermittlung vernetzten Wissens eine zentrale Funktion wahrnimmt.

Die Ergebnisqualität eines an solchen Zielen orientierten psychotherapeutischen Curriculums sicherzustellen, ist freilich nicht trivial. Schwierigkeiten ergeben sich bereits bei der Bestimmung objektiver Standards zur Beurteilung von Therapieausbildungen: Auch wenn intuitiv ein Verständnis darüber vorhanden ist, was einen guten Therapeuten ausmacht (Reinecker & Schindler (2000) sowie Freedheim & Overholser (1998) listen eine Reihe von Aspekten auf), bleibt die Operationalisierung dieser Faktoren ein kritisches Unterfangen. Auffällig oft werden Verfahren der umfassenden Selbsteinschätzung zur Evaluation von Bildungsprogrammen verwendet, wobei sich die Datengrundlage meist auf ein deskriptives Niveau beschränkt. Selbsteinschätzungen behelfen sich beispielsweise der Erfassung von Merkmalen wie der allgemeinen Zufriedenheit mit der Ausbildung (Foulkes, 2003), des subjektiv beurteilten Lernfortschritts (Schulz, 2000; Schulz et al. 1998), der allgemeinen Erweiterung bzw. Steigerung des Selbstvertrauens (Selbsteinschätzung der therapeutischen Kompetenz: Beitman & Yue, 1999; Elizur et al., 1994; Jagsch et al., 2002; Milne et al., 1999; Schulz, 2000; Schulz et al., 1998; Willutzki & Botermans, 1997) oder des auf die Aus- bzw. Weiterbildung zurückzuführenden beruflichen Aufstiegs (Schulz, 2000; Schulz et al., 1998).

Diese – zumeist über Fragebögen oder Interviews erhobenen – Selbsteinschätzungen bilden für den Ausbildungspfeiler der Erfahrung und, mit Abstrichen, auch für den der Fertigkeiten sicherlich die geeignete Datengrundlage für eine Evaluation des Erreichten. Sofern es aber um die Messung des vernetzten Wissens geht, welches ja als Grundlage für die weiteren, praktischen Ausbildungsschritte gesehen wird, sollten objektivere Evaluationskriterien den Bildungsprozess begleiten. Zu schnell wird das eigene deklarative Fachwissen als adäquat vernetzt eingeschätzt, denn der aktuelle Wissensstand definiert ja das, was als Wirklichkeit erlebt (und damit als Vergleichskriterium herangezogen) wird. Aus diesem Grund ist es positiv zu sehen, dass die meisten Bildungsinstitutionen für den Bereich des deklarativen Wissens Prüfungen verlangen.

Leider fragen derartige Prüfungen – sofern sie standardisiert sind – in aller Regel nur einzelne Details ab, und das notwendigerweise auch nur stichprobenartig. Die Vernetzung des Wissens (d.h. das Sehen von Zusammenhängen und Unterschieden) und die Wissensstruktur als Ganze lassen sich auf diese Weise aber nur unzureichend überprüfen. Aus diesem Grund bietet es sich an, ergänzend zum traditionellen Abfragen von Fakten auch relationales Wissens zu erheben, und zwar auf eine genauso standardisierte und damit vergleichbare Weise wie es die üblichen Prüfungsformate ermöglichen.

Die im Folgenden vorgestellte Studie wird dazu einen methodischen Vorschlag unterbreiten. Sie stützt sich exemplarisch auf die Messung des relationalen Wissens über psychopathologische Störungsbilder, welches, wie Reinecker & Schindler (2000) anmerken, einen wichtigen Pfeiler therapeutischer Kompetenz bildet. Das störungsspezifische Wissen mit den Dimensionen Phänomenologie, Ätiologie, Prävalenz usw. umfasst eine Vielzahl an einzelnen Aspekten, die vom Lernenden im Laufe der Ausbildung zueinander in Beziehung zu setzen und zu einem gesamthaften relationalen Wissensbild zu integrieren sind. Deswegen erscheint es vertretbar, das vorgeschlagene Evaluationsverfahren an diesem Inhaltsbereich des deklarativen therapeutischen Wissens zu entwickeln und erstmalig zu überprüfen.

Das Verfahren folgt der strukturellen Idee von Mapping-Techniken (vgl. Mandl & Fischer, 2000), überlässt die Anforderung des gesamthaften Strukturierens aber nicht dem Befragten (was nur zu oft zu Überforderungen führt), sondern zerlegt die Gesamtaufgabe des Wiedergebens relationalen Wissens in kleinstmögliche Einheiten. Dazu erhebt es über numerische relationale Urteile die subjektiv eingeschätzte Ähnlichkeit zwischen je zwei psychopathologischen Störungsbildern, wobei das Verfahren bewusst die Kriterien für die Ähnlichkeitseindrücke dem Wissensstand der evaluierten Person überlässt und nicht durch das Frageformat vorgibt. Im Laufe einer längerfristigen Ausbildung können sich diese Kriterien nicht nur erweitern, sondern auch fundamental verändern (in der Tradition Piagets (1976) würde das dann als Akkomodationsleistung bezeichnet).

Die durch numerische Ähnlichkeitsurteile ermittelten Proximitäten zwischen Paaren von Objekten lassen sich statistisch durch Nonmetrische Multidimensionale Skalierung (NMDS) als Distanzwerte verrechnen und so zur Grundlage von Wissenskarten machen (die Logik dieses Verfahrens ist z.B. bei Borg & Groenen, 1997 beschrieben; vgl. auch Läge, Daub, Bosia, Ryf & Jäger, in review). Auf diese Weise gewinnt man die relationale Information über das aktuelle Sachwissen auf personenspezifischer Ebene. Solche individuellen Wissenskarten können durch Prokrustes-Transformation mit einem Expertenmodell verglichen werden (Gower & Dijksterhuis, 2004; Läge, 2001). Bereits ohne großen mathematischen Aufwand springt ins Auge, welche Objekte der Lerner gut kennt (d.h. korrekt verortet) und welche er falsch platziert (Abbildung 70). Die Abweichungen zwischen der Lernerkarte und der Expertenkarte werden auf der Objektebene über die Distanz zwischen den beiden korrespondierenden Punkten (= Object Loss oder kurz ObjLoss) quantifiziert. Die gemittelte Summe aller Einzelabweichungen ergibt das Gesamtmaß des Average Loss (kurz als AvgLoss bezeichnet). Der AvgLoss dient also der Beurteilung des gesamten Wissensstandes einer einzelnen Person, verglichen mit einem Norm- / Expertenmodell.

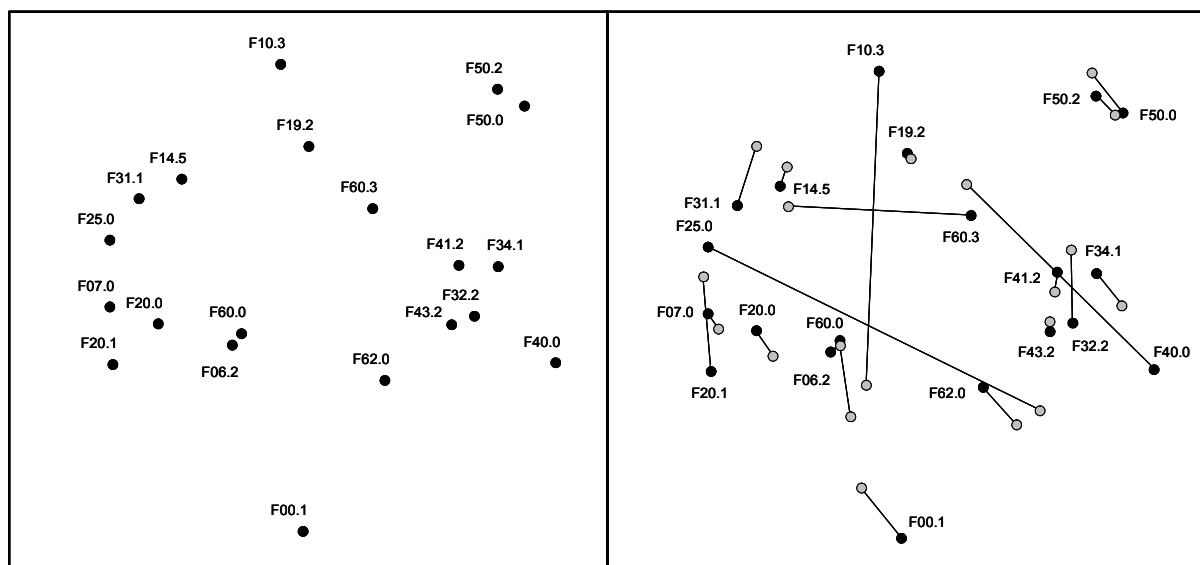


Abbildung 70: Auf Ähnlichkeitsurteilen beruhende NMDS-Lösung (links) und Ergebnis einer Prokrustes-Transformation (rechts). Die linke Karte steht für das in dieser Studie verwendete Expertenmodell. Abgebildet sind 20 Störungsbilder, codiert nach ICD-10 (s. Tabelle 5). Die grauen Punkte in der rechten Abbildung stehen für die Objektpositionen einer Lerner-Karte. Die korrespondierenden Punkte zwischen der Normkarte (links) und der Lernerkarte sind mit einer Verbindungslinie gekennzeichnet (= ObjLoss). Die Gesamtabweichung beträgt in diesem Beispiel AvgLoss = 0.37. Eine strukturelle Ähnlichkeit der beiden Karten ist erkennbar, jedoch weisen einige Objektpositionen des Lerners große Abweichungen auf (z.B. F25.0, F10.3, F40.0).

Der Erstellung des Vergleichskriteriums (Normmodell) muss großes Gewicht beigemessen werden. Im Gegensatz zu anderen Inhaltsgebieten, in denen Ähnlichkeitsrelationen zwischen Objekten über die mathematische Berechnung bestimmt werden können, werden im hier vorgestellten Bereich der psychopathologischen Störungsbilder die Sichtweisen mehrerer Fachpersonen empirisch erfragt und zu einem gemeinsamen Modell verrechnet. Damit verschwinden individuelle Gewichtungunterschiede in den Beurteilungsdimensionen und das „Gemeinsame“ in den psychopathologischen Weltbildern dieser Experten kristallisiert sich heraus. Relevanz und Validität einer solchen Mittelwertsbildung lässt sich aus den Resultaten kontrollierter Lernexperimente herauslesen, in denen gezeigt werden konnte, dass über das eben beschriebene Verfahren Faktenwissen und Wissensveränderungen über einen kontrollierten und für alle Probanden gleichsam strukturierten Lernprozess im Sinne einer Konvergenz individueller Wissenskarten an ein Expertenmodell gemessen werden können (Läge, Rüfenacht, Fichter & Streule, eingereicht; Läge, Schlatter & Streule, eingereicht; Läge, Streule & Schlatter, eingereicht). Das in der hier präsentierten Studie verwendete Normmodell (Abbildung 70, links) stammt aus einer Expertenbefragung von Egli, Schlatter, Streule & Läge (submitted). 20 praktizierende Therapeuten (8 w, 12 m) aus der Umgebung von Zürich hatten dazu die Ähnlichkeit dieser Störungsbilder paarweise eingeschätzt. Die gemittelten Ähnlichkeitsdaten wurden mit dem robusten NMDS-Algorithmus RobuScal (Läge, Daub, Bosia, Ryf & Jäger, in review) skaliert und von den Experten qualitativ evaluiert. Die Auswahl der Störungsbilder erfolgte zum einen aufgrund der Repräsentativität (aus allen F-Kategorien des ICD-10 ohne F7, F8 und F9) und Variabilität und zum anderen aufgrund der klinischen Relevanz (Prävalenz).

Im Verlauf einer psychotherapeutischen Ausbildung sollten sich die individuellen strukturellen Sichtweisen nach und nach an das Bild annähern, welches als gemeinsame Essenz der Expertensicht gesehen werden kann. Diese Hypothese folgt zumindest aus dem typischen konvergenten Lernstoff einer Ausbildung, und das Erreichen einer Annäherung setzt zugleich voraus, dass einzelne gelernte Fakten nicht nur abgespeichert, sondern zu einem relationalen Wissen verknüpft (also integriert und elaboriert) sein sollte. Konkret auf den gewählten Objektbereich bezogen bedeutet dies, dass die adäquate Verarbeitung der Merkmalsinformation es einer Person erlaubt, Unterschiede und Gemeinsamkeiten zwischen Störungsbildern zu entdecken. Dieses wiederum sollte in letzter Konsequenz zu differenzierten und qualitativ stimmigen Ähnlichkeitsurteilen führen. Mit zunehmendem Expertisegrad (in der folgenden Pilotstudie operationalisiert über das Weiterbildungsjahr) sollten sich somit die Wissensunterschiede (operationalisiert über die AvgLoss der individuellen Lernerkarten vom Normmodell) verkleinern. Des Weiteren ist zu erwarten, dass sich durch die Auseinandersetzung aller Auszubildenden mit dem gleich strukturierten Lehrgang eine Homogenisierung der Wissenskarten mit zunehmendem Expertisegrad ergibt. Die individuellen Karten sollten sich also mit fortschreitender Ausbildungszeit auch aufeinander zu bewegen.

Diese Verfahrenslogik des Vergleichs zwischen Lernerkarten und Expertenkarten wird in der vorgestellten Pilotstudie nicht nur erstmalig für das Inhaltsgebiet der Psychopathologie überprüft, sondern es handelt sich überhaupt um den ersten probenhaften Einsatz dieses Evaluationsinstrumentes im „real existierenden“ Unterricht (d.h. außerhalb von kontrollierten Laborexperimenten). Sollten sich die aufgestellten Hypothesen einer systematischen Konvergenz der gemessenen Wissenskarten bestätigen lassen, so läge damit ein Indiz vor, dass der Ausbau dieses Verfahren für einen breiteren Praxiseinsatz lohnenswert sein könnte.

10.2. Methode

Stichprobe. Eine erste Pilotstudie wurde in Kooperation mit dem Lehrstuhl Klinische Psychologie und Psychotherapie der Universität Zürich im Rahmen der postgradualen psychotherapeutischen Weiterbildung mit den Schwerpunkten Verhaltensmedizin und Kognitiver Verhaltenstherapie³⁹ (Ehlert & Kaiser, 2001; Kaiser & Ehlert, 2003) im Wintersemester 03/04 durchgeführt. Der Personenkreis besteht aus Psychologen und Ärzten. Die 4-jährige, berufsbegleitende Weiterbildung für Psychologen ist von der Föderation Schweizer Psychologinnen und Psychologen (FSP) anerkannt.

Von den 37 Teilnehmern in den vier Weiterbildungsjahren konnten insgesamt 26 Personen (24 w, 2 m) für die freiwillige Teilnahme an der Untersuchung gewonnen werden. Aus dem ersten Weiterbildungsjahr stammen 14 (von 15) Personen (13 w, 1 m), aus dem zweiten 8 (von 8) Personen (8 w), aus dem dritten 2 (von 9) Personen (1 w, 1 m) und aus dem vierten 2 (von 5) Personen (2 w). Die einzelnen Stichproben werden in der Folge mit WB1 (= 1. Weiterbildungsjahr) bis WB4 (= 4. Weiterbildungsjahr) gekennzeichnet.

Erhebungsdesign. Das Untersuchungsdesign ist als Quasi-Längsschnitt über die vier Weiterbildungsjahrgänge hin konzipiert. Die computerbasierte Datenerhebung fand in der zweiten Woche des Wintersemesters 03/04 statt. Die ersten beiden probandenstarken Weiterbildungsjahrgänge (WB1, WB2) führten den Versuch im Plenum in Laborräumen durch; die Datenerhebung mit den vier Versuchspersonen des dritten und vierten Jahrgangs (WB3, WB4) erfolgte individuell in Laborräumen. Die Probanden führten die Ähnlichkeitseinschätzung der 20 ausgewählten Störungsbilder (Tabelle 5) nach dem Beurteilungskriterium Phänomenologie durch (der höheren Kontrollierbarkeit wegen wurde auf andere Beurteilungsdimensionen wie Ätiologie usw. zunächst einmal verzichtet). Alle 190 Paare von Störungsbildern wurden in randomisierter Reihenfolge am Computer dargeboten und waren jeweils auf einer neunstufigen Ähnlichkeitsskala (1 = minimale Ähnlichkeit; 9 = maximale Ähnlichkeit) zu beurteilen. Die Befragten benötigten für diese Aufgabe zwischen 15 und 25 Minuten.

Einschränkende Rahmenbedingungen. Der Rückgriff auf ein bereits vorliegendes Expertenmodell bedeutet, dass nicht alle vorkommenden Störungsbilder in gleicher Weise für die postgraduale psychotherapeutische Weiterbildung mit den Schwerpunkten Verhaltensmedizin und Kognitiver Verhaltenstherapie der Universität Zürich relevant sind. Die in der Studie von Egli et al. (submitted) angefragten Experten konnten zudem die Beurteilungsdimensionen selbst wählen (im Gegensatz zur ausdrücklichen Aufforderung an die Teilnehmer der postgradualen Weiterbildung, sich bei den Beurteilungen an das Kriterium Phänomenologie zu halten). Nachbefragungen der Experten hatten zwar eine starke Fokussierung auf dieses Kriterium zu Tage gebracht, ein Einfluss anderer Beurteilungsdimensionen auf das Ähnlichkeitssurteil ist jedoch nicht auszuschließen. Diese partielle Inkompatibilität von Expertenmodell und Lernzielen der Weiterbildung führt dazu, dass die hier vorgelegte Pilotstudie zwar das zentrale Testkriterium überprüfen kann (nämlich ob die Weiterbildung zu einer systematischen Annäherung der individuellen Kognitiven Karten an ein allgemeines Expertenmodell führt), dass sie jedoch die tatsächliche Leistungsfähigkeit dieses Verfahrens als begleitende Längsschnittdiagnostik systematisch unterschätzen dürfte.

³⁹ Als Lehrinhalt werden im 1. und 2. Weiterbildungsjahr Grundlagen, theoretisches Basiswissen und störungsspezifisches Wissen vermittelt, durchgängig begleitet von Selbsterfahrung und klinischer Praxis. Im weiteren Verlauf (2. - 4. Weiterbildungsjahr) kommen eigene therapeutische Erfahrungen, Gruppensupervision und Fallseminare hinzu. Mehrere Experten von verschiedenen Institutionen dozieren zu ihren jeweiligen Fachgebieten.

F00.1	Demenz bei Alzheimer-Krankheit mit spätem Beginn
F06.2	organische wahnhafte (schizophreniforme) Störung
F07.0	organische Persönlichkeitsstörung
F10.3	Störung durch Alkohol, Entzugssyndrom
F14.5	Störung durch Kokain, psychotische Störung
F19.2	Störung durch multiplen Substanzgebrauch und Konsum sonst. psychotr. Subst., Abhängigkeitssyndrom
F20.0	paranoide Schizophrenie
F20.1	hebephrene Schizophrenie
F25.0	schizomanische Störung
F31.1	bipolare affektive Störung, gegenwärtig manische Episode, ohne psychotische Symptome
F32.2	schwere depressive Episode ohne psychotische Symptome
F34.1	Dysthymia
F40.0	Agoraphobie
F41.2	Angst und depressive Störung, gemischt
F43.2	Anpassungsstörung
F50.0	Anorexia nervosa
F50.2	Bulimia nervosa
F60.0	paranoide Persönlichkeitsstörung
F60.3	emotional instabile Persönlichkeitsstörung
F62.0	andauernde Persönlichkeitsänderung nach Extrebelastung

Tabelle 5: Diagnostizierte Wissensobjekte: 20 psychopathologische Störungsbilder aus der ICD-10-Kategorie F. Linke Spalte: F-Codes; Rechte Spalte: ausführliche Bezeichnung.

10.3. Resultate

Interpretierbarkeit der Wissenskarten. Die individuellen Ähnlichkeitsmatrizen wurden mit dem NMDS-Algorithmus RobuScal (Läge, Daub, Bosia, Ryf & Jäger, in review) zweidimensional skaliert. Zur Überprüfung der statistischen Qualität der resultierenden NMDS-Karten wird der NMDS-Stress herangezogen (vgl. Borg & Groenen, 1997). Dieser Wert steht für die Abbildungsgüte der einer Karte zugrunde liegenden Ähnlichkeitseinschätzungen im Vergleich zu den relativen Verhältnissen der Objektpositionen in der Karte. Die Werte der individuellen NMDS-Lösungen bewegen sich durchgängig unter dem von Gigerenzer (1981, S. 353) für eine zweidimensionale, aus 20 Objekten bestehende NMDS-Lösung als kritisch erachteten Niveau von 0.28 (WB1: $M = 0.16$, $SD = 0.02$; WB2: $M = 0.15$, $SD = 0.02$; WB3: $M = 0.16$, $SD = 0.01$; WB4: $M = 0.14$, $SD = 0.03$). Die individuellen Wissenskarten bauen sich somit aus systematisch zueinander passenden Ähnlichkeitsurteilen auf. Das bedeutet, bereits im ersten Jahr geben die Kursteilnehmer in sich hinlänglich konsistente Urteile ab (was natürlich nichts über die inhaltliche Qualität aussagt). Damit ist eine wichtige Grundvoraussetzung gegeben, die relationalen Ähnlichkeitsurteile überhaupt als Messinstrument für das strukturelle Wissen einsetzen zu können.

Unterschiede im Expertisegrad. Die gemittelten AvgLoss-Werte nehmen über die einzelnen Stichproben hinweg kontinuierlich ab: WB1 zeigte einen gemittelten AvgLoss-Wert von 0.471 ($SD = 0.090$), gefolgt von WB2 mit einem AvgLoss von 0.416 ($SD = 0.095$), WB3 mit einem AvgLoss von 0.365 ($SD = 0.004$) und WB4 mit 0.317 ($SD = 0.065$). Diese Abnahme wird auf dem 5%-Niveau signifikant (Kruskal-Wallis: $X^2(3, 26) = 7.893$, $p < .05$). Zur Verdeutlichung niedriger und hoher struktureller Unterschiede zwischen Lernerkarten und der Normkarte finden sich in Abbildung 71 je ein Beispiel eines Auszubildenden aus WB1 und aus WB4.

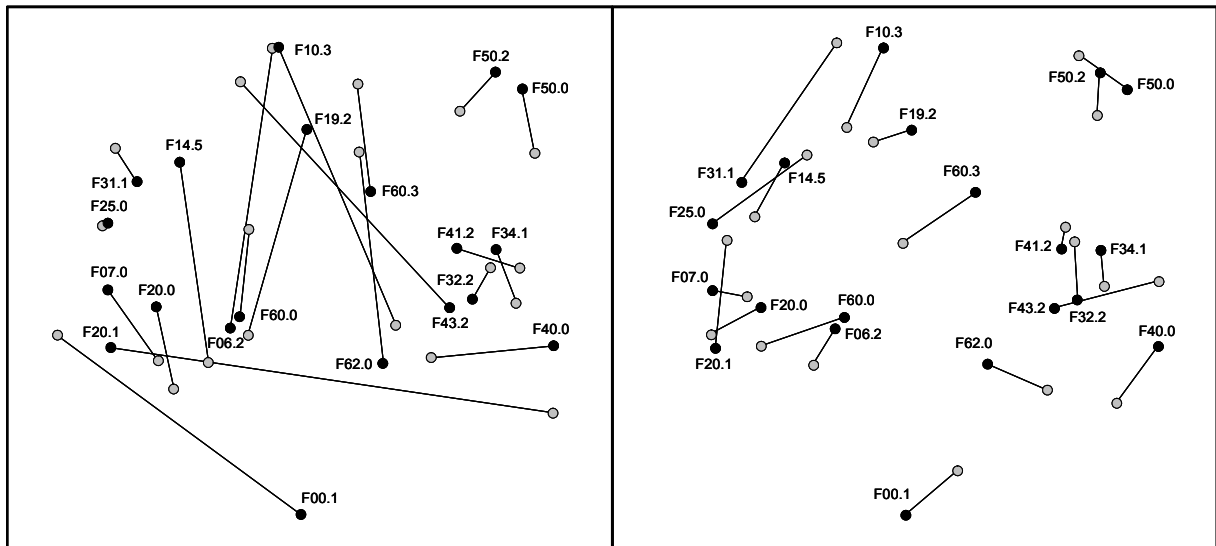


Abbildung 71: Zwei Prokrustes-Vergleiche individueller Karten mit dem Normmodell. Links Versuchsperson WB1-4, rechts Versuchsperson WB4-1. Die Abweichungswerte betragen für WB1-4 AvgLoss = 0.61, für WB4-1 AvgLoss = 0.29. Schwarze Punkte stehen für die Objektpositionen im Normmodell, graue für diejenigen der Lernenden.

Person WB1-4 weist in ihrer Wissenskarte hohe strukturelle Unterschiede zum Expertenmodell auf. Nur vereinzelte Störungsbilder befinden sich in der korrekten Region (F25.0, F31.1, F32.2, F34.1, F41.2, F50.0, F50.2). Mehrere von ihnen verletzen die Clusterzugehörigkeit. Die kognitive Struktur der psychopathologischen Störungsbilder hat sich also noch nicht adäquat ausgebildet. Im Gegensatz dazu besitzt Versuchsperson WB4-1 bereits eine sehr gute Vorstellung bezüglich der relationalen Beziehungen zwischen den Störungsbildern. Die Clusterzugehörigkeit der Objekte wird nur von ca. drei Störungsbildern verletzt (F20.1, F25.0, F31.1). Die geringen Unterschiede lassen auf ein hohes Verständnis in die Zusammenhänge und Unterschiede schließen.

Homogenisierung der Wissenskarten. Neben dem Vergleich der einzelnen Personen mit dem Normmodell ist vor allem auch interessant, ob es im Verlauf der Weiterbildung zu einer Homogenisierung der Sichtweisen zwischen den Personen kommt. Zu erwarten ist ja nicht nur, dass sich die individuellen Wissenskarten mit zunehmendem Expertisegrad dem Normmodell annähern, sondern dass sich infolge des gemeinsam besuchten Curriculums auch die strukturellen Sichtweisen untereinander systematisch ähnlicher werden. Um das zu überprüfen, werden alle individuellen Karten durch Prokrustes-Transformation paarweise miteinander verglichen. Die daraus generierte Matrix der AvgLoss-Werte zwischen je zwei individuellen Karten werden wiederum mit RobuScal skaliert und als Personenkarte dargestellt (sog. Prokrustesbasierte Personenskalierung; ProperScal; Läge, 2001). Abbildung 72 zeigt die zweidimensionale Darstellung der relationalen Beziehungen zwischen den Personenkarten.

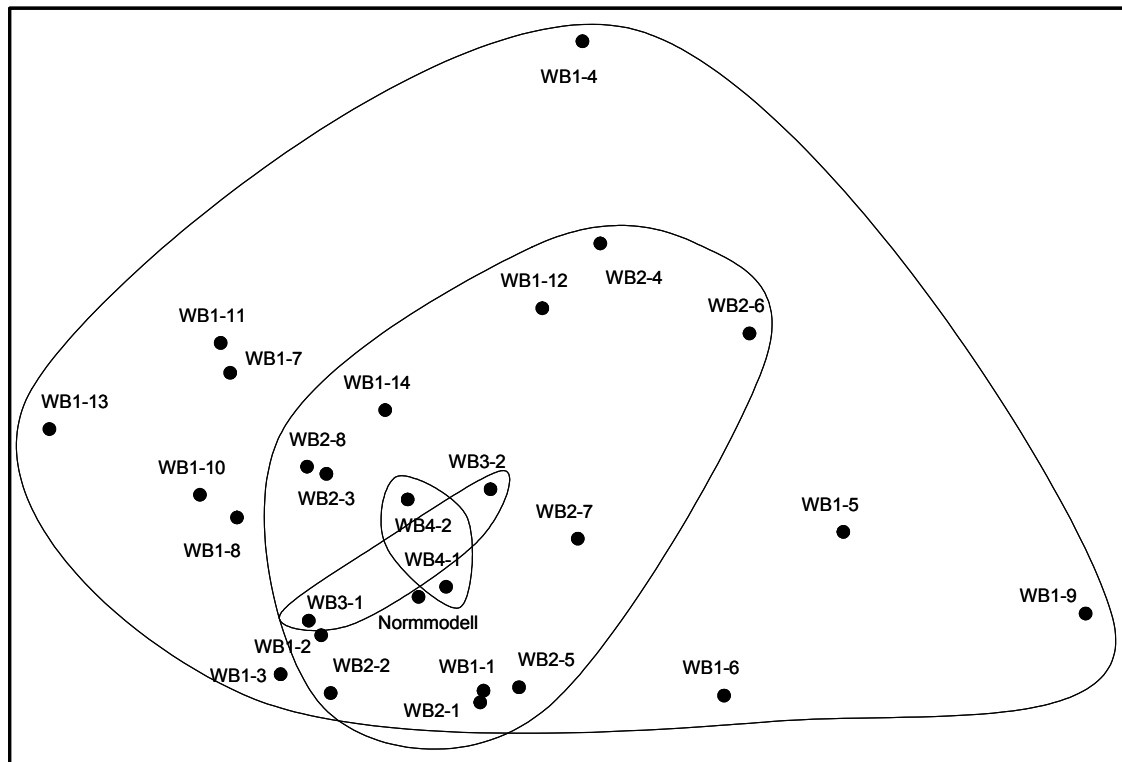


Abbildung 72: NMDS-Lösung aller individuellen Kartenpaarvergleiche (ProperScal). Die einzelnen Markierungen umfassen je die Teilnehmer der vier Stichproben (WB1-1 bis WB1-14; WB2-1 bis WB2-8; WB3-1 bis WB3-2; WB4-1 bis WB4-2). Ebenfalls eingetragen ist das Normmodell. Mittlerer AvgLoss aller Vergleiche = 0.52; Minimum AvgLoss = 0.27; Maximum AvgLoss = 0.80.

Deutlich zu erkennen ist die große Streuung der Kandidaten aus WB1. Mit zunehmender Weiterbildungszzeit verringert sich diese Heterogenität. Alle vier Personen aus WB3 und WB4 liegen schließlich recht nahe beieinander (und auch am Normmodell).

10.4. Diskussion

Diese Arbeit stellt einen innovativen Ansatz zur Messung individuellen strukturellen Sachwissens im Bereich psychopathologischer Störungsbilder vor. Die Verfahrenslogik der Visualisierung struktureller Wissenscharakteristika einer Person über die Nonmetrische Multidimensionale Skalierung und der Performanzabschätzung über den Vergleich einer Lernerkarte mit einem Expertenmodell (Prokrustes-Transformation) soll einen Beitrag zur Standardisierung und Objektivierung von Methoden liefern, welche im Zusammenhang mit der Evaluation der Ergebnisqualität eines psychotherapeutischen Curriculums stehen. Das Verfahren trägt den Namen „Zürcher Wissensstrukturmodell für Psychopathologie (ZWP)“.

Dass die NMDS merkmalsbasiertes Sachwissen zu messen und darzustellen sowie die Prokrustes-Transformation Wissensveränderungen über einen Lernverlauf aufzuzeigen in der Lage ist, konnte bereits in mehreren kontrollierten Lernexperimenten gezeigt werden (Läge, Rüfenacht, Fichter & Streule, eingereicht; Läge, Schlatter & Streule, eingereicht; Läge, Streule & Schlatter, eingereicht). In dieser Studie nun wurde nicht nur der Versuch unternommen, dieses Vorgehen auf das Inhaltsgebiet der psychopathologischen Störungsbilder auszuweiten, sondern zudem die Leistung der Methode im erstmaligen Einsatz in einem realen curricularen Umfeld auszuloten. Gezeigt werden konnten systematische strukturelle Wissensunterschiede zwischen den einzelnen Ausbildungsjahrgängen, und zwar ganz in der prognostizierten Ab-

hängigkeit vom Expertisegrad. Außerdem ließ sich die Hypothese bestätigen, dass zunehmende Ausbildungszeit zu einer Homogenisierung der individuellen Sichtweisen führt. Dies bedeutet eine hinreichende Verfahrensleistung zur Wissensevaluation und ermöglicht den Ausbau von ZWP für den Praxiseinsatz als standardisiertes Messinstrument.

Die Messung der Ergebnisqualität eines Curriculums mit Hilfe von ZWP geschieht auf zwei Ebenen. Unter der Voraussetzung eines validen Vergleichskriteriums (Expertenmodell) können zum einen interindividuelle Stichprobenvergleiche helfen, den augenblicklichen Stand eines Ausbildungsjahrgangs festzuhalten. Schneidet beispielsweise ein fortgeschrittener Jahrgang im Mittel schlechter ab als ein jüngerer, so sollte entsprechend nach Ursachen gesucht und darauf reagiert werden. Sind die Wissenskarten einzelner Personen besonders auffällig, dann sind spezielle Förderungsmassnahmen im Aufbau des deklarativen Wissens angezeigt.

In einer Erweiterung der hier vorgestellten Auswertung der individuellen Wissenskarten kann über die Analyse der einzelnen Objektabweichungen Außerdem festgestellt werden, mit welchen Störungsbildern die betreffende Person besondere Schwierigkeiten bei der relationalen Einordnung hat. Läge, Rüfenacht, Fichter & Streule (eingereicht) evaluierten die Verfahrenslogik von ZWP auf genau dieser Ebene. Nicht allein das Gesamtmaß AvgLoss, sondern die einzelnen Abweichungen der Objektpositionen einer Lernerkarte (ObjLoss) wurden auf ihre Vorhersagegenauigkeit bezüglich der Wissensqualität hin untersucht. Unter der Voraussetzung eines gewissen Wissensniveaus ist die Prokrustes-Transformation sogar in der Lage, auf Objektebene Wissensmängel (Lücken, Falschwissen) zu detektieren. Auf der Basis dieses Vergleichs könnte neben der verstärkten Behandlung einzelner schwieriger Störungsbilder im Plenum zudem adaptiv auf den Wissensstand eines Lerner eingegangen werden. Dies steigert die Effizienz eines Curriculums, weil somit Mehrfachlernen bereits gewussten Materials praktisch komplett vermieden werden kann, indem man den Lerner dort abholt, wo er sich gerade befindet (vgl. Ausubel, 1963 und 2000; sowie zur Adaptivität von Lernprogrammen Leutner, 1992).

Die NMDS integriert ebenso einen didaktischen Aspekt. Über die Betrachtung eigener Wissenskarten und deren Abweichung zur Normkarte entsteht ein „externes“ Feedback über die individuellen Gedächtnisinhalte. Eine bewusste und aktive Beschäftigung mit den möglichen Ursachen von Kriteriumsabweichungen fördert den Lernprozess dahingehend, dass explizit nach Gemeinsamkeiten und Unterschieden gesucht wird und somit Stabilität und Differenziertheit des eigenen Wissens stark zunimmt. Dies kann sowohl in Eigenregie (selbstgesteuert) als auch gemeinsam mit Dozenten oder Kommilitonen geschehen. Die Ausbildung komplexer struktureller Facetten in dem für alles weitere fundamentale deklarative Wissen führt zu einer optimalen Vorbereitung des angehenden Therapeuten auf die zu erlernende praktische Tätigkeit.

Der Einsatz der NMDS zum Zwecke der Wissensmessung stellt unserer Ansicht nach einen innovativen und bislang einzigartigen Versuch dar, einen genau spezifizierten Teil der Ergebnisqualität eines psychotherapeutischen Curriculums über relationale Facetten von Gedächtnisinhalten zu überprüfen. Der im Vergleich zu herkömmlichen Prüfungsformen (z.B. Multiple-Choice, Forced-Choice) zusätzliche Nutzen dieses Verfahrens liegt in der Ausdrücklichkeit von Zusammenhängen. Expertenwissen zeichnet sich nicht allein durch eine große Wissensmenge aus, sondern lässt sich vielmehr anhand komplexer und differenzierter Wissenscharakteristiken erkennen (Gewichtungen, Prototypikalitäten, Auswahl der relevanten Beurteilungsdimensionen usw.). Das Bewusstsein bezüglich feiner Unterschiede und Gemeinsamkeiten ermöglicht die Grundlage für ein elaboriertes und adäquates Verhalten. Vor allem in einem sich in Entwicklung befindlichen und nicht kategorial strukturierten Inhaltsgebiet wie der Psychopathologie ist dies wichtig. Die sich vielfach überschneidende Phänomenologie mehrerer Störungen innerhalb eines Patienten benötigt ein gleichzeitiges Aufrechterhalten verschie-

dener Hypothesen. Nur dann ist der Therapeut in der Lage, das bestmögliche Therapieprogramm zu erstellen.

Wie bereits erwähnt, steht und fällt die Verfahrenslogik von ZWP mit der Qualität des Normmodells. Die Auswahl der zu beurteilenden Störungsbilder sollte dazu an die vermittelten Inhalte des Curriculums angepasst sein. Dieser Punkt ist sicherlich an der vorliegenden Pilotstudie zu bemängeln. Dass das zentrale Testkriterium der Annäherung der Lernerkarten an ein Normmodell mit zunehmender Expertise trotzdem bestätigt werden konnte, stimmt im Hinblick auf einen möglichen Praxiseinsatz von ZWP zuversichtlich. Es scheint möglich, damit ein zu den bereits bestehenden Verfahren der Evaluation komplementäres und vielseitig einsetzbares Evaluationswerkzeug für die relationale Komponente des deklarativen therapeutischen Wissens zu entwickeln.

10.5. Literaturverzeichnis

- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune & Stratton.
- Ausubel, D. P. (2000). *The acquisition and retention of knowledge: A cognitive view*. Dordrecht: Kluwer Academic Publishers.
- Beitman, B. D. & Yue, D. (1999). A new psychotherapy training program: Description and preliminary results. *Academic Psychiatry*, 23 (2), 95-102.
- Borg, I. & Groenen, P. (1997). *Modern multidimensional scaling – Theory and applications*. New York: Springer.
- Buchkremer, G. & Schneider, F. (1995). Qualitätssicherung in der Psychotherapie. In H.-J. Haug & R.-D. Stieglitz (Hrsg.), *Qualitätssicherung in der Psychiatrie*. Stuttgart: Ferdinand Enke Verlag. S. 141-150.
- Eckert, H. & Westhoff, K. (2000). Psychologiedidaktik und Evaluation am Beispiel der universitären Ausbildung im Fach "Diagnostik und Intervention". In G. Krampen & H. Zayer (Hrsg.), *Psychologiedidaktik und Evaluation II. Neue Medien, Psychologiedidaktik und Evaluation in der psychologischen Haupt- und Nebenfachausbildung*. Bonn: Deutscher Psychologen Verlag GmbH. S. 143-158.
- Egli, S., Schlatter, K., Streule, R. & Läge, D. (submitted). *A structure based expert model of the ICD-10 mental disorders*.
- Ehlert, U. & Kaiser, N. (2001). Curriculum der Ausbildung in Verhaltensmedizin. *Praxis Klinische Verhaltensmedizin und Rehabilitation*, 55, 199-202.
- Elizur, A., Kretsch, R., Spaizer, N. & Sorek, Y. (1994). Self-evaluation of psychotherapeutic competence. *British Journal of Medical Psychology*, 67 (3), 231-235.
- Foulkes, P. (2003). Trainee perceptions of teaching of different psychotherapies. *Australasian Psychiatry*, 11 (2), 209-214.
- Freedheim, D. K. & Overholser, J. C. (1998). Training in psychotherapy during graduate school. *Psychotherapy in Private Practice*, 17 (1), 3-18.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. München: Reinhardt.
- Gower, J. C. & Dijksterhuis, G. B. (2004). *Procrustes Problems*. New York: Oxford University Press.
- Jagsch, R., Kryspin-Exner, I. & Jandl-Jager, E. (2002). Selbsteinschätzung der Kompetenzen von Absolventen des psychotherapeutischen Propädeutikums. *Psychotherapie-Forum*, 10 (4), 214-222.
- Kaiser, N. & Ehlert, U. (2003). Weiterbildungsmöglichkeiten in der Verhaltensmedizin. *Psychotherapeut*, 48, 361-367.

- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen*. Habilitationsschrift an der Philosophischen Fakultät der Universität Zürich.
- Läge, D., Daub, S., Bosia, L., Ryf, S. & Jäger, C. (in review). *Robustness in nonmetric multidimensional scaling*.
- Läge, D., Rüfenacht, I., Fichter, C. & Streule, R. (eingereicht). *Der ObjectLoss einer Prokrustes-Transformation als Maß für Wissen*.
- Läge, D., Schlatter, K. & Streule, R. (eingereicht). *Der NMDS-Stresswert als Maß für Assimilation und Akkomodation beim Lernen von Wissensstrukturen*.
- Läge, D., Streule, R. & Schlatter, K. (eingereicht). *Die Veränderung von Wissenskarten beim Merkmalslernen*.
- Leutner, D. (1992). *Adaptive Lernsysteme. Instruktionspsychologisches Grundlagen und experimentelle Analysen*. Weinheim: Psychologie Verlags Union.
- Mandl, H. & Fischer, F. (Hrsg.). (2000). *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken*. Göttingen: Hogrefe.
- Milne, D. L., Baker, C., Blackburn, I.-M., James, I. & Reichelt, K. (1999). *Effectiveness of cognitive therapy training*. Journal of Behavior Therapy and Experimental Psychiatry, 30, 81-92.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Reinecker, H. & Schindler, L. (2000). Aus- und Weiterbildung. In J. Margraf (Hrsg.), *Lehrbuch der Verhaltenstherapie – Band 1 (2. vollst. über. & erw. Aufl.)*. Berlin: Springer. S. 655-662.
- Schulz, W. (2000). *Deutlich mehr Kompetenz. Erste Ergebnisse der Dokumentation und Evaluation der „Weiterbildung zum Sozialtherapeuten klientenzentriert/gesprächspsychotherapeutisch orientiert“*. Gesprächspsychotherapie und Personenzentrierte Beratung, 31 (2), 146-151.
- Schulz, W., Flotho, B., Lübke, A. & Hahlweg, K. (1998). *Dokumentation und Evaluation des Weiterbildungsganges Braunschweig in Klinischer Psychologie/Psychotherapie*. Verhaltenstherapie, 8 (4), 234-244.
- Willutzki, U. & Botermans, J.-F. (1997). Ausbildung in Psychotherapie in Deutschland und der Schweiz und ihre Bedeutung für die therapeutische Kompetenz. Psychotherapeut, 42, 282-289.

Lebenslauf Katharina Schlatter

Persönliche Informationen

Staatsangehörigkeit: Schweiz
Geburtsdatum: 13.03.1973
Bürgerorte: Elfingen AG und Buchs ZH
Geburtsort: Brugg AG
e-mail: katharina.schlatter@gmx.ch

Ausbildung

- | | |
|-------------------|--|
| 06.2002 – 06.2005 | Promotionsstudium (Universität Zürich)
Dissertation Juni 2005. Titel: Adaptives Lernen mit Wissensstrukturkarten – Mess- und Skalierungsmodelle zur Modellierung von klassifikatorischem Expertenwissen auf dem Prüfstand |
| 08.2003 | Kurs “Multidimensional Scaling”, Summer School on Advanced Methods in the Social Sciences 2003, Lugano |
| 10.1996 – 06.2002 | Studium der Psychologie (Universität Zürich), Vertiefungsrichtung Allgemeine Psychologie, Nebenfächer: Psychopathologie für das Kindes- und Jugendalter, Neurophysiologie |
| 10.1994 – 05.1996 | Biologie (3 Semester), Universität Zürich |
| 08.1990 – 06.1994 | Kantonsschule Wettingen (Matura Typus PSG) |

Berufserfahrung

- | | |
|-------------------|--|
| 06.2002 – 12.2004 | Projektmitarbeiterin an Schweizerischem Nationalfondsprojekt („Merkmalsorientierte Diagnostik in adaptiven Lernsystemen“) und Doktorandin, Abteilung für Angewandte Kognitionspsychologie, Universität Zürich |
| 05.2001 – 05.2002 | Assistentin am Psychologischen Institut der Universität Zürich, Abteilung Allgemeine Psychologie, Datenerhebung und Auswertung |
| 09.2001 – 11.2001 | Praktikum am Kinder- und Jugendpsychiatrischen Dienst, Baden |
| 1999 - 2001 | Semesterassistentin am Psychologischen Institut der Universität Zürich (4Semester), Abteilung Allgemeine Psychologie |
| 1994 – 2001 | Temporäranstellungen während dem Studium: Telesales, Verkauf von Upgrades (NOVELL Schweiz, Zürich), Schmuckverkauf (DEVINE, Globus Zürich); Kinderkrippe (Chinderhuus, Zürich-Affoltern), Gebäude Reinigung (Waidspital, Zürich) |